DocCode – SCORE

SCORE Placeholder Sheet for IFW Content

Application Number: 16272487 Document Date: 02/11/2019

The presence of this form in the IFW record indicates that the following document type was received in electronic format on the date identified above. This content is stored in the SCORE database.

Since this was an electronic submission, there is no physical artifact folder, no artifact folder is recorded in PALM, and no paper documents or physical media exist. The TIFF images in the IFW record were created from the original documents that are stored in SCORE.

- Drawing

At the time of document entry (noted above):

- USPTO employees may access SCORE content via eDAN using the Supplemental Content tab, or via the SCORE web page.
- External customers may access SCORE content via PAIR using the Supplemental Content tab.

Form Revision Date: August 26, 2013

10x Genomics EX1012 (Part 1 of 2)

DocCode – SEQ.TXT

SCORE Placeholder Sheet for IFW Content

Application Number: 16272487 Document Date: 02/11/2019

The presence of this form in the IFW record indicates that the following document type was received in electronic format on the date identified above. This content is stored in the SCORE database.

Since this was an electronic submission, there is no physical artifact folder, no artifact folder is recorded in PALM, and no paper documents or physical media exist. The TIFF images in the IFW record were created from the original documents that are stored in SCORE.

- Sequence Listing

At the time of document entry (noted above):

- USPTO employees may access SCORE content via eDAN using the Supplemental Content tab, or via the SCORE web page.
- External customers may access SCORE content via PAIR using the Supplemental Content tab.

Form Revision Date: August 26, 2013

PTO/AIA/15 (10-17)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995 no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	NATE-037/001US 321329-2523
First Named Inventor	Joseph M. BEECHEM
Title	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION
Priority Mail Express® Label No.	

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. [ ] **Fee Transmittal Form** (PTO/SB/17 or equivalent)
2. [ ] **Applicant asserts small entity status.** See 37 CFR 1.27
3. [ ] **Applicant certifies micro entity status.** See 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.
4. [X] **Specification** [*Total Pages* 109]
   Both the claims and abstract must start on a new page.
   (See MPEP § 608.01(a) for information on the preferred arrangement)
5. [X] **Drawing(s)** (35 U.S.C. 113) [*Total Sheets* 34]
6. **Inventor's Oath or Declaration** [*Total Pages*]
   (including substitute statements under 37 CFR 1.64 and assignments serving as an oath or declaration under 37 CFR 1.63(e))
   a. [ ] Newly executed (original or copy)
   b. [ ] A copy from a prior application (37 CFR 1.63(d))
7. [X] **Application Data Sheet** ** See note below.*
   See 37 CFR 1.76 (PTO/AIA/14 or equivalent)
8. **CD-ROM or CD-R** in duplicate, large table, or Computer Program (*Appendix*)
   [ ] Landscape Table on CD
9. **Nucleotide and/or Amino Acid Sequence Submission** (*if applicable, items a. – c. are required*)
   a. [X] Computer Readable Form (CRF)
   b. [ ] Specification Sequence Listing on:
      i. [ ] CD-ROM or CD-R (2 copies); or
      ii. [ ] Paper
   c. [X] Statements verifying identity of above copies

ADDRESS TO:
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

ACCOMPANYING APPLICATION PAPERS

10. [ ] **Assignment Papers** (cover sheet & document(s))
    Name of Assignee
11. [ ] **37 CFR 3.73(c) Statement** (*when there is an assignee*) [X] **Power of Attorney**
12. [ ] **English Translation Document** (*if applicable*)
13. [ ] **Information Disclosure Statement** (PTO/SB/08 or PTO-1449)
    [ ] Copies of citations attached
14. [ ] **Preliminary Amendment**
15. [ ] **Return Receipt Postcard** (*MPEP § 503*) (*Should be specifically itemized*)
16. [ ] **Certified Copy of Priority Document(s)** (*if foreign priority is claimed*)
17. [ ] **Nonpublication Request** Under 35 U.S.C. 122(b)(2)(B)(i). Applicant must attach form PTO/SB/35 or equivalent.
18. [X] **Other:** Petition under 37 CFR 1.84(a)(2) to Accept Color Drawings

***Note:** (1) Benefit claims under 37 CFR 1.78 and foreign priority claims under 1.55 **must** be included in an Application Data Sheet (ADS).
(2) For applications filed under 35 U.S.C. 111, the application must contain an ADS specifying the applicant if the applicant is an assignee, person to whom the inventor is under an obligation to assign, or person who otherwise shows sufficient proprietary interest in the matter. See 37 CFR 1.46(b).

19. CORRESPONDENCE ADDRESS

[X] The address associated with Customer Number: 141429 **OR** [ ] Correspondence address below

Name					
Address					
City		State		Zip Code	
Country		Telephone		Email	

Signature	/Matthew Pavao/	Date	February 11, 2019
Name (Print/Type)	Matthew Pavao, Ph.D.	Registration No. (Attorney/Agent)	50,572

198895705 v1

Doc Code: PA..
Document Description: Power of Attorney

PTO/AIA/82B (07-13)
Approved for use through 11/30/2014. OMB 0651-0051
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

POWER OF ATTORNEY BY APPLICANT

I hereby revoke all previous powers of attorney given in the application identified in either the attached transmittal letter or the boxes below.

Application Number	Filing Date

(Note: The boxes above may be left blank if information is provided on form PTO/AIA/82A.)

☑ I hereby appoint the Patent Practitioner(s) associated with the following Customer Number as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above: 141429

OR

☐ I hereby appoint Practitioner(s) named in the attached list (form PTO/AIA/82C) as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the patent application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above. (Note: Complete form PTO/AIA/82C.)

Please recognize or change the correspondence address for the application identified in the attached transmittal letter or the boxes above to:

☑ The address associated with the above-mentioned Customer Number

OR

☐ The address associated with Customer Number:

OR

Firm or Individual Name					
Address					
City		State		Zip	
Country					
Telephone		Email			

I am the Applicant (if the Applicant is a juristic entity, list the Applicant name in the box):

NanoString Technologies, Inc.

☐ Inventor or Joint Inventor (title not required below)

☐ Legal Representative of a Deceased or Legally Incapacitated Inventor (title not required below)

☑ Assignee or Person to Whom the Inventor is Under an Obligation to Assign (provide signer's title if applicant is a juristic entity)

☐ Person Who Otherwise Shows Sufficient Proprietary Interest (e.g., a petition under 37 CFR 1.46(b)(2) was granted in the application or is concurrently being filed with this document) (provide signer's title if applicant is a juristic entity)

SIGNATURE of Applicant for Patent

The undersigned (whose title is supplied below) is authorized to act on behalf of the applicant (e.g., where the applicant is a juristic entity).

Signature	/Andrew J. Leon/	Date (Optional)	June 9, 2016
Name	Andrew J. Leon		
Title	Director of Intellectual Property, NanoString Technologies, Inc.		

NOTE: Signature - This form must be signed by the applicant in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. If more than one applicant, use multiple forms.

☑ Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.131, 1.32, and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of:	Joseph M. Beechem *et al*	Confirmation No.:	TBD
Application No.:	TBD	Group Art Unit:	TBD
Filed:	HEREWITH	Examiner:	TBD

FOR: **BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION**

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

STATEMENT IN SUPPORT OF COMPUTER READABLE FORM SUBMISSION UNDER 37 C.F.R. § 1.821

I hereby state that the contents of the computer readable form of the Sequence Listing, submitted in the above-identified application in accordance with 37 C.F.R. § 1.821, does not include any new matter that goes beyond the disclosure of the application as filed. The Sequence Listing is supported by the specification and references incorporated therein. Therefore, no new matter is added.

Dated: February 11, 2019

Respectfully submitted,
COOLEY LLP

Cooley LLP
ATTN: Patent Group
1299 Pennsylvania Avenue NW, Suite 700
Washington, DC 20004

By: / Robert E. Powers /
Robert E. Powers, Ph.D.

Tel: (617) 937-2490
Fax: (202) 842-7899

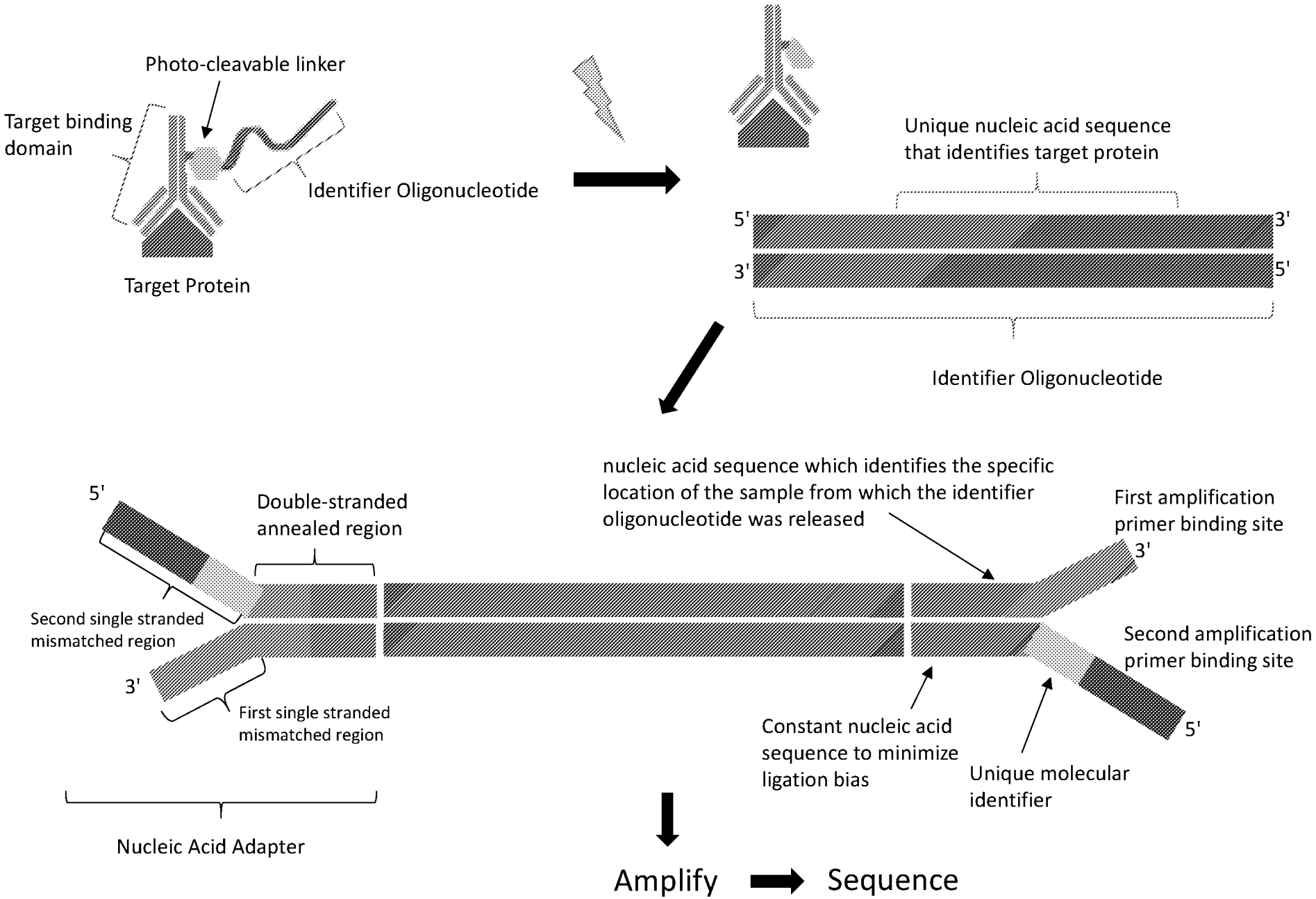

Photo-cleavable linker
Target binding domain
Identifier Oligonucleotide
Target Protein
Unique nucleic acid sequence that identifies target protein
5'
3'
3'
5'
Identifier Oligonucleotide
5'
Double-stranded annealed region
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
First amplification primer binding site
3'
Second single stranded mismatched region
Second amplification primer binding site
3'
First single stranded mismatched region
Constant nucleic acid sequence to minimize ligation bias
Unique molecular identifier
5'
Nucleic Acid Adapter
Amplify
Sequence

Figure 1

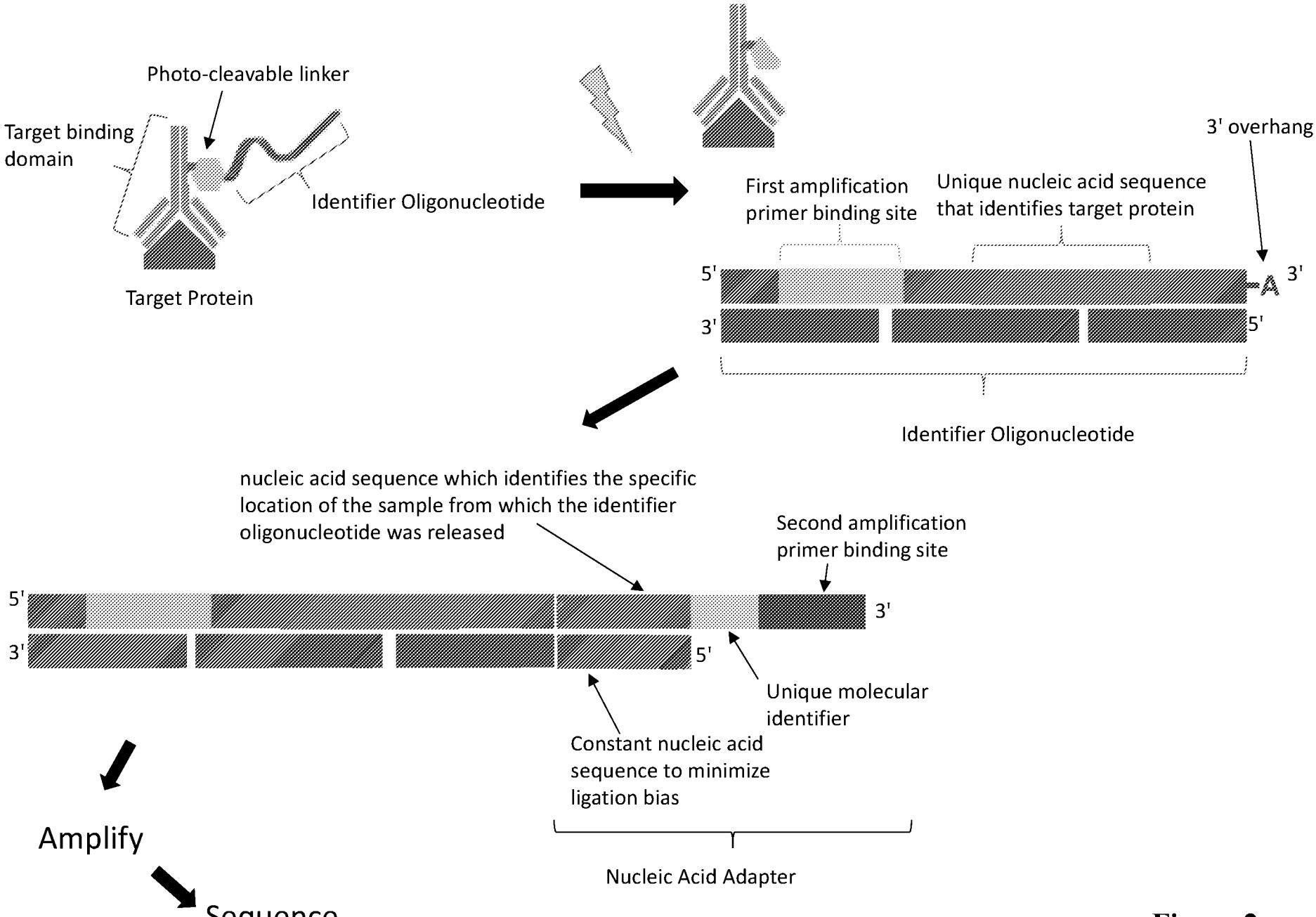

Photo-cleavable linker
Target binding domain
Identifier Oligonucleotide
Target Protein
3' overhang
First amplification primer binding site
Unique nucleic acid sequence that identifies target protein
5'
3'
A
3'
5'
Identifier Oligonucleotide
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Second amplification primer binding site
5'
3'
3'
5'
Unique molecular identifier
Constant nucleic acid sequence to minimize ligation bias
Nucleic Acid Adapter
Amplify
Sequence

Figure 2

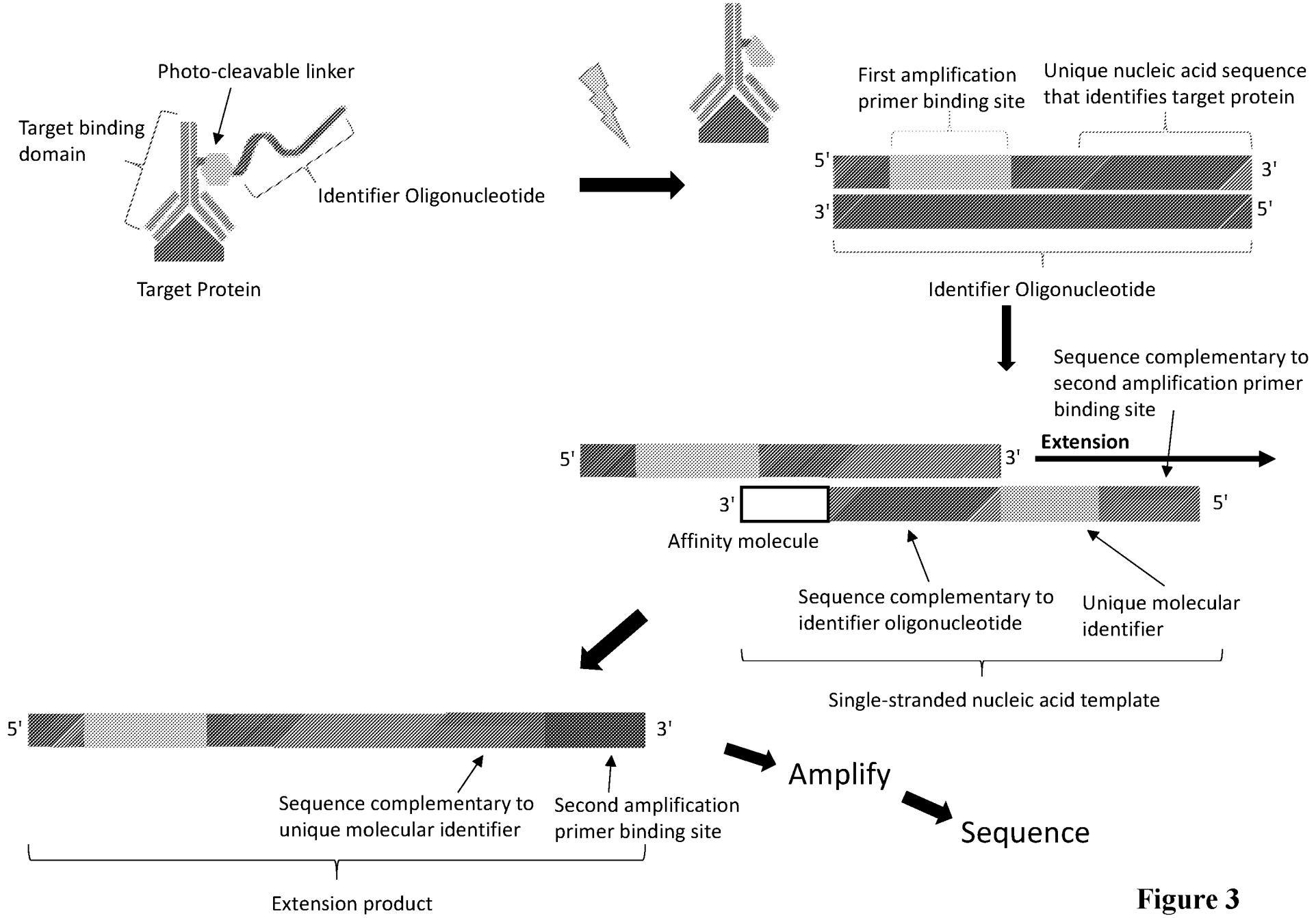

Photo-cleavable linker
Target binding domain
Identifier Oligonucleotide
Target Protein
First amplification primer binding site
Unique nucleic acid sequence that identifies target protein
5'
3'
3'
5'
Identifier Oligonucleotide
Sequence complementary to second amplification primer binding site
Extension
5'
3'
3'
5'
Affinity molecule
Sequence complementary to identifier oligonucleotide
Unique molecular identifier
Single-stranded nucleic acid template
5'
3'
Sequence complementary to unique molecular identifier
Second amplification primer binding site
Extension product
Amplify
Sequence

Figure 3

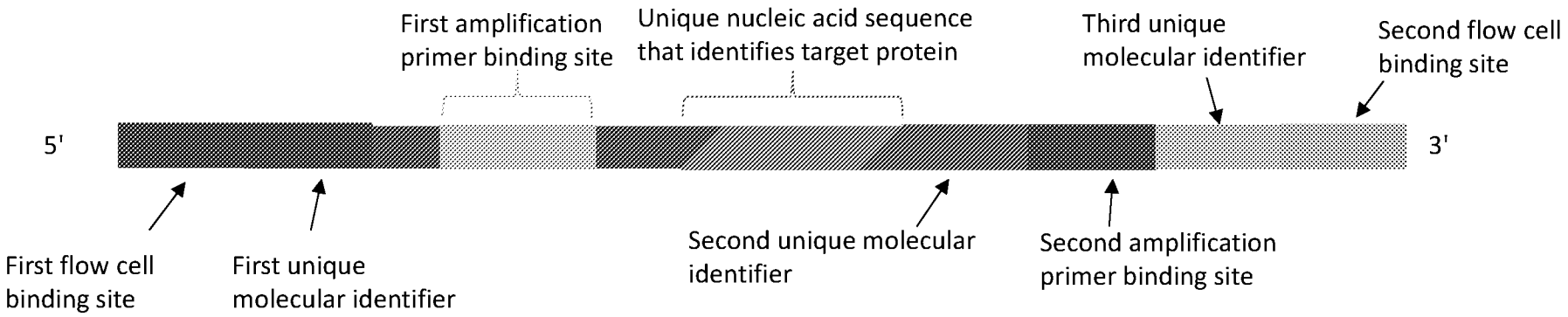

First amplification primer binding site
Unique nucleic acid sequence that identifies target protein
Third unique molecular identifier
Second flow cell binding site
5'
3'
First flow cell binding site
First unique molecular identifier
Second unique molecular identifier
Second amplification primer binding site

Figure 4

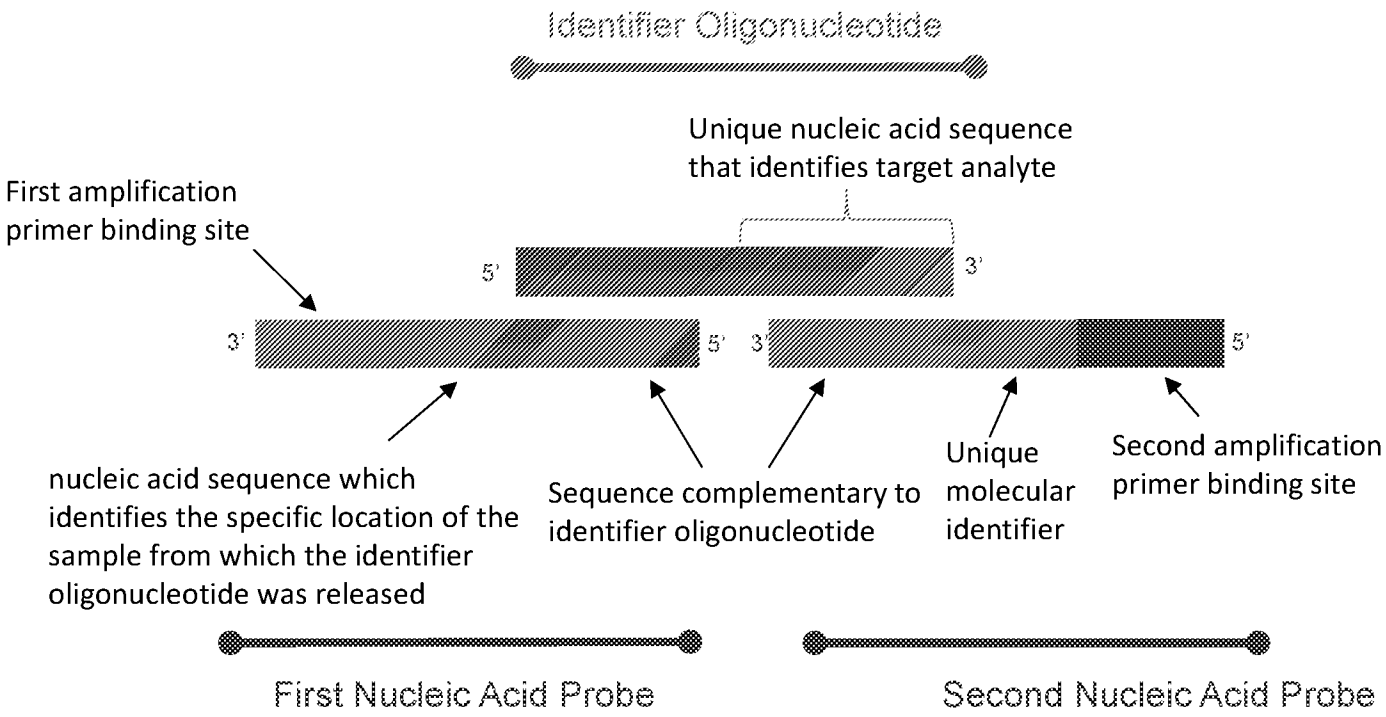

Identifier Oligonucleotide
Unique nucleic acid sequence that identifies target analyte
First amplification primer binding site
5'
3'
3'
5'
3'
5'
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Sequence complementary to identifier oligonucleotide
Unique molecular identifier
Second amplification primer binding site
First Nucleic Acid Probe
Second Nucleic Acid Probe

Figure 5

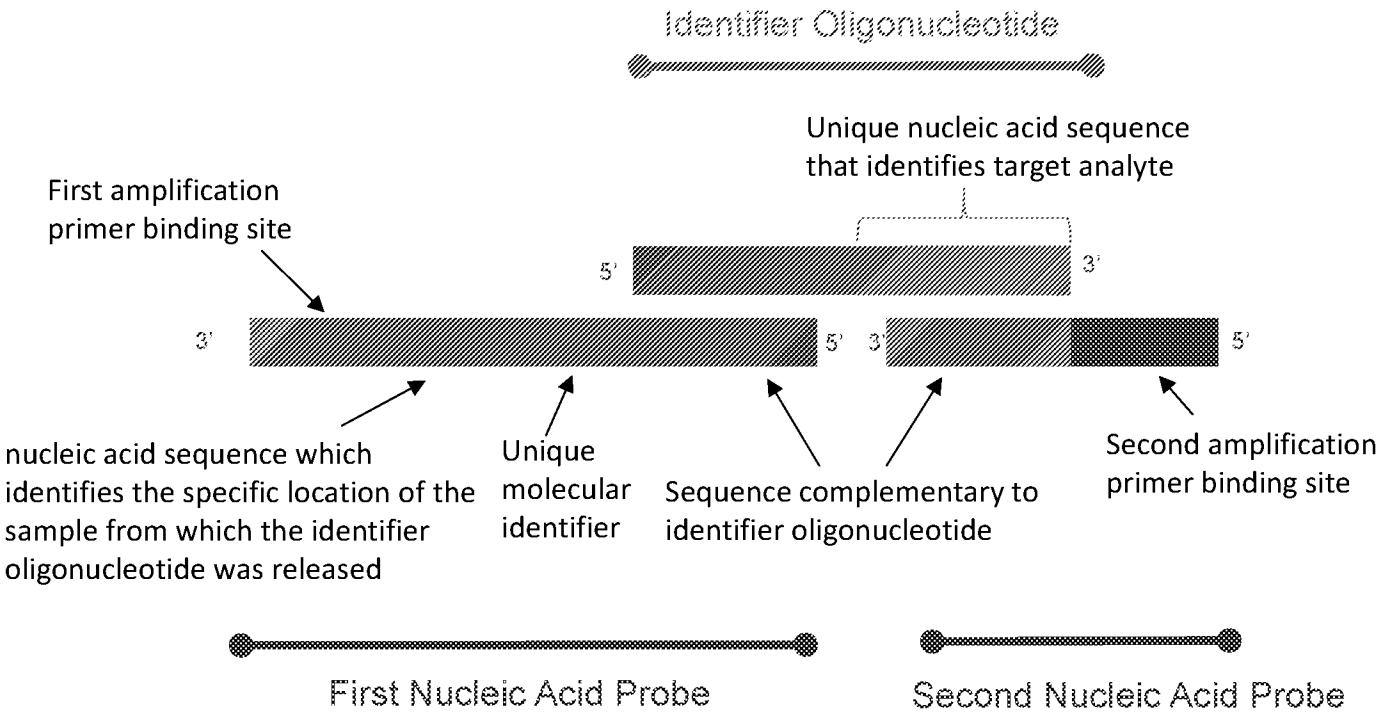

Identifier Oligonucleotide
Unique nucleic acid sequence that identifies target analyte
First amplification primer binding site
5'
3'
3'
5'
3'
5'
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Unique molecular identifier
Sequence complementary to identifier oligonucleotide
Second amplification primer binding site
First Nucleic Acid Probe
Second Nucleic Acid Probe

Figure 6

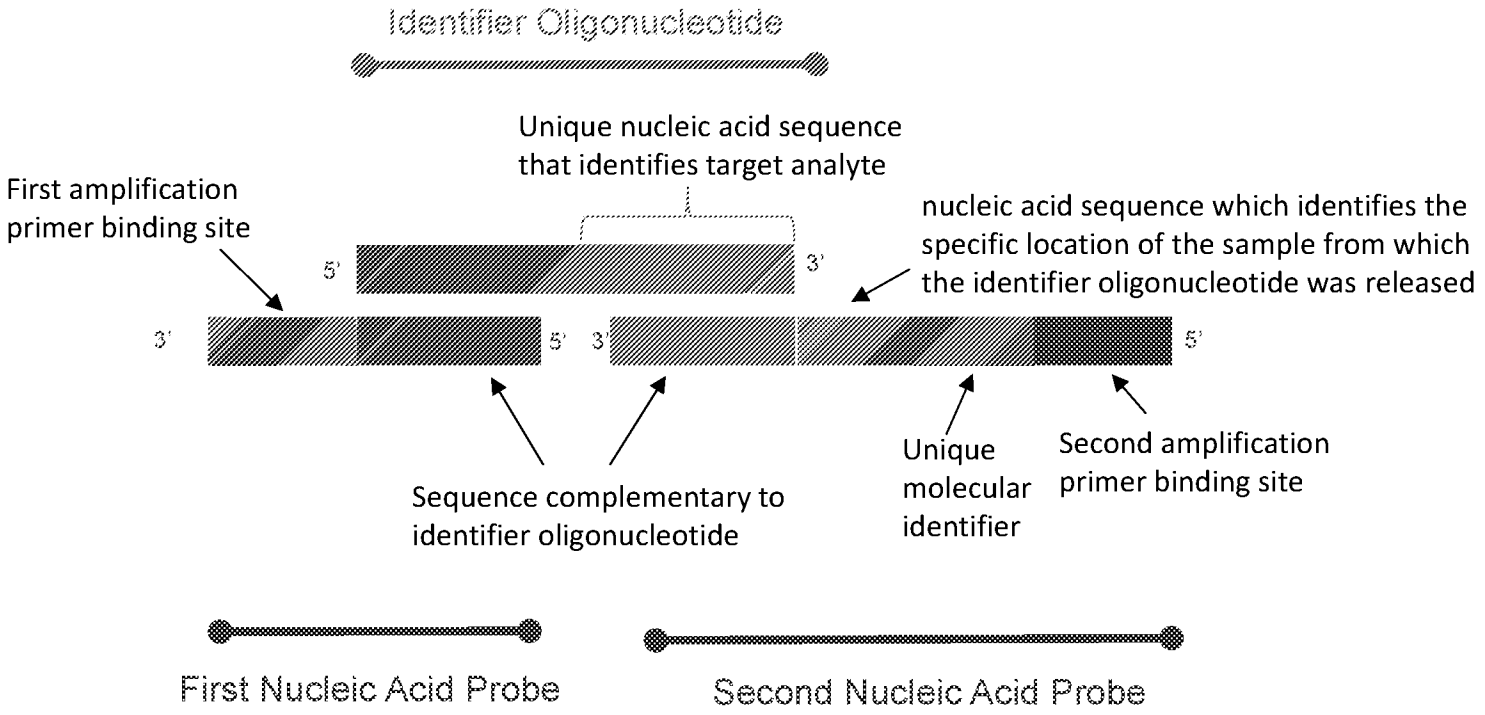

Identifier Oligonucleotide
Unique nucleic acid sequence that identifies target analyte
First amplification primer binding site
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
5'
3'
3'
5'
3'
5'
Sequence complementary to identifier oligonucleotide
Unique molecular identifier
Second amplification primer binding site
First Nucleic Acid Probe
Second Nucleic Acid Probe

Figure 7

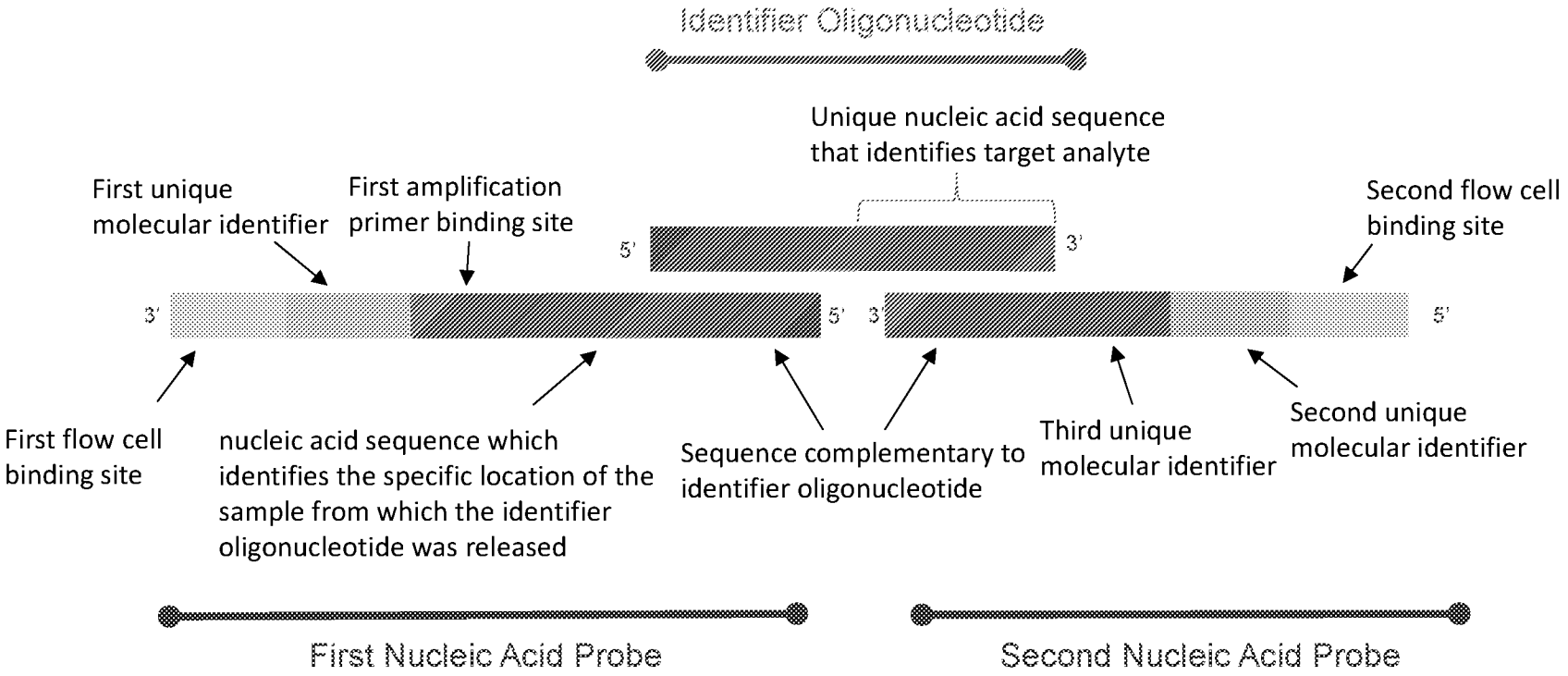

Identifier Oligonucleotide
Unique nucleic acid sequence that identifies target analyte
First unique molecular identifier
First amplification primer binding site
Second flow cell binding site
5'
3'
3'
5'
3
5'
First flow cell binding site
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Sequence complementary to identifier oligonucleotide
Third unique molecular identifier
Second unique molecular identifier
First Nucleic Acid Probe
Second Nucleic Acid Probe

Figure 8

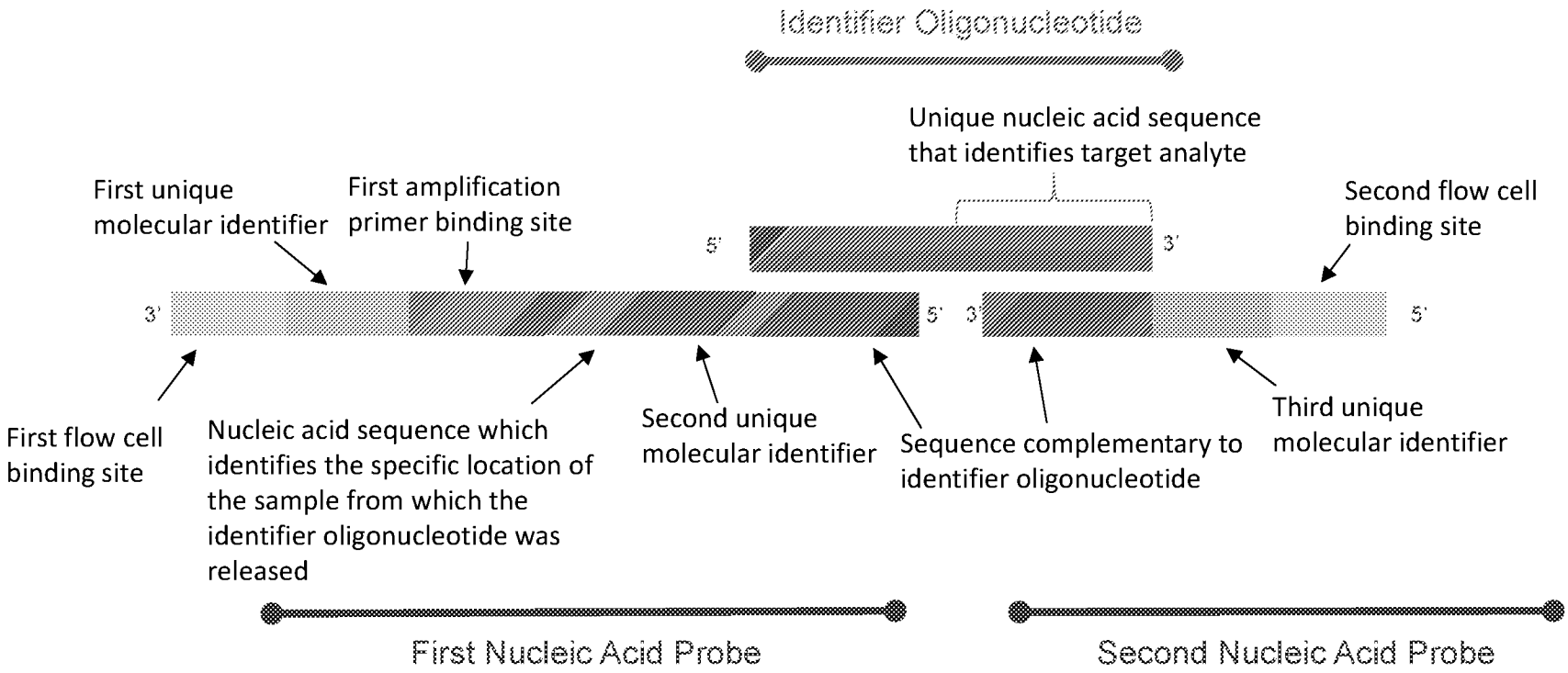

Identifier Oligonucleotide
Unique nucleic acid sequence that identifies target analyte
First unique molecular identifier
First amplification primer binding site
Second flow cell binding site
5'
3'
3'
5'
3'
5'
First flow cell binding site
Nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Second unique molecular identifier
Sequence complementary to identifier oligonucleotide
Third unique molecular identifier
First Nucleic Acid Probe
Second Nucleic Acid Probe

Figure 9

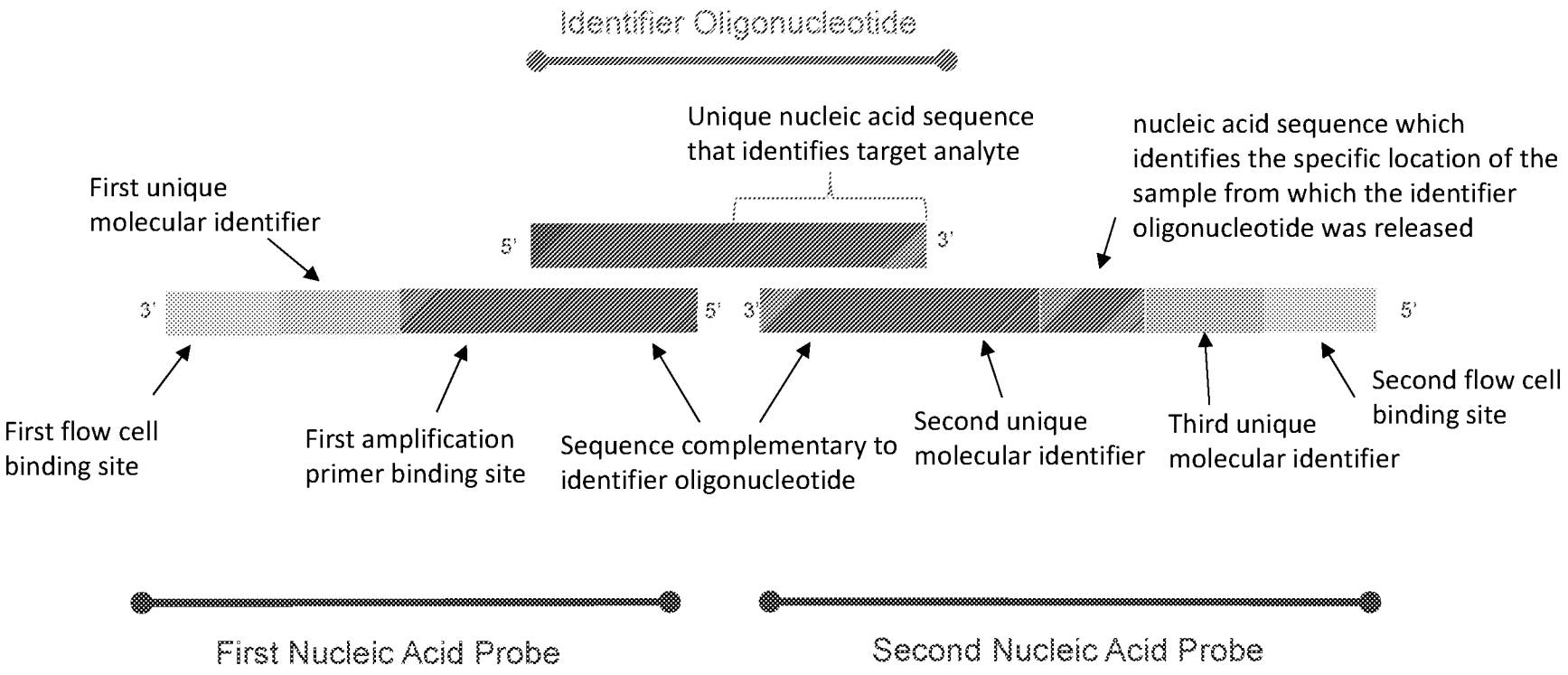

Identifier Oligonucleotide
Unique nucleic acid sequence that identifies target analyte
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
First unique molecular identifier
5'
3'
3'
5'
3'
5'
First flow cell binding site
First amplification primer binding site
Sequence complementary to identifier oligonucleotide
Second unique molecular identifier
Third unique molecular identifier
Second flow cell binding site
First Nucleic Acid Probe
Second Nucleic Acid Probe

Figure 10

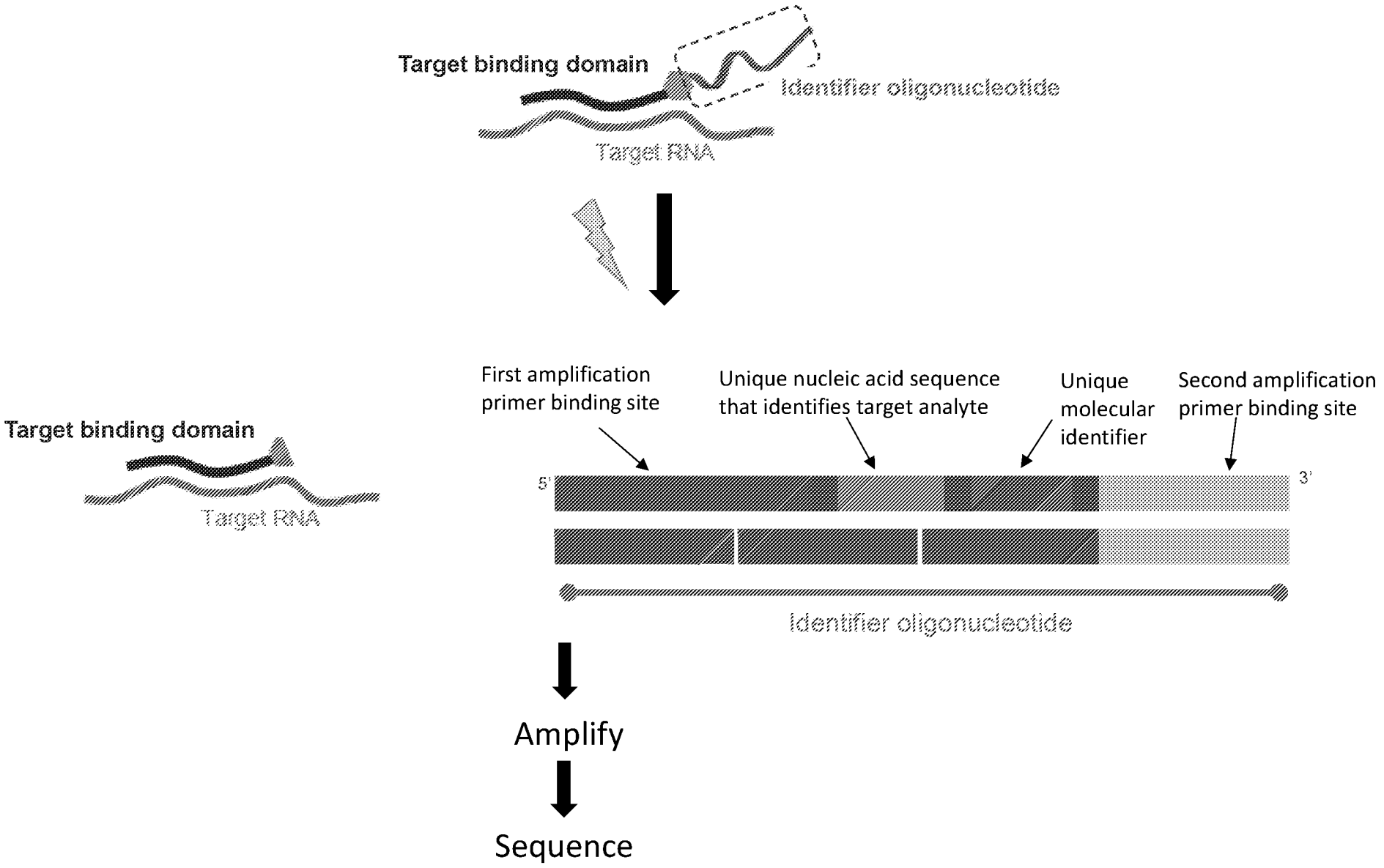

Target binding domain
Identifier oligonucleotide
Target RNA
Target binding domain
Target RNA
First amplification primer binding site
Unique nucleic acid sequence that identifies target analyte
Unique molecular identifier
Second amplification primer binding site
5'
3'
Identifier oligonucleotide
Amplify
Sequence

Figure 11

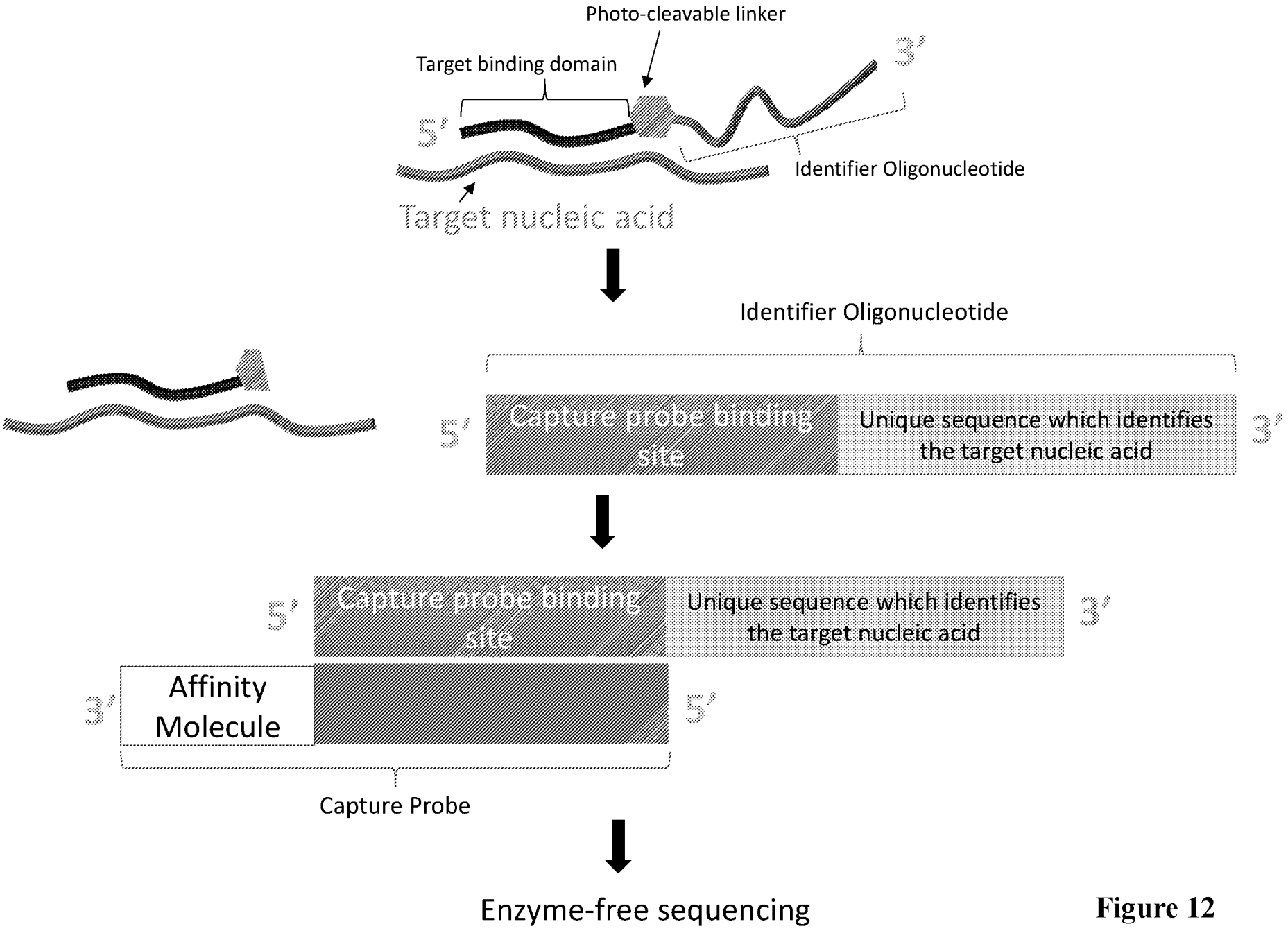

Photo-cleavable linker
Target binding domain
3′
5′
Identifier Oligonucleotide
Target nucleic acid
Identifier Oligonucleotide
5′
Capture probe binding site
Unique sequence which identifies the target nucleic acid
3′
5′
Capture probe binding site
Unique sequence which identifies the target nucleic acid
3′
3′
Affinity Molecule
5′
Capture Probe
Enzyme-free sequencing

Figure 12

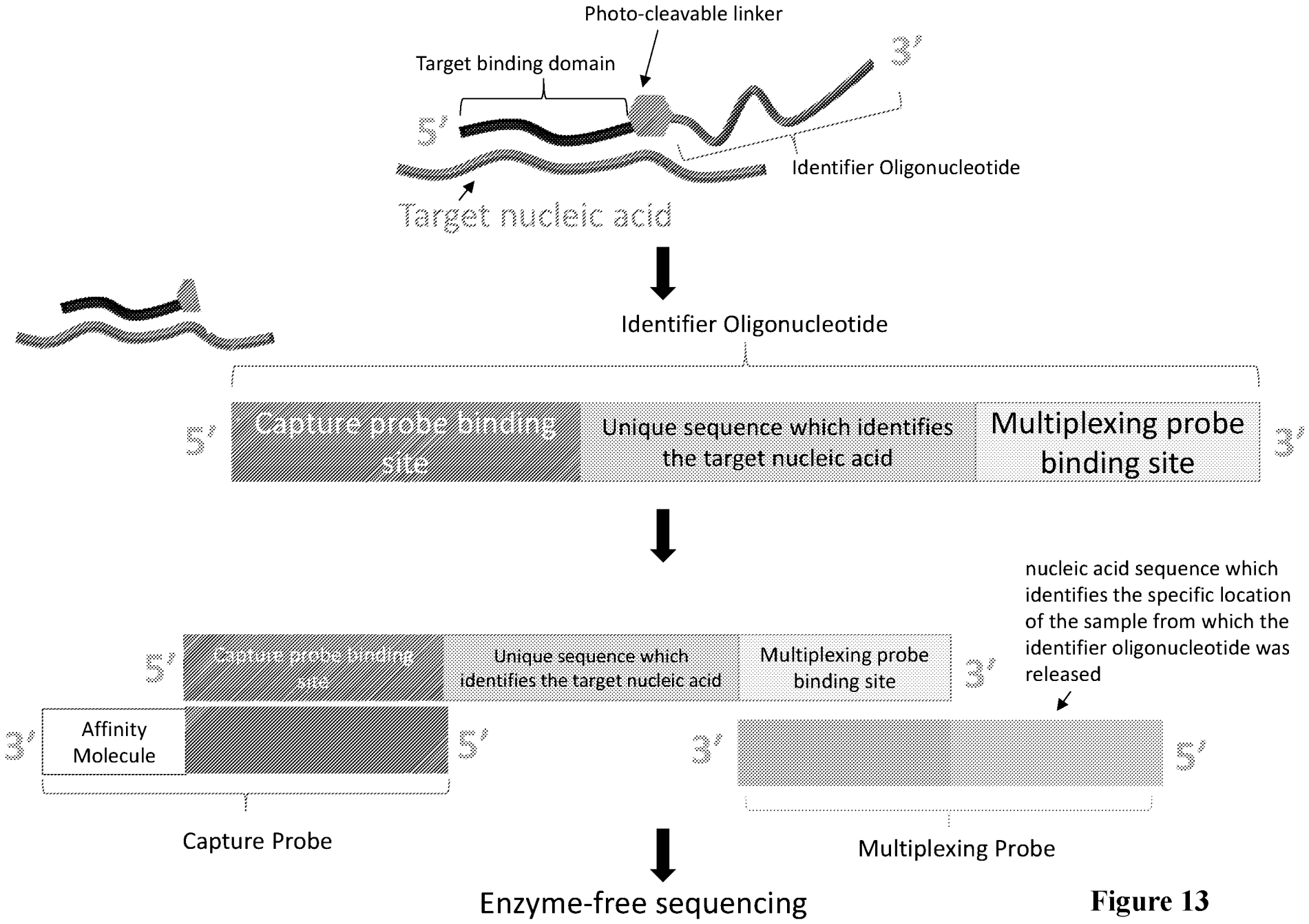

Photo-cleavable linker
Target binding domain
3′
5′
Identifier Oligonucleotide
Target nucleic acid
Identifier Oligonucleotide
5′
Capture probe binding site
Unique sequence which identifies the target nucleic acid
Multiplexing probe binding site
3′
nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
5′
Capture probe binding site
Unique sequence which identifies the target nucleic acid
Multiplexing probe binding site
3′
3′
Affinity Molecule
5′
3′
5′
Capture Probe
Multiplexing Probe
Enzyme-free sequencing

Figure 13

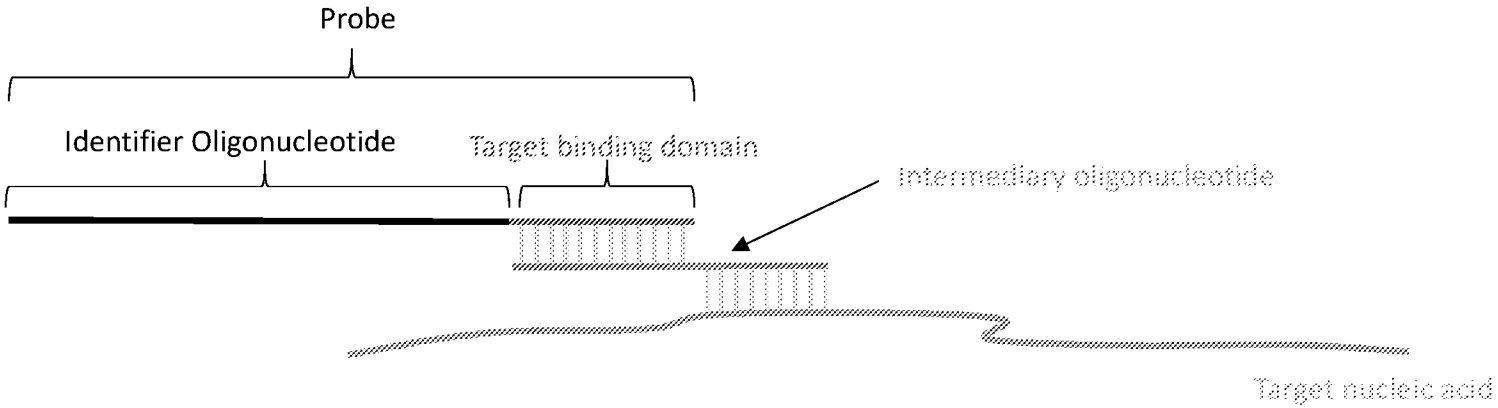

Probe
Identifier Oligonucleotide
Target binding domain
Intermediary oligonucleotide
Target nucleic acid

Figure 14

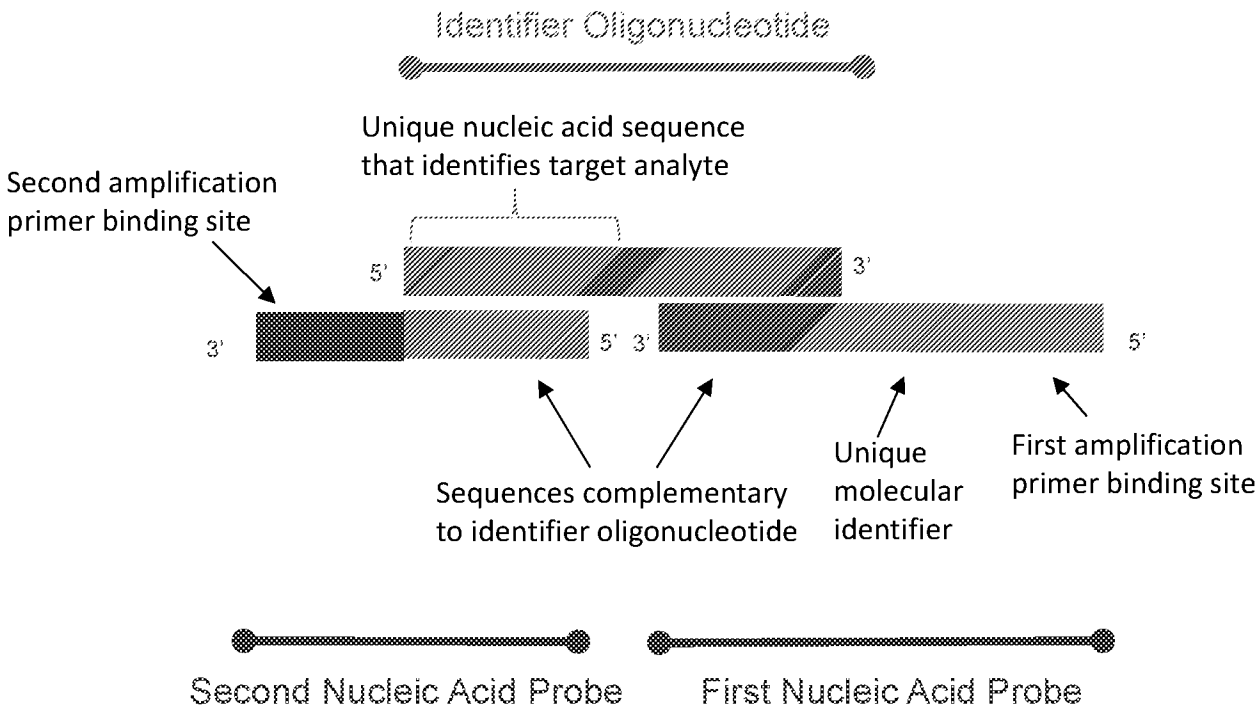

Identifier Oligonucleotide
Unique nucleic acid sequence that identifies target analyte
Second amplification primer binding site
5'
3'
3'
5'
3'
5'
Sequences complementary to identifier oligonucleotide
Unique molecular identifier
First amplification primer binding site
Second Nucleic Acid Probe
First Nucleic Acid Probe

Figure 15

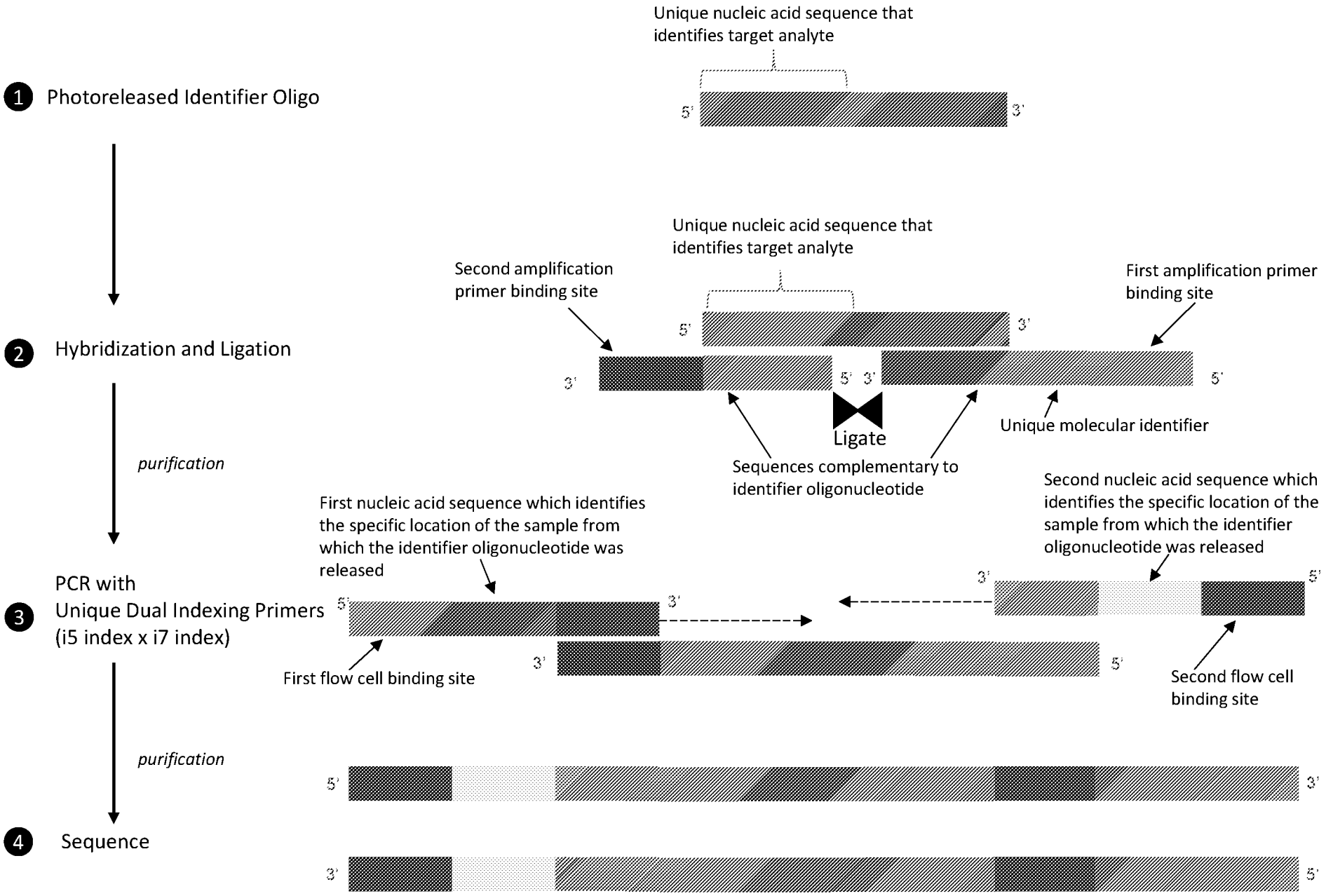

1 Photoreleased Identifier Oligo
Unique nucleic acid sequence that identifies target analyte
5'
3'
2 Hybridization and Ligation
Second amplification primer binding site
Unique nucleic acid sequence that identifies target analyte
First amplification primer binding site
Ligate
Unique molecular identifier
Sequences complementary to identifier oligonucleotide
purification
3 PCR with Unique Dual Indexing Primers (i5 index x i7 index)
First nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Second nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
First flow cell binding site
Second flow cell binding site
purification
4 Sequence

Figure 16

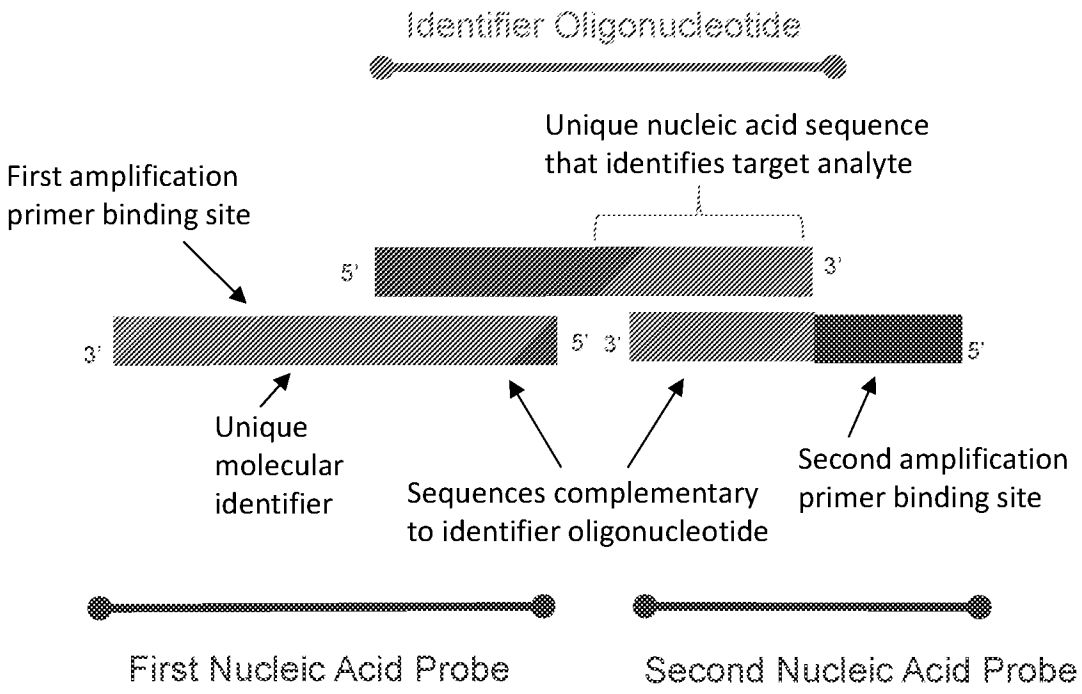

Identifier Oligonucleotide
Unique nucleic acid sequence
that identifies target analyte
First amplification
primer binding site
5'
3'
3'
5'
3'
5'
Unique
molecular
identifier
Sequences complementary
to identifier oligonucleotide
Second amplification
primer binding site
First Nucleic Acid Probe
Second Nucleic Acid Probe

Figure 17

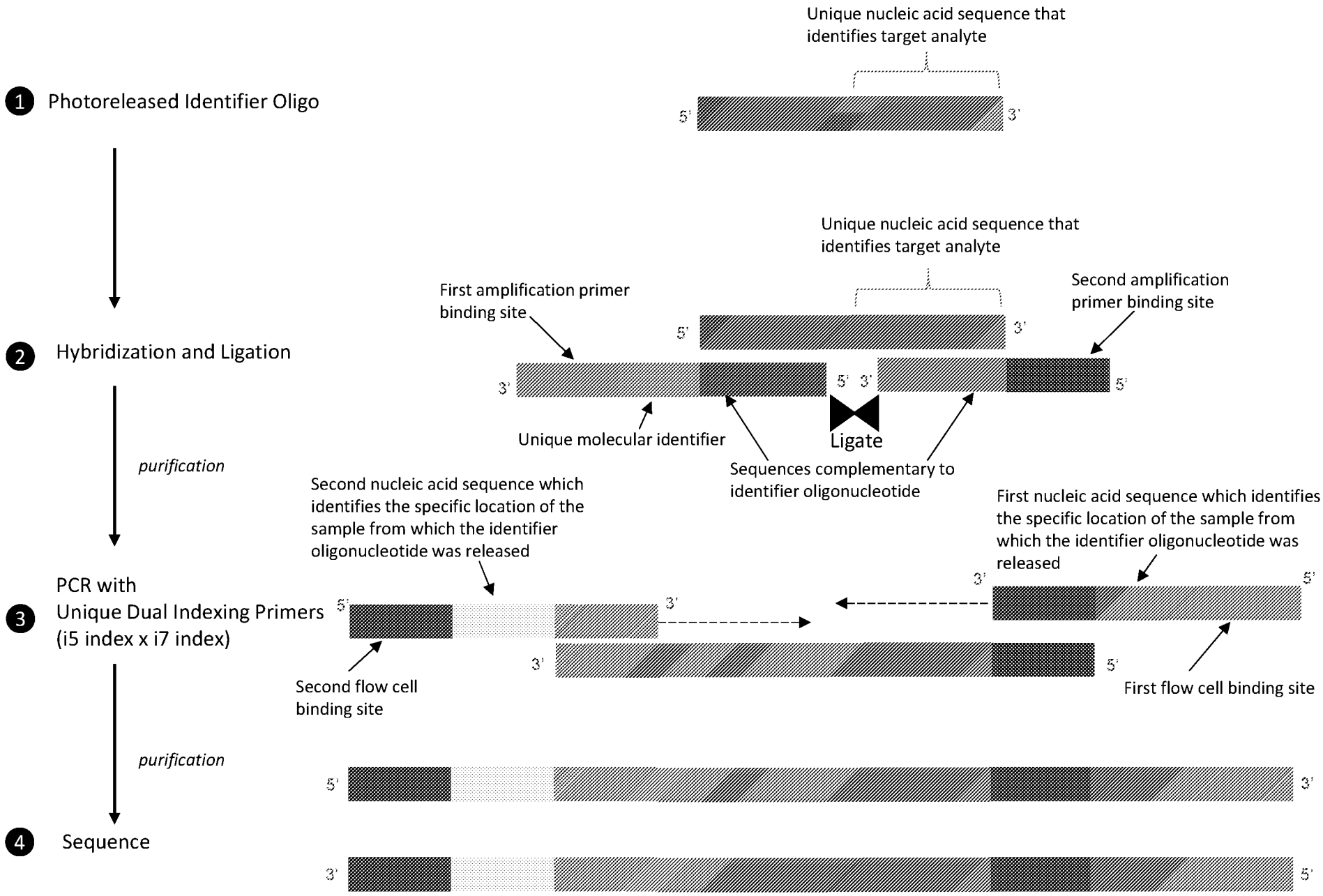

1 Photoreleased Identifier Oligo
Unique nucleic acid sequence that identifies target analyte
5'
3'
2 Hybridization and Ligation
First amplification primer binding site
Unique nucleic acid sequence that identifies target analyte
Second amplification primer binding site
Unique molecular identifier
Ligate
Sequences complementary to identifier oligonucleotide
purification
3 PCR with Unique Dual Indexing Primers (i5 index x i7 index)
Second nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
First nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Second flow cell binding site
First flow cell binding site
purification
4 Sequence

Figure 18

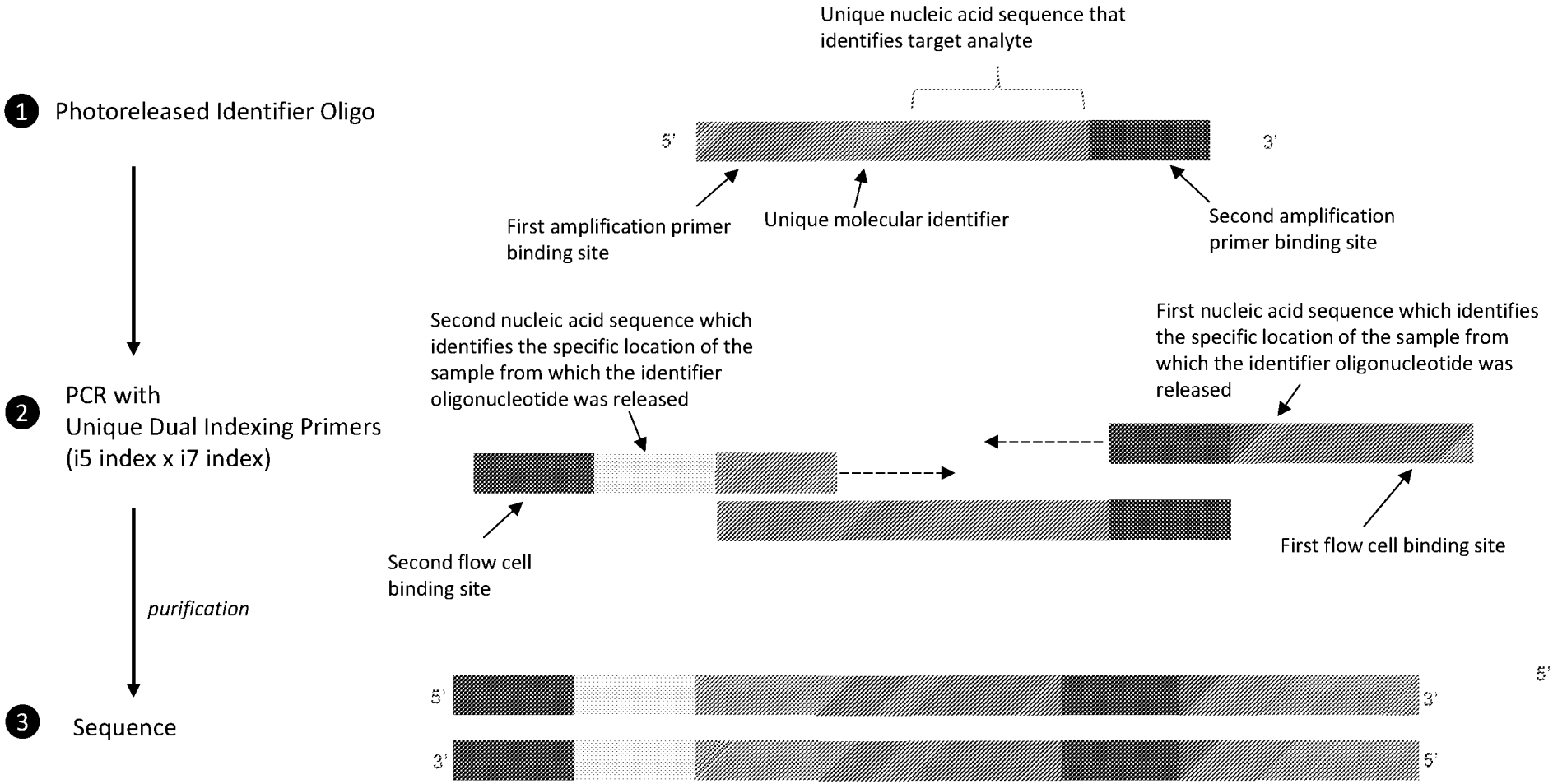

1 Photoreleased Identifier Oligo
2 PCR with
Unique Dual Indexing Primers
(i5 index x i7 index)
purification
3 Sequence
Unique nucleic acid sequence that identifies target analyte
5'
3'
First amplification primer binding site
Unique molecular identifier
Second amplification primer binding site
Second nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
First nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released
Second flow cell binding site
First flow cell binding site
5'
3'
3'
5'
5'

Figure 19

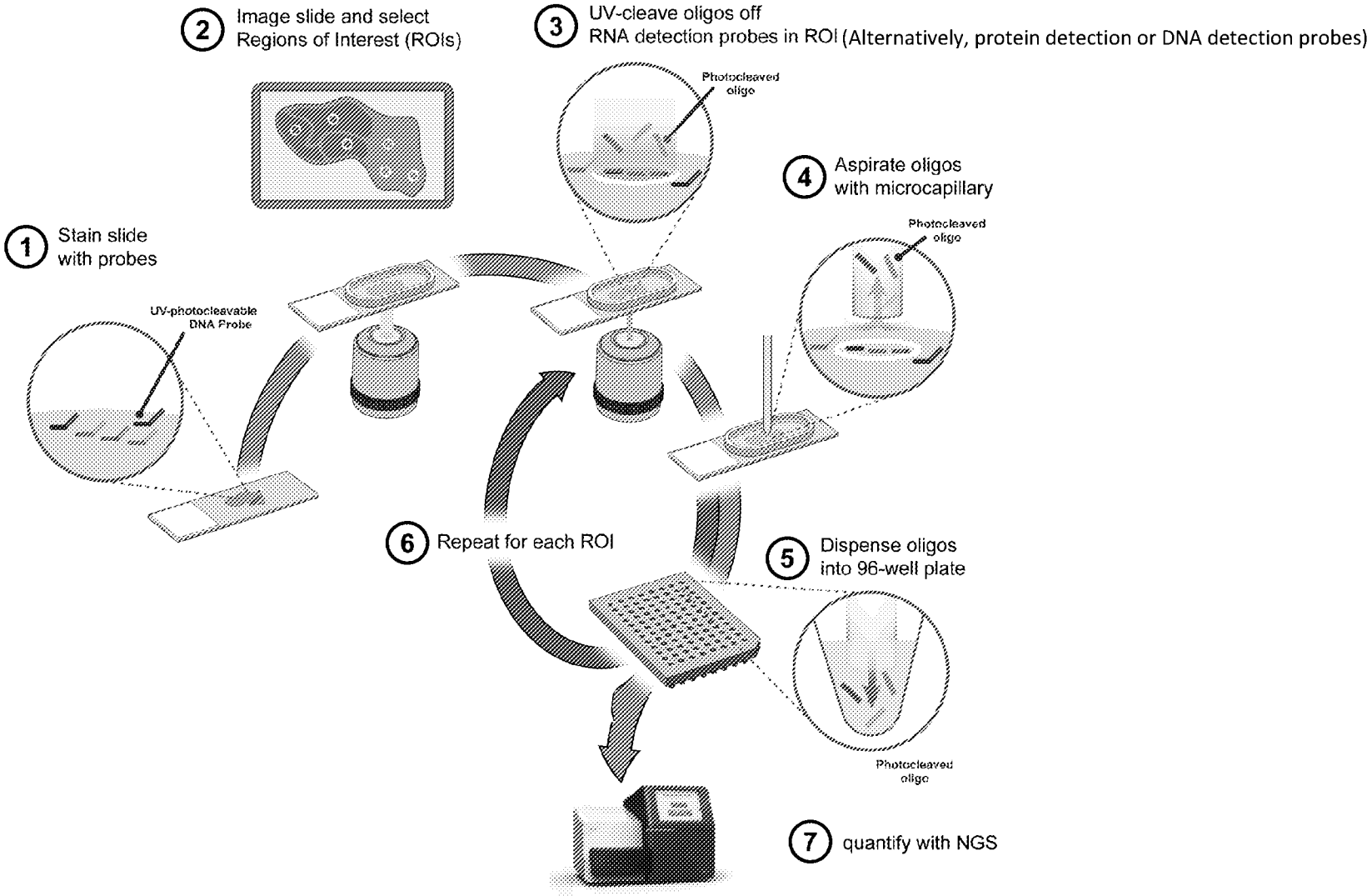

2 Image slide and select Regions of Interest (ROIs)
3 UV-cleave oligos off RNA detection probes in ROI (Alternatively, protein detection or DNA detection probes)
Photocleaved oligo
4 Aspirate oligos with microcapillary
1 Stain slide with probes
UV-photocleavable DNA Probe
Photocleaved oligo
6 Repeat for each ROI
5 Dispense oligos into 96-well plate
Photocleaved oligo
7 quantify with NGS

Figure 20

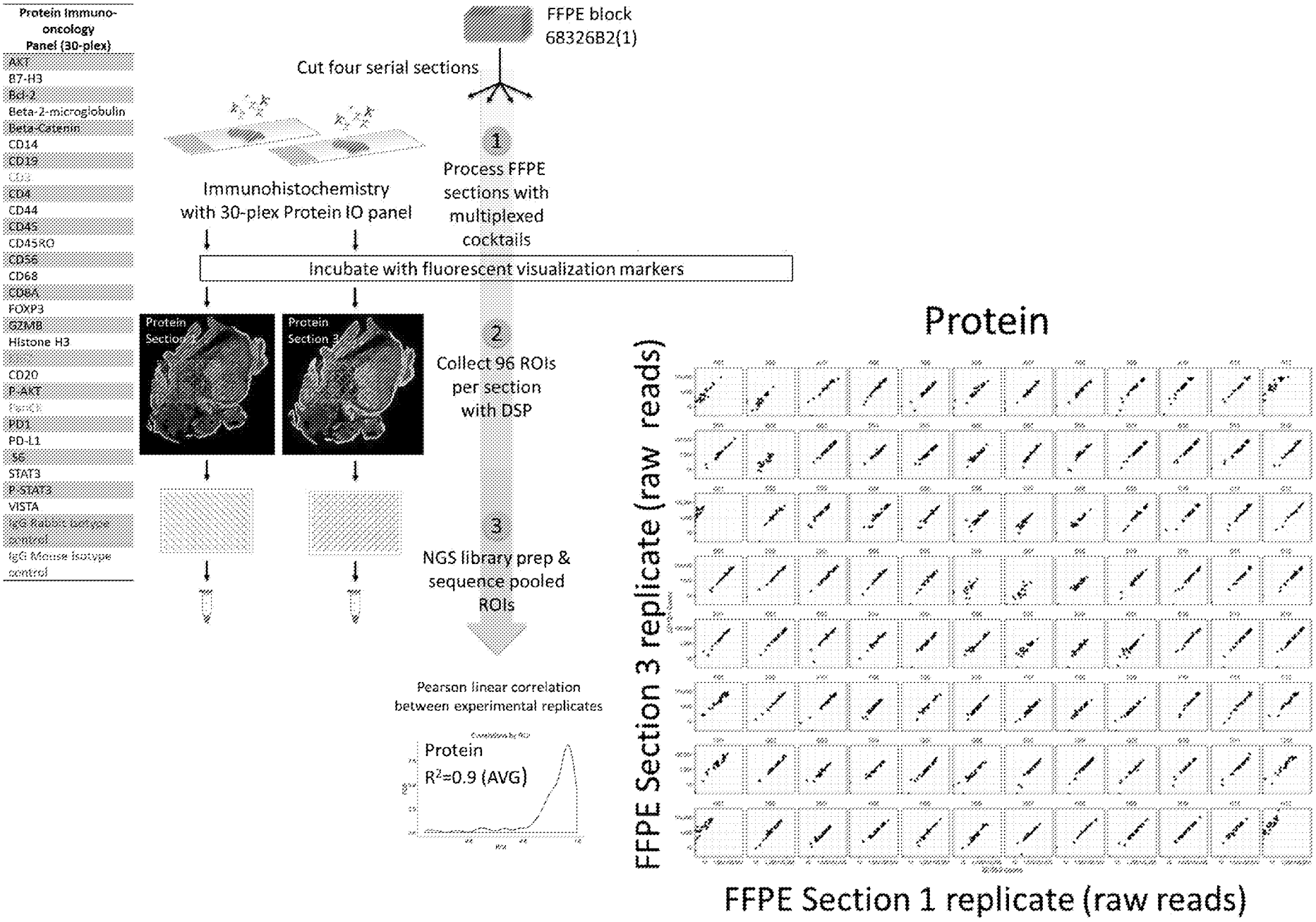

Protein Immuno-oncology Panel (30-plex)
AKT
B7-H3
Bcl-2
Beta-2-microglobulin
Beta-Catenin
CD14
CD19
CD3
CD4
CD44
CD45
CD45RO
CD56
CD68
CD8A
FOXP3
GZMB
Histone H3
CD20
P-AKT
PanCK
PD1
PD-L1
S6
STAT3
P-STAT3
VISTA
IgG Rabbit isotype control
IgG Mouse isotype control
FFPE block 68326B2(1)
Cut four serial sections
1
Process FFPE sections with multiplexed cocktails
Immunohistochemistry with 30-plex Protein IO panel
Incubate with fluorescent visualization markers
Protein Section 1
Protein Section 3
2
Collect 96 ROIs per section with DSP
3
NGS library prep & sequence pooled ROIs
Pearson linear correlation between experimental replicates
Protein
R2=0.9 (AVG)
Protein
FFPE Section 3 replicate (raw reads)
FFPE Section 1 replicate (raw reads)

Figure 21

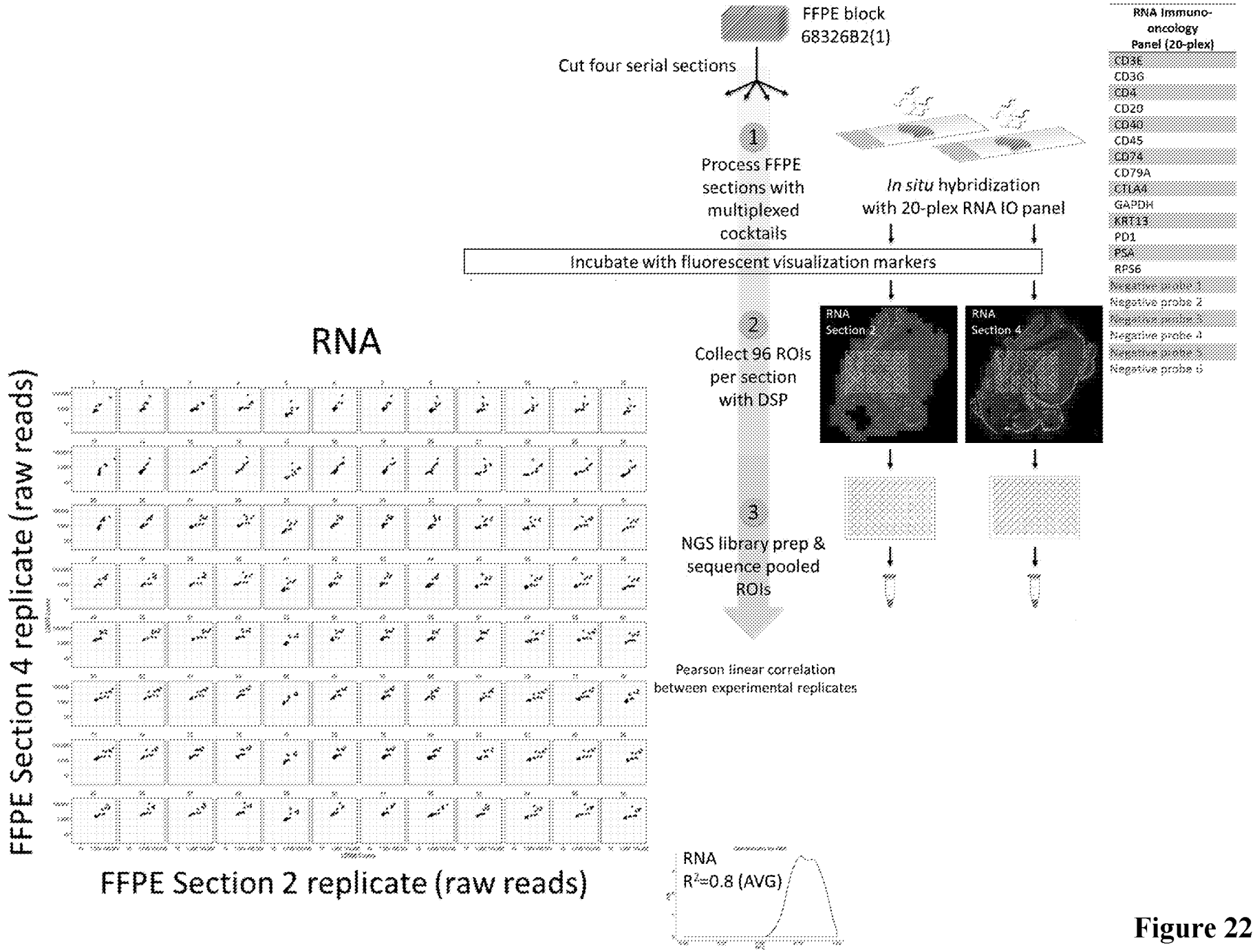

FFPE block
68326B2(1)
Cut four serial sections
1
Process FFPE sections with multiplexed cocktails
In situ hybridization with 20-plex RNA IO panel
Incubate with fluorescent visualization markers
RNA Immuno-oncology Panel (20-plex)
CD3E
CD3G
CD4
CD20
CD40
CD45
CD74
CD79A
CTLA4
GAPDH
KRT13
PD1
PSA
RPS6
Negative probe 1
Negative probe 2
Negative probe 3
Negative probe 4
Negative probe 5
Negative probe 6
RNA Section 2
RNA Section 4
2
Collect 96 ROIs per section with DSP
3
NGS library prep & sequence pooled ROIs
Pearson linear correlation between experimental replicates
RNA
FFPE Section 4 replicate (raw reads)
FFPE Section 2 replicate (raw reads)
RNA
R²=0.8 (AVG)

Figure 22

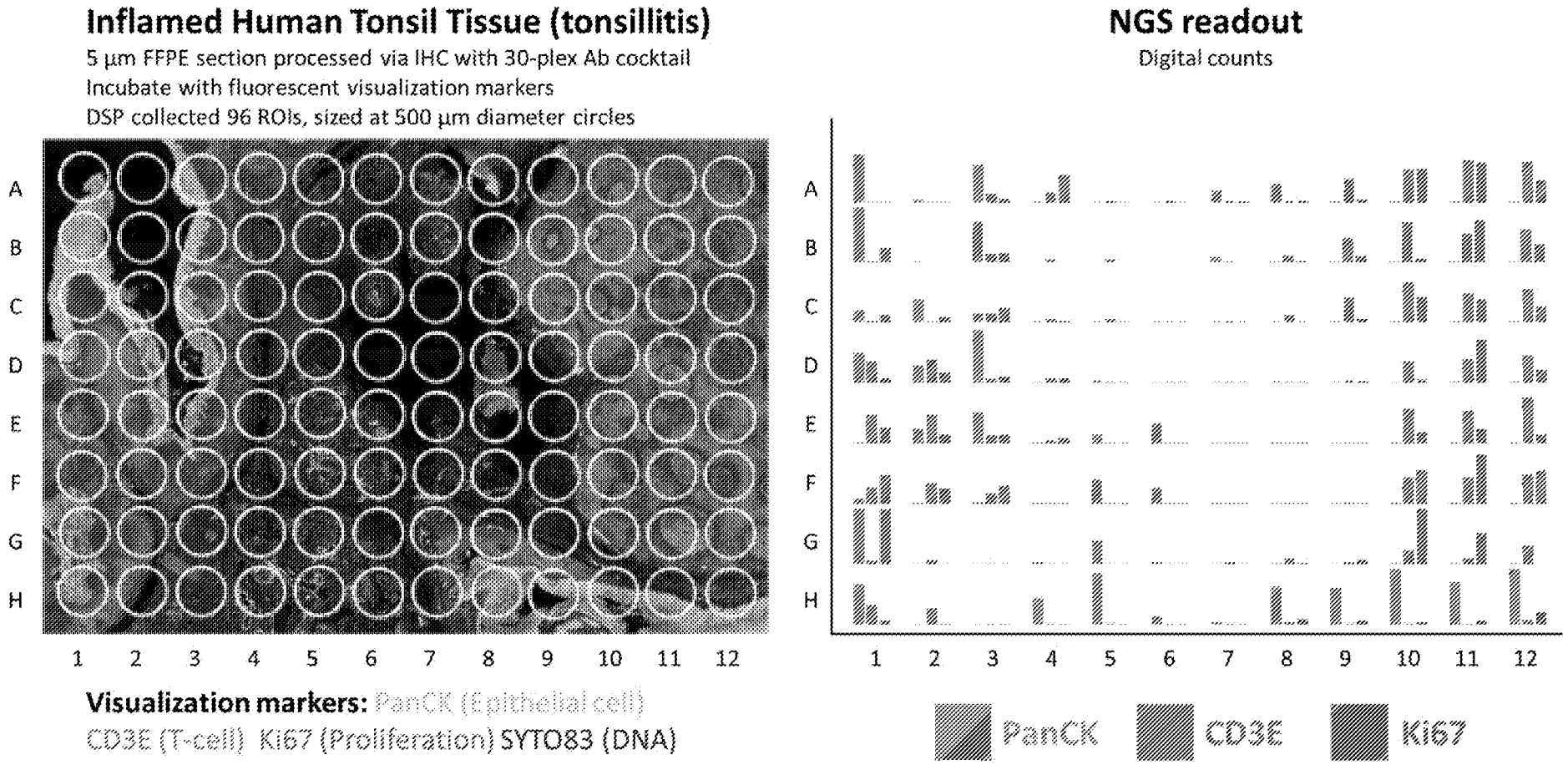

Inflamed Human Tonsil Tissue (tonsillitis)
5 µm FFPE section processed via IHC with 30-plex Ab cocktail
Incubate with fluorescent visualization markers
DSP collected 96 ROIs, sized at 500 µm diameter circles
A
B
C
D
E
F
G
H
1 2 3 4 5 6 7 8 9 10 11 12
Visualization markers: PanCK (Epithelial cell)
CD3E (T-cell) Ki67 (Proliferation) SYTO83 (DNA)
NGS readout
Digital counts
A
B
C
D
E
F
G
H
1 2 3 4 5 6 7 8 9 10 11 12
PanCK
CD3E
Ki67

Figure 23

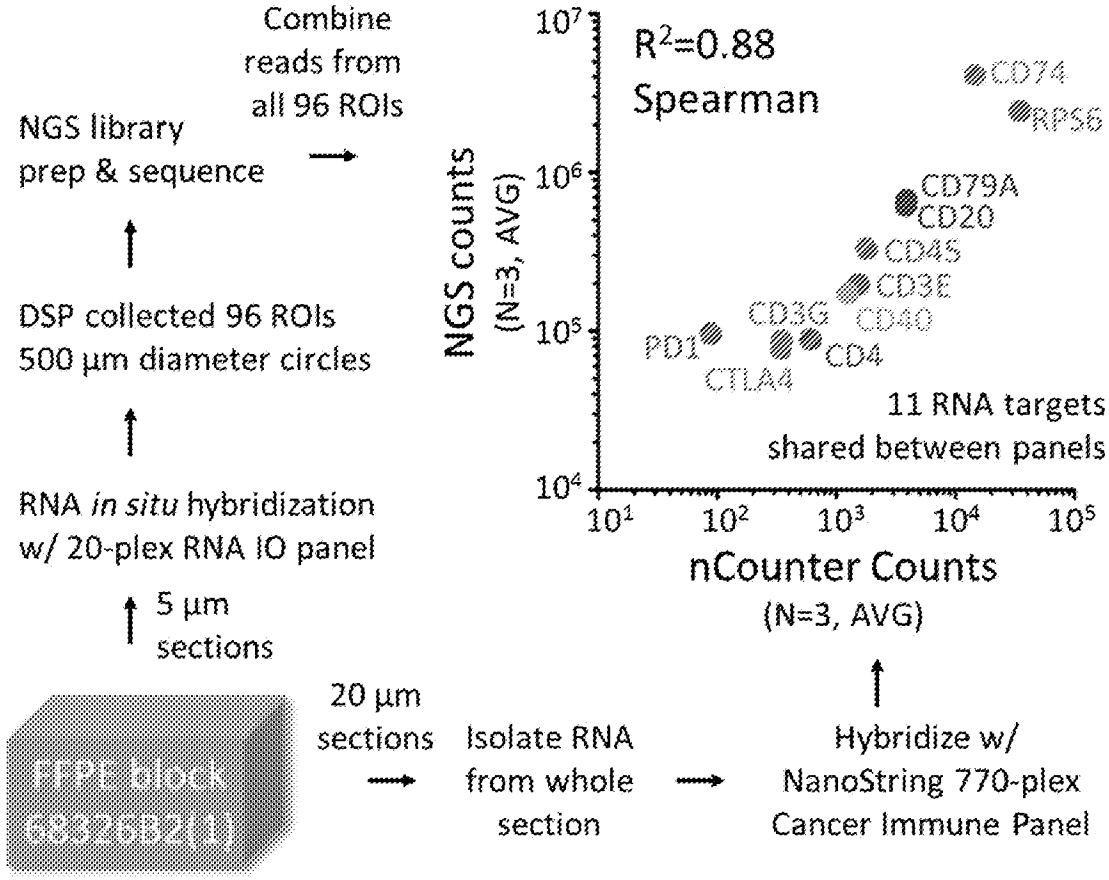

Combine reads from all 96 ROIs
NGS library prep & sequence
DSP collected 96 ROIs 500 µm diameter circles
RNA in situ hybridization w/ 20-plex RNA IO panel
5 µm sections
FFPE block 68326B2(1)
20 µm sections
Isolate RNA from whole section
Hybridize w/ NanoString 770-plex Cancer Immune Panel
R²=0.88
Spearman
NGS counts (N=3, AVG)
nCounter Counts (N=3, AVG)
CD74
RPS6
CD79A
CD20
CD45
CD3E
CD3G
CD40
PD1
CD4
CTLA4
11 RNA targets shared between panels

Figure 24

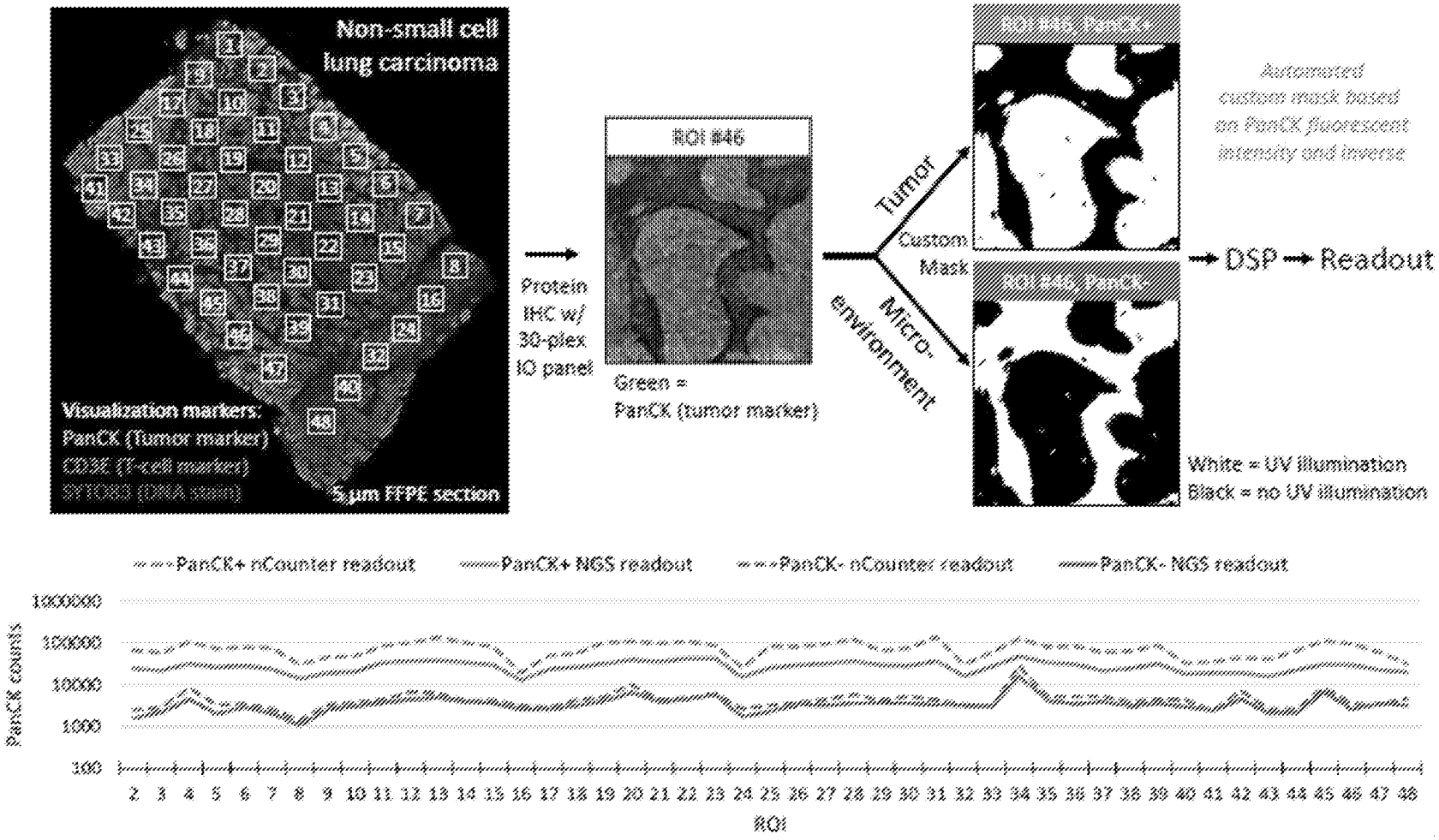
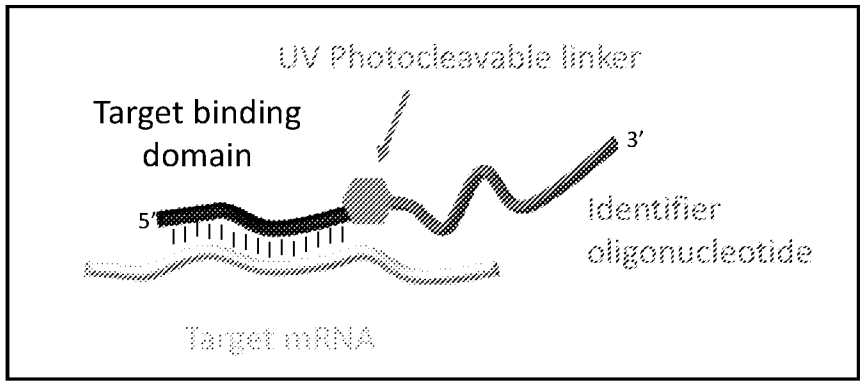

Non-small cell lung carcinoma
Visualization markers:
PanCK (Tumor marker)
CD3E (T-cell marker)
SYTO13 (DNA stain)
5 µm FFPE section
Protein IHC w/ 30-plex IO panel
ROI #46
Green = PanCK (tumor marker)
Tumor
Custom Mask
Micro-environment
ROI #46, PanCK+
ROI #46, PanCK-
Automated custom mask based on PanCK fluorescent intensity and inverse
DSP → Readout
White = UV illumination
Black = no UV illumination
PanCK+ nCounter readout
PanCK+ NGS readout
PanCK- nCounter readout
PanCK- NGS readout
PanCK counts
1000000
100000
10000
1000
100
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48
ROI

Figure 25


UV Photocleavable linker
Target binding domain
3'
5'
identifier oligonucleotide
Target mRNA

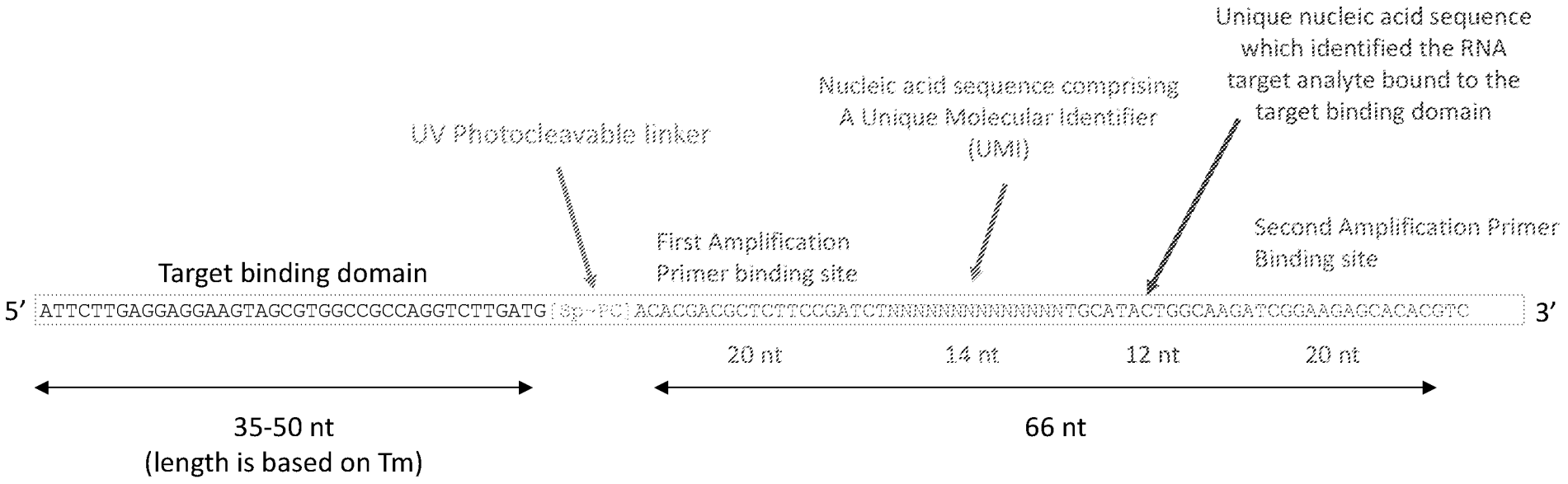

Unique nucleic acid sequence which identified the RNA target analyte bound to the target binding domain
Nucleic acid sequence comprising A Unique Molecular Identifier (UMI)
UV Photocleavable linker
Second Amplification Primer Binding site
First Amplification Primer binding site
Target binding domain
5'
ATTCTTGAGGAGGAAGTAGCGTGGCCGCCAGGTCTTGATG
iSp-PC
ACACGACGCTCTTCCGATCTNNNNNNNNNNNNNNNNTGCATACTGGCAAGATCGGAAGAGCACACGTC
3'
20 nt
14 nt
12 nt
20 nt
35-50 nt
(length is based on Tm)
66 nt

Figure 26

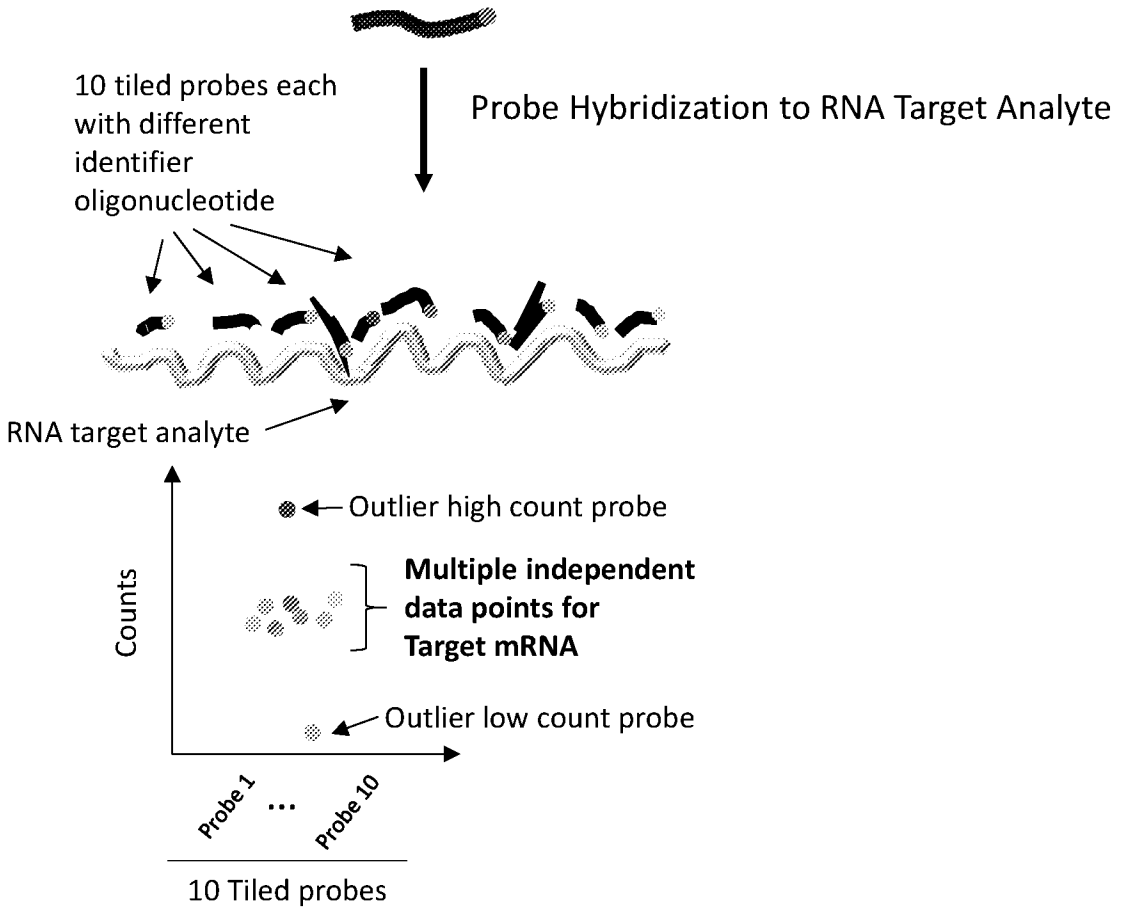

10 tiled probes each with different identifier oligonucleotide
Probe Hybridization to RNA Target Analyte
RNA target analyte
Outlier high count probe
Multiple independent data points for Target mRNA
Counts
Outlier low count probe
Probe 1
...
Probe 10
10 Tiled probes

Figure 27

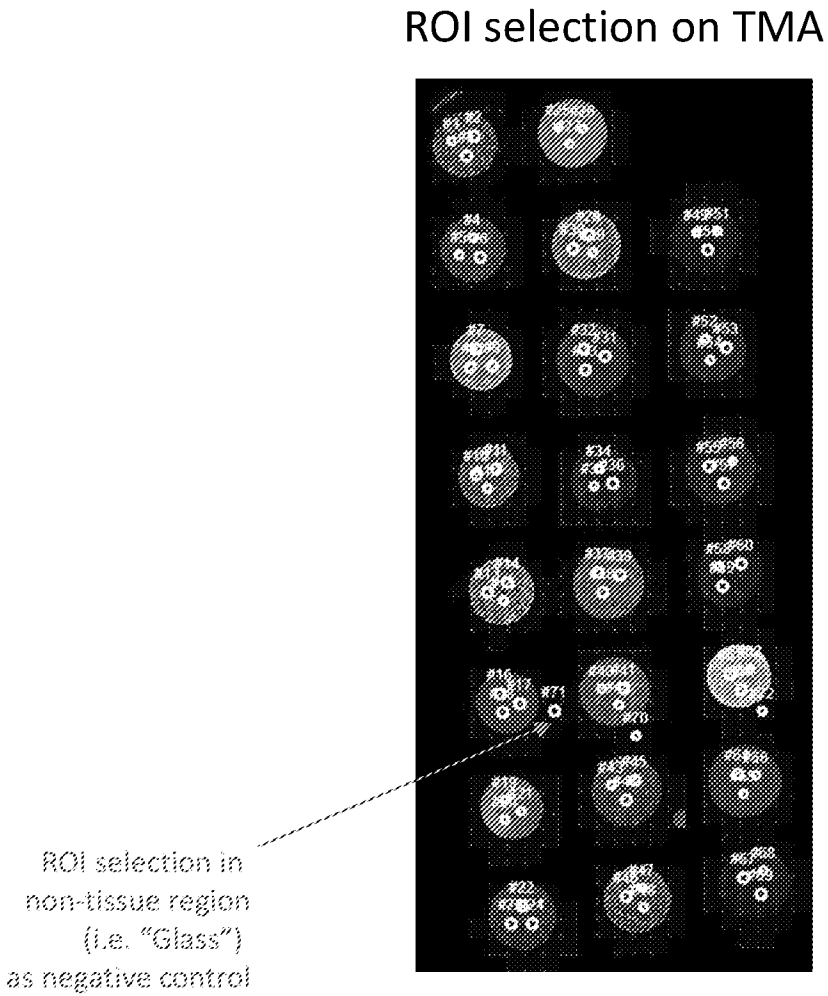

ROI selection on TMA
ROI selection in non-tissue region (i.e. "Glass") as negative control

Figure 28

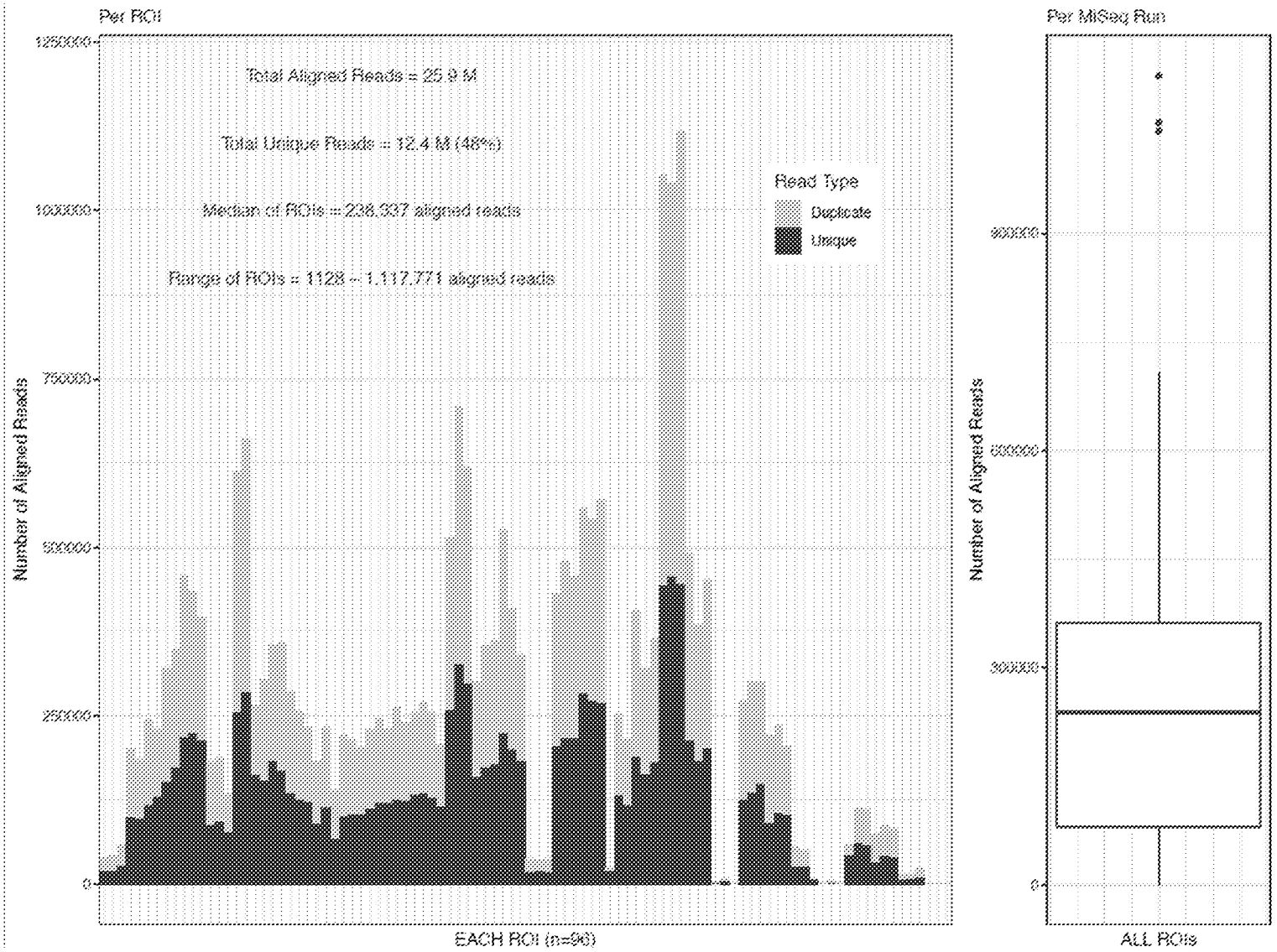
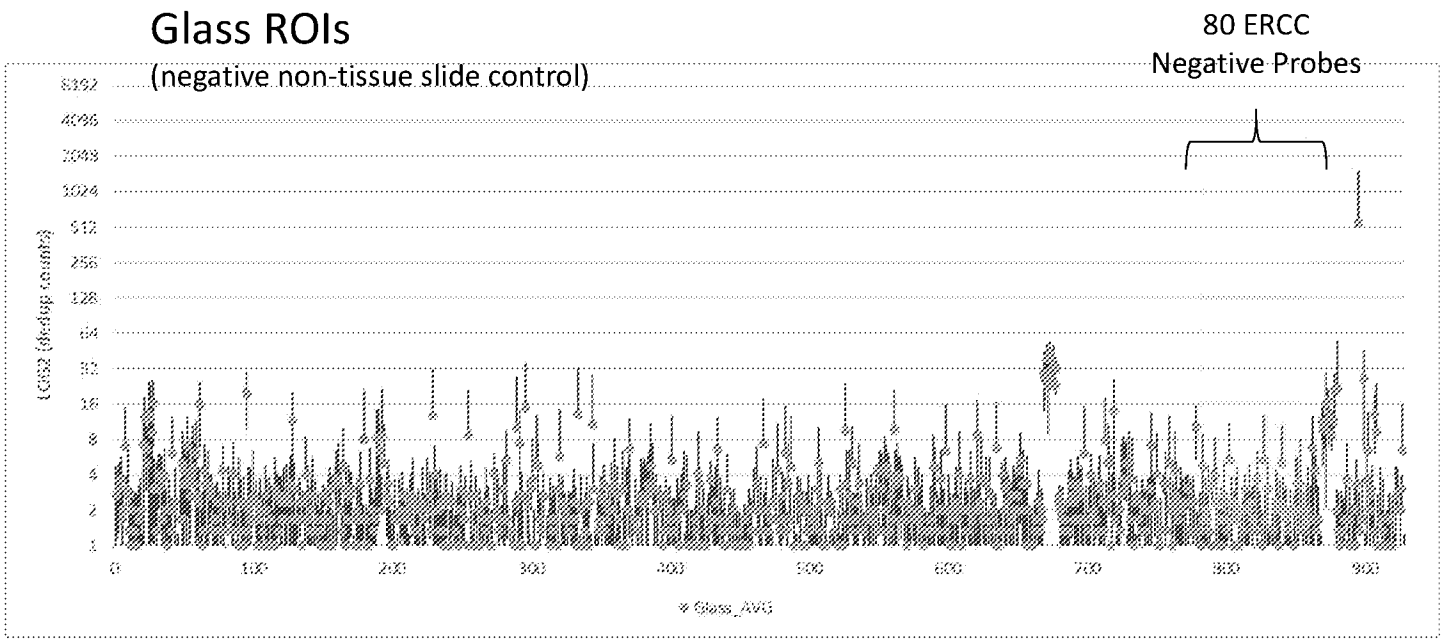

Per ROI
Total Aligned Reads = 25.9 M
Total Unique Reads = 12.4 M (48%)
Median of ROIs = 238,337 aligned reads
Range of ROIs = 1128 – 1,117,771 aligned reads
Read Type
Duplicate
Unique
Number of Aligned Reads
EACH ROI (n=96)
Per MiSeq Run
Number of Aligned Reads
ALL ROIs

Figure 29


Glass ROIs
(negative non-tissue slide control)
80 ERCC
Negative Probes
8192
4096
2048
1024
512
256
128
64
32
16
8
4
2
1
LOG2 (dedup counts)
0
100
200
300
400
500
600
700
800
900
Glass_AVG

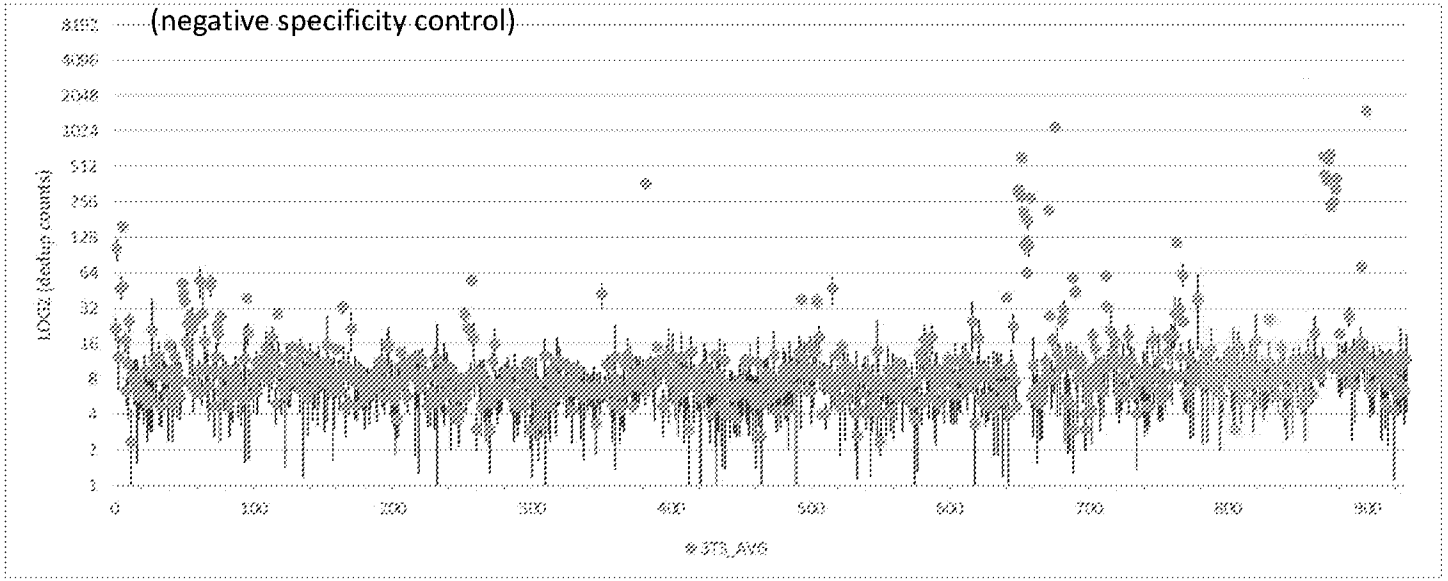

3T3 Mouse Embryonic Fibroblasts ROIs
(negative specificity control)
8192
4096
2048
1024
512
256
128
64
32
16
8
4
2
1
LOG2 (dedup counts)
0
100
200
300
400
500
600
700
800
900
3T3_AVG

Figure 30

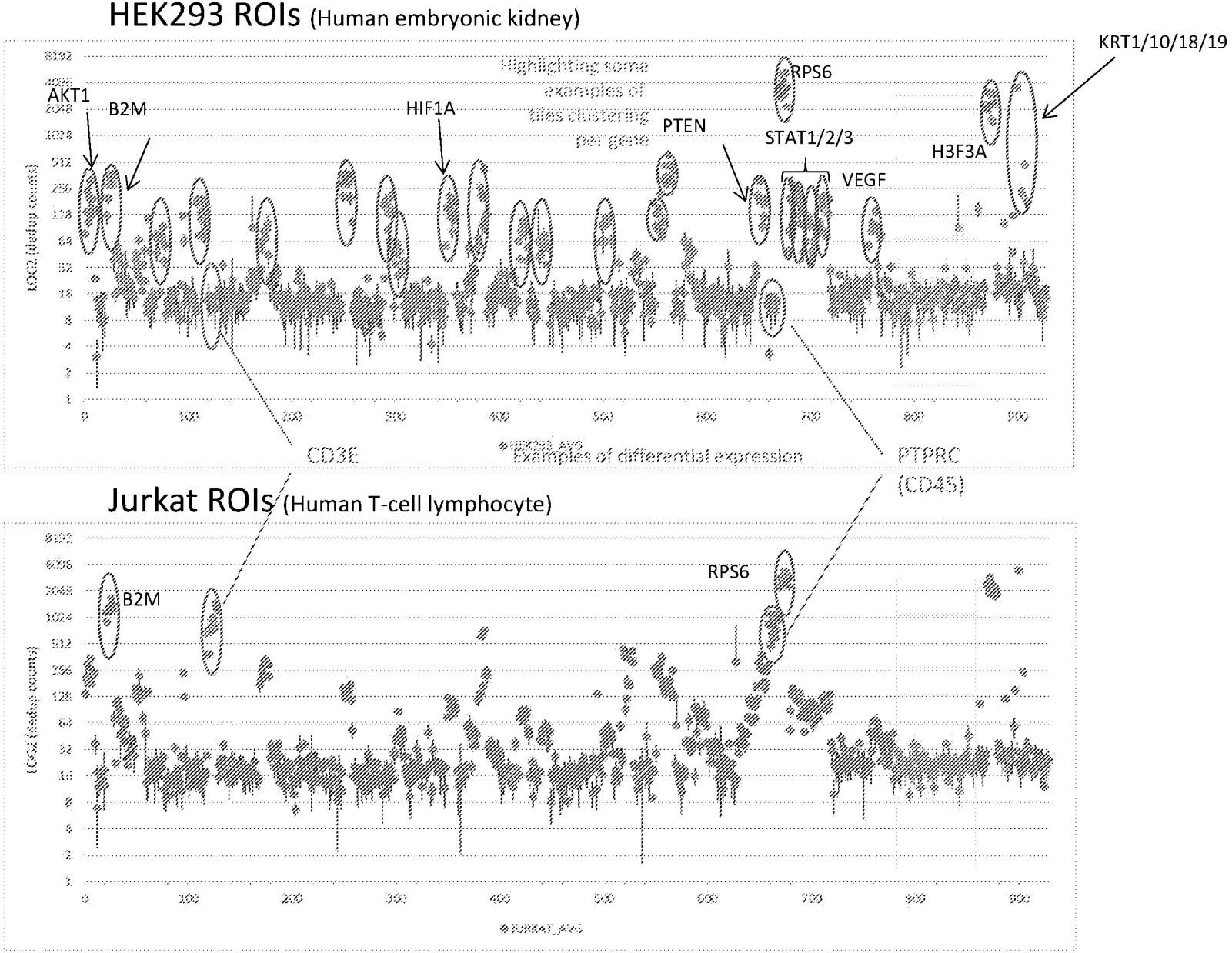

HEK293 ROIs (Human embryonic kidney)
KRT1/10/18/19
Highlighting some examples of tiles clustering per gene
AKT1
B2M
HIF1A
PTEN
RPS6
STAT1/2/3
VEGF
H3F3A
CD3E
PTPRC (CD45)
Examples of differential expression
Jurkat ROIs (Human T-cell lymphocyte)
B2M
RPS6

Figure 31

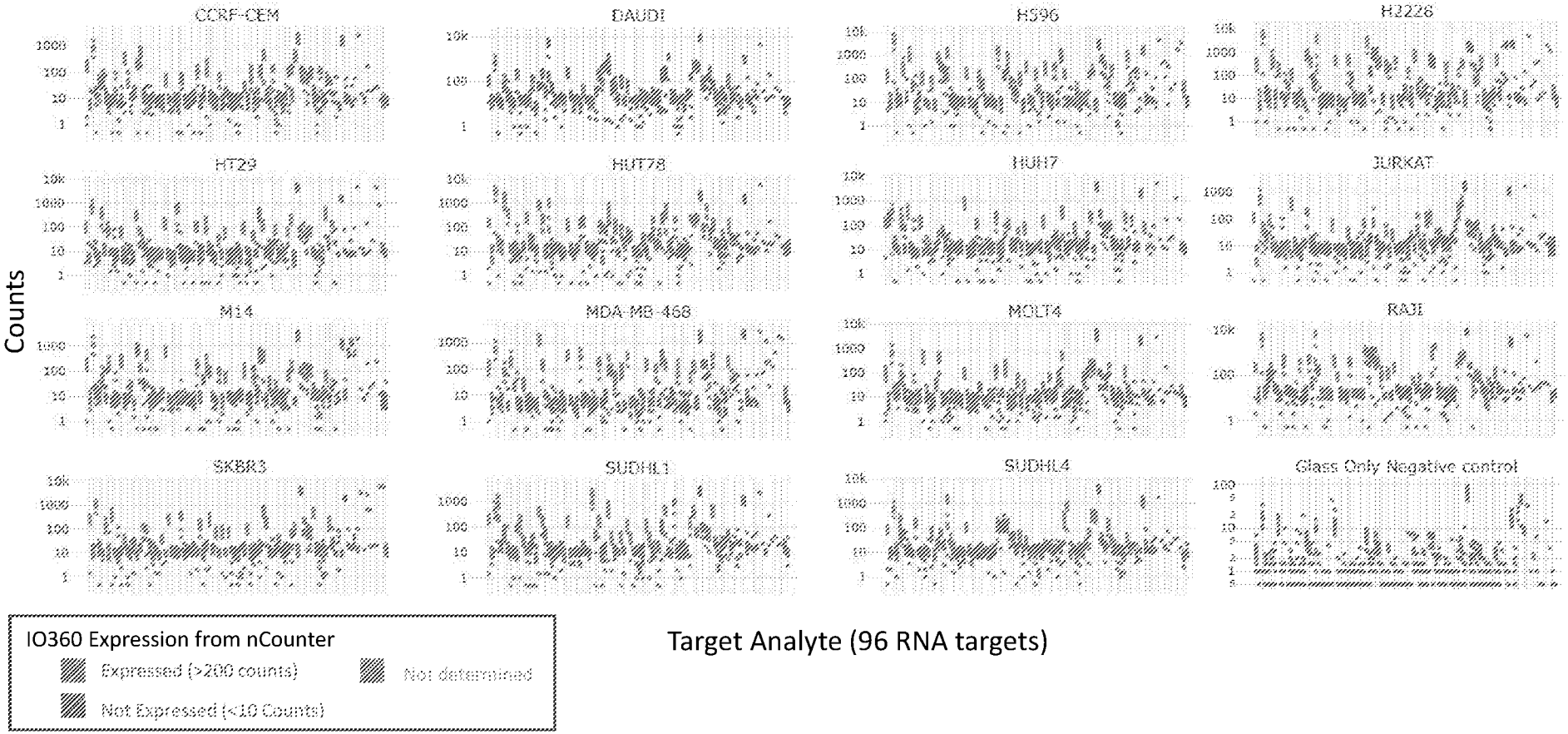

CCRF-CEM
DAUDI
HS96
H2228
HT29
HUT78
HUH7
JURKAT
M14
MDA-MB-468
MOLT4
RAJI
SKBR3
SUDHL1
SUDHL4
Glass Only Negative control
Counts
IO360 Expression from nCounter
Expressed (>200 counts)
Not determined
Not Expressed (<10 Counts)
Target Analyte (96 RNA targets)

Figure 32

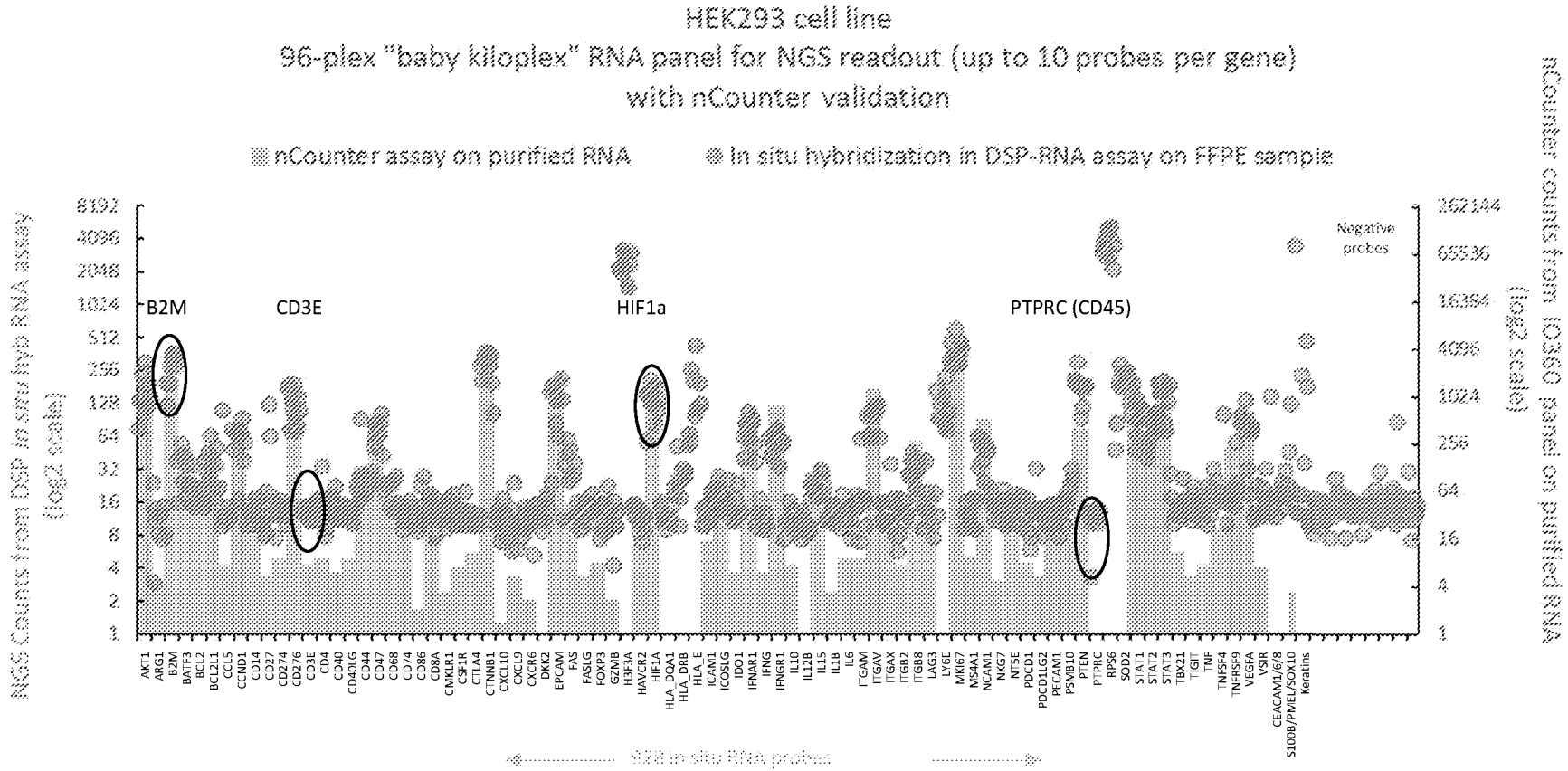

HEK293 cell line
96-plex "baby kiloplex" RNA panel for NGS readout (up to 10 probes per gene)
with nCounter validation
nCounter assay on purified RNA
In situ hybridization in DSP-RNA assay on FFPE sample
NGS Counts from DSP in situ hyb RNA assay (log2 scale)
nCounter counts from IO360 panel on purified RNA (log2 scale)
B2M
CD3E
HIF1a
PTPRC (CD45)
Negative probes
828 in situ RNA probes

Figure 33

Jurkat cell line
96-plex "baby kiloplex" RNA panel for NGS Readout (up to 10 probes per gene)
with nCounter validation
nCounter assay on purified RNA
In situ hybridization in DSP-RNA assay on FFPE sample
nCounter counts from IO360 panel on purified RNA
(log2 scale)
NGS Counts from DSP in situ hyb RNA assay
(log2 scale)
B2M
CD3E
HIF1a
PTPRC (CD45)
Negative probes
928 in situ RNA probes
AKT1
ARG1
B2M
BATF3
BCL2
BCL2L1
CCL5
CCND1
CD14
CD27
CD274
CD276
CD3E
CD4
CD40
CD40LG
CD44
CD47
CD68
CD74
CD86
CD8A
CMKLR1
CSF1R
CTLA4
CTNNB1
CXCL10
CXCL9
CXCR6
DKK2
EPCAM
FAS
FASLG
FOXP3
GZMB
H3F3A
HAVCR2
HIF1A
HLA_DQA1
HLA_DRB
HLA_E
ICAM1
ICOSLG
IDO1
IFNAR1
IFNG
IFNGR1
IL10
IL12B
IL15
IL18
IL6
ITGAM
ITGAV
ITGAX
ITGB2
ITGB8
LAG3
LY6E
MKI67
MS4A1
NCAM1
NKG7
NT5E
PDCD1
PDCD1LG2
PECAM1
PSMB10
PTEN
PTPRC
RPS6
SOD2
STAT1
STAT2
STAT3
TBX21
TIGIT
TNF
TNFSF4
TNFRSF9
VEGFA
VSIR
CEACAM1/6/8
S100B/PMEL/SOX10
Keratins

Figure 34

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

Secrecy Order 37 CFR 5.2:

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Inventor Information:

Inventor 1 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Joseph	M.	BEECHEM	

Residence Information (Select One) ● US Residency Non US Residency Active US Military Service

City	State/Province	Country of Residence
Eugene	OR	US

Mailing Address of Inventor:

Address 1	46 Ridgeline Drive
Address 2	
City	Eugene
State/Province	OR
Postal Code	97405
Country	US

Inventor 2 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Dae		KIM	

Residence Information (Select One) ◉ US Residency Non US Residency Active US Military Service

City	State/Province	Country of Residence
Bellevue	WA	US

Mailing Address of Inventor:

Address 1	16729 SE 63rd Place
Address 2	
City	Bellevue
State/Province	WA
Postal Code	98006
Country	US

Inventor 3 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Margaret		HOANG	

Residence Information (Select One) ◉ US Residency Non US Residency Active US Military Service

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

City	Seattle	State/Province	WA	Country of Residence	US

Mailing Address of Inventor:

Address 1	1111 South Nevada Street
Address 2	
City	Seattle
State/Province	WA
Postal Code	98108
Country	US

Inventor 4 — Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Mark		GREGORY	

Residence Information (Select One) ◉ US Residency ○ Non US Residency ○ Active US Military Service

City	Seattle	State/Province	WA	Country of Residence	US

Mailing Address of Inventor:

Address 1	14327 24th Avenue NE
Address 2	
City	Seattle
State/Province	WA
Postal Code	98125
Country	US

Inventor 5 — Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Erin		PIAZZA	

Residence Information (Select One) ◉ US Residency ○ Non US Residency ○ Active US Military Service

City	Edmonds	State/Province	WA	Country of Residence	US

Mailing Address of Inventor:

Address 1	15314 57th Pl W
Address 2	
City	Edmonds
State/Province	WA
Postal Code	98026
Country	US

Inventor 6 — Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Denise		ZHOU	

Residence Information (Select One) ◉ US Residency ○ Non US Residency ○ Active US Military Service

City	Seattle	State/Province	WA	Country of Residence	US

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

Mailing Address of Inventor:

Address 1	826 24th Ave S		
Address 2			
City	Seattle	**State/Province**	WA
Postal Code	98144	**Country** i	US

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below.
For further information see 37 CFR 1.33(a).

☐ **An Address is being provided for the correspondence Information of this application.**

Customer Number	141429		
Email Address	zIPPatentDocketingMailboxUS@Cooley.com	Add Email	Remove Email

Application Information:

Title of the Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION		
Attorney Docket Number	NATE-037001US 321329-2523	**Small Entity Status Claimed**	☐
Application Type	Nonprovisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	34	**Suggested Figure for Publication (if any)**	

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country i

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	● Customer Number	US Patent Practitioner	○ Limited Recognition (37 CFR 11.9)
Customer Number	141429		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.
When referring to the current application, please leave the "Application Number" field blank.

Prior Application Status	Pending		Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)
	Claims benefit of provisional	62/771212	2018-11-26
Prior Application Status	Pending		Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)
	Claims benefit of provisional	62/629180	2018-02-12

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button. Add

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX)[i] the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Remove

Application Number	Country[i]	Filing Date (YYYY-MM-DD)	Access Code[i] (if applicable)

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button. Add

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.
NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☒ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Applicant 1 — Remove

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.

Clear

Assignee	Legal Representative under 35 U.S.C. 117	Joint Inventor
● Person to whom the inventor is obligated to assign.	Person who shows sufficient proprietary interest	

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here. ☒

Organization Name	NanoString Technologies, Inc.		
Mailing Address Information For Applicant:			
Address 1	530 Fairview Avenue North		
Address 2	Suite 2000		
City	Seattle	**State/Province**	WA
Country	US	Postal Code	98109
Phone Number		Fax Number	
Email Address			

Additional Applicant Data may be generated within this form by selecting the Add button. Add

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

Assignee 1

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

Remove

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

Prefix	**Given Name**	Middle Name	**Family Name**	Suffix

Mailing Address Information For Assignee including Non-Applicant Assignee:

Address 1			
Address 2			
City		**State/Province**	
Country i		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button. Add

Signature:

Remove

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet must be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Matthew Pavao/			Date (YYYY-MM-DD)	2019-02-11
First Name	Matthew	Last Name	Pavao	Registration Number	50,572

Additional Signature may be generated within this form by selecting the Add button. Add

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NATE-037001US 321329-2523
	Application Number	
Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION	

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1 The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3 A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent CooperationTreaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

ABSTRACT OF THE INVENTION

The present invention relates to, among other things, probes, compositions, methods, and kits for simultaneous, multiplexed detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell that are adaptable for use with existing sequencing technologies.

BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

BACKGROUND OF THE INVENTION

[0001] This application claims priority to, and the benefit of, U.S. Provisional Application No. 62/629,180, filed February 12, 2018 and U.S. Provisional Application No. 62/771,212, filed November 26, 2018. The contents of each of the aforementioned patent applications are incorporated herein by reference in their entireties.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on February 1, 2019, is named "NATE-037_001US_Sequence_Listing.txt" and is 48.5 KB in size.

BACKGROUND OF THE INVENTION

[0003] Standard immunohistochemical and *in situ* hybridization methods allow for simultaneous detection of, at most, six to ten protein or nucleic acid targets, with three to four targets being typical. There exists a need for probes, compositions, methods, and kits for simultaneous, multiplexed detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell. Furthermore, there is a need for such systems to be adaptable for use with existing sequencing technologies that are already available to a large number of end users.

SUMMARY OF THE INVENTION

[0004] The present disclosure relates to probes, compositions, methods, and kits for simultaneous, multiplexed, spatial detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell.

[0005] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the ligation product produced in step (4); and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0006] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises: a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias, and wherein at least one of the first or second amplification primer binding sites identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; (5) amplifying the ligation product produced in step (4); and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0007] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises: a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the extension product produced in step (4) using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0008] The nucleic acid adapter of step (4) can be a partially double-stranded nucleic acid molecule. A partially double-stranded nucleic acid adapter can comprise a double-stranded annealed region, a first single-stranded mismatched region and a second single-stranded mismatched region. The first single-stranded mismatched region and the second single stranded mismatched region can be present on opposing sides of the double-stranded annealed region.

[0009] The nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter.

[0010] The constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter.

[0011] A unique molecular identifier can be present in at least one of the first or second single-stranded mismatched regions of a partially double-stranded nucleic acid adapter.

[0012] The first amplification primer binding site can be present in the first single-stranded mismatched region of a partially double-stranded nucleic acid adapter and the second amplification primer binding site can be present in the second single-stranded mismatched region of the same partially double-stranded nucleic acid adapter.

[0013] The methods of the present disclosure described in the preceding can further comprise prior to step (4), performing an end repair reaction. The methods can also further comprise prior to step (4), performing a tailing reaction to attach a single nucleotide overhang to the 3' ends of the identifier oligonucleotide. The methods can further comprise, prior to step (4), performing an end repair reaction and a tailing reaction to attach a single nucleotide overhang to the 3' ends of the identifier oligonucleotide. The tailing reaction and the end repair reaction can be performed sequentially or concurrently.

[0014] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a first amplification primer binding site, and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the ligation product produced in step (4); and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0015] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises: a first amplification primer binding site and a unique nucleic acid sequence which identifies the target

analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises a nucleic acid sequence comprising a unique molecular identifier, a second amplification primer binding site, and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the extension product produced in step (4) using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0016] The nucleic acid adapter of step (4) can be a partially double-stranded nucleic acid molecule. A partially double-stranded nucleic acid adapter can comprise a double-stranded annealed region and a single-stranded mismatched region.

[0017] The nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter.

[0018] The constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter. The constant nucleic acid sequence can also comprise a cleavable moiety. The cleavable moiety can be an enzymatically cleavable moiety. The enzymatically cleavable moiety can be a USER sequence.

[0019] A unique molecular identifier can be present in the single-stranded mismatched region of a partially double-stranded nucleic acid adapter.

[0020] The second amplification primer binding site can be present in the single-stranded mismatched region of a partially double-stranded nucleic acid adapter.

[0021] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a first

amplification primer binding site, and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a single stranded nucleic acid template, wherein the nucleic acid template comprises a region complementary to the unique nucleic acid sequence of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier, a nucleic acid sequence complementary to a second amplification primer binding site and optionally, an affinity molecule; (5) extending the identifier oligonucleotide of step (4) to form an extension product complementary to the single stranded nucleic acid template, wherein the extension product comprises the identifier oligonucleotide, the nucleic acid sequence complementary to the unique molecular identifier, and the second amplification primer binding site; (6) amplifying the extension product produced in step (5) using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0022] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a first amplification primer binding site and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a single stranded nucleic acid template, wherein the nucleic acid template comprises a region complementary to the unique nucleic acid sequence of the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier, a nucleic acid

sequence complementary to a second amplification primer binding site and optionally, an affinity molecule; (5) extending the identifier oligonucleotide of step (4) to form an extension product complementary to the single stranded nucleic acid template, wherein the extension product comprises the identifier oligonucleotide, the nucleic acid sequence complementary to the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, the nucleic acid sequence complementary to the unique molecular identifier and the second amplification primer binding site; (6) amplifying the extension product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0023] The single stranded nucleic acid template can further comprise an affinity molecule. In aspects in which the single stranded nucleic acid template comprises an affinity molecule, the methods of the present disclosure described in the preceding can further comprise an affinity purification step between steps (4) and (5).

[0024] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding site, and wherein the first the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying

the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0025] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a second amplification primer binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0026] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier

oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding site, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0027] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a second amplification primer binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes

hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0028] In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released is located on the first nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be located 5' to the first amplification primer binding site.

[0029] In aspects in which a unique molecular identifier is located on the second nucleic acid probe, the unique molecular identifier can be located 3' to the second amplification primer binding site.

[0030] In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a unique molecular identifier are present in the first nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a unique molecular identifier can be located 5' to the first amplification primer binding site.

[0031] In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a unique molecular identifier are present in the second nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the unique molecular identifier can be located 3' to the second amplification primer binding site.

[0032] In aspects in which a unique molecular identifier is present in the first nucleic acid probe and the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released is present in the second nucleic acid probe, the unique molecular identifier can be located 5' to the first amplification primer binding site and the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be located 3' to the second amplification binding site.

[0033] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier, a nucleic acid sequence comprising a third unique molecular identifier and a second flow cell binding site, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0034] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the

identifier oligonucleotide, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier and a second flow cell binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a third unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0035] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier, a nucleic acid sequence comprising a third unique molecular identifier and a second flow cell binding site, and wherein the first and the second nucleic acid

probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0036] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier and a second flow cell binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a third unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0037] In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the first unique molecular identifier are present in the first nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the first unique molecular identifier can be located 5' to the first flow cell binding site.

[0038] In aspects in which the second and the third unique molecular identifiers are present in the second nucleic acid probe, the second and the third unique molecular identifiers can be located 3' to the second flow cell binding site.

[0039] In some aspects, the first unique molecular identifier can be present in the first nucleic acid probe and can be located 5' to the first flow cell binding site. In other aspects, the second unique molecular identifier can be present in the second nucleic acid probe and can be located 3' to the second flow cell binding site.

[0040] In some aspects, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the third unique molecular identifier can be present in the first nucleic acid probe and can be located 5' to the first flow cell binding site.

[0041] In some aspects, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the third unique molecular identifier can be present in the second nucleic acid probe and can be located 3' to the second flow cell binding site.

[0042] In some aspect, the third unique molecular identifier can be present in the first nucleic acid probe and can be located 5' to the first flow cell binding site. In this same aspect, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be present in the second nucleic acid probe and can be located 3' to the second flow cell binding site.

[0043] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding

domain, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, and a second amplification primer binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) amplifying the released identifier oligonucleotide using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; (5) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0044] The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a first amplification primer binding site and a second amplification primer binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) amplifying the released identifier oligonucleotide using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein at least one of the amplification primers comprises a nucleic acid sequence comprising a unique molecular identifier; and (5) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0045] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding

domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain and a capture probe binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a capture probe, wherein the capture probe comprises an affinity molecule and a region complementary to the capture probe binding site; and (5) Identifying the released identifier oligonucleotide by sequencing the amplified hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0046] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a capture probe binding site and a multiplexing probe binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a capture probe and a multiplexing probe, wherein the capture probe comprises an affinity molecule and a region complementary to the capture probe binding site, and wherein the multiplexing probe comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a region complementary to the multiplexing probe binding site; and (5) Identifying the released identifier oligonucleotide by sequencing the hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0047] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the

released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises: a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, and wherein the second nucleic acid probe comprises: a nucleic acid complementary to a portion of the identifier oligonucleotide, and a second amplification primer binding site, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) ligating the hybridized first and second nucleic acid probes together; (6) amplifying the ligation product produced in step (5); and (7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0048] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises: a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, and a second amplification primer binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) amplifying the collected identifier oligonucleotide; (5) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[0049] A In all methods of the present disclosure, the ligation process can be a nick ligation process. The nick ligation process can be a nick repair process.

[0050] In all methods of the present disclosure, the sequencing can be an enzyme free sequencing method.

[0051] In all methods of the present disclosure, the identifier oligonucleotide can be double-stranded. In aspects in which the identifier oligonucleotide is double-stranded, at least one of the two strands of the identifier oligonucleotide can comprise at least two separate nucleic acid

molecules. Alternatively, at least one 3' end of an identifier oligonucleotide can comprise a single nucleotide overhang.

[0052] In all methods of the present disclosure, the identifier oligonucleotide can be single-stranded.

[0053] In all methods of the present disclosure, the unique nucleic acid sequence which identifies the target analyte bound to a target binding domain can comprise between about 5 nucleotides and about 40 nucleotides preferably about 35 nucleotides, preferably still about 10 nucleotides.

[0054] In all methods of the present disclosure, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can comprise between about 6 nucleotides and about 15 nucleotides, preferably about 12 nucleotides, preferably still about 10 nucleotides.

[0055] In all methods of the present disclosure, at least one of a first nucleic acid probe or a second nucleic acid probe can comprise an affinity molecule. For example, at least one of a first nucleic acid probe or a second nucleic acid probe can comprise a biotin.

[0056] In all methods of the present disclosure, an amplification primer binding site can comprise between about 18 nucleotides and about 40 nucleotides, preferably about 32 nucleotides, preferably still about 25 nucleotides. An amplification primer binding site can comprise an i7 sequence, wherein the i7 sequence comprises the sequence set forth in SEQ ID NO: 1. An amplification primer binding site can comprise an i5 sequence, wherein the i5 sequence comprises the sequence set forth in SEQ ID NO: 2.

[0057] In all methods of the present disclosure, an amplification primer can comprise a flow cell adapter sequence, wherein the flow cell adapter sequence is suitable for sequencing. An amplification primer can comprise a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3. An amplification primer can comprise a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

[0058] In all methods of the present disclosure, a flow cell binding site can comprise a flow cell adapter sequence, wherein the flow cell adapter sequence is suitable for sequencing. A flow cell binding site can comprise a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3. A flow cell binding site can

comprise a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

[0059] In all methods of the present invention, at least one of the amplification primers can comprise an affinity molecule. For example, at least one of the amplification primers cam comprise a biotin.

[0060] In all methods of the present disclosure, amplification can comprise performing PCR. Performing PCR can comprise an amplification primer.

[0061] An amplification primer can comprise a flow cell binding site. An amplification primer can comprise a nucleic sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released. An amplification primer can comprise a nucleic acid sequence complementary to an amplification primer binding site.

[0062] Any of the above aspects can be combined with any other aspect.

[0063] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. In the Specification, the singular forms also include the plural unless the context clearly dictates otherwise; as examples, the terms "a," "an," and "the" are understood to be singular or plural and the term "or" is understood to be inclusive. By way of example, "an element" means one or more element. Throughout the specification the word "comprising," or variations such as "comprises" or "comprising," will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term "about."

[0064] Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. The references cited herein are not admitted to be prior art to the claimed invention. In the case of conflict, the present Specification, including definitions, will control. In addition, the materials, methods, and examples are

illustrative only and are not intended to be limiting. Other features and advantages of the disclosure will be apparent from the following detailed description and claim.

BRIEF DESCRIPTION OF THE DRAWINGS

[0065] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

[0066] The above and further features will be more clearly appreciated from the following detailed description when taken in conjunction with the accompanying drawings.

[0067] Figure 1 is a schematic of a two-ended adapter ligation method of the present disclosure.

[0068] Figure 2 is a schematic of a one-ended adapter ligation method of the present disclosure.

[0069] Figure 3 is a schematic of a templated-primer extension method of the present disclosure.

[0070] Figure 4 is a schematic of a template-extended identifier oligonucleotide of the present disclosure.

[0071] Figure 5 is a schematic of a short probe hybridization method of the present disclosure.

[0072] Figure 6 is a schematic of a short probe hybridization method of the present disclosure.

[0073] Figure 7 is a schematic of a short probe hybridization method of the present disclosure.

[0074] Figure 8 is a schematic of a long probe hybridization method of the present disclosure.

[0075] Figure 9 is a schematic of a long probe hybridization method of the present disclosure.

[0076] Figure 10 is a schematic of a long probe hybridization method of the present disclosure.

[0077] Figure 11 is a schematic of a direct-PCR method of the present disclosure.

[0078] Figure 12 is a schematic of an enzyme free method of the present disclosure.

[0079] Figure 13 is a schematic of a multiplexed enzyme free method of the present disclosure.

[0080] Figure 14 is a schematic of a probe of the present disclosure indirectly binding to a target nucleic acid.

[0081] Figure 15 is a schematic of an identifier oligonucleotide-short nucleic acid probe complex of the present disclosure.

[0082] Figure 16 is a schematic of a short probe hybridization method of the present disclosure.

[0083] Figure 17 is a schematic of an identifier oligonucleotide-short nucleic acid probe complex of the present disclosure.

[0084] Figure 18 is a schematic of a short probe hybridization method of the present disclosure.

[0085] Figure 19 is a schematic of a direct-PCR method of the present disclosure.

[0086] Figure 20 is a schematic overview of the methods of the present disclosure.

[0087] Figure 21 shows the spatial detection of protein target analytes using the methods of the present disclosure.

[0088] Figure 22 shows the spatial detection of RNA target analytes using the methods of the present disclosure.

[0089] Figure 23 shows the spatial detection of protein target analytes using the methods of the present disclosure.

[0090] Figure 24 shows the spatial detection of RNA target analytes using the methods of the present disclosure.

[0091] Figure 25 shows the spatial detection of protein target analytes using the methods of the present disclosure.

[0092] Figure 26 is a schematic of a probe of the present disclosure.

[0093] Figure 27 shows the use of probe tiling in the methods of the present disclosure.

[0094] Figure 28 shows the regions of interest selected on a tissue microarray.

[0095] Figure 29 is a series of graphs showing the read depth achieved using the methods of the present disclosure.

[0096] Figure 30 is a series of graphs showing the spatial detection of RNA target analytes in negative control samples using the methods of the present disclosure.

[0097] Figure 31 is a series of graphs showing the spatial detection of RNA target analytes in a HEK293 sample (top panel) and a Jurkat cell sample (bottom panel) using the methods of the present disclosure.

[0098] Figure 32 is a series of graphs showing the spatial detection of RNA target analytes in sixteen FFPE samples using the methods of the present disclosure.

[0099] Figure 33 is a graph showing the spatial detection of RNA target analytes in a HEK293 sample using the methods of the present disclosure.

[00100] Figure 34 is a graph showing the spatial detection of RNA target analytes in a Jurkat cell sample using the methods of the present disclosure.

DETAILED DESCRIPTION OF THE INVENTION

[00101] The present disclosure is based in part on probes, compositions, methods, and kits for simultaneous, multiplexed spatial detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell using existing sequencing methods.

[00102] The present disclosure provides a comparison of the identity and abundance of target proteins and/or target nucleic acids present in a first region of interest (*e.g.*, tissue type, a cell (including normal and abnormal cells), and a subcellular structure within a cell) and the identity and abundance of target proteins and/or target nucleic acids present in a second region of interest. There is no pre-defined upper limit to the number of regions of interest and comparisons that can be made; the upper limit relates to the size of the region of interest relative the size of the sample. As examples, when a single cell represents a region of interest, then a section may have hundreds to thousands of regions of interest; however, if a tissue section includes only two cell types, then the section may have only two regions of interest (each including only one cell type).

[00103] The present disclosure provides a higher degree of multiplexing than is possible with standard immunohistochemical or *in situ* hybridization methods. Standard immunohistochemical methods allow for maximal simultaneous detection of six to ten protein targets, with three to four protein targets being more typical. Similarly, *in situ* hybridization methods are limited to simultaneous detection of fewer than ten nucleic acid targets. The present disclosure provides detection of large combinations of nucleic acid targets and/or protein targets from a defined region of a sample. The present disclosure provides an increase in objective measurements by digital quantification and increased reliability and consistency, thereby enabling comparison of results among multiple centers.

[00104] Various compositions and methods of the present disclosure are described in full detail herein.

[00105] In one aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "two-ended adapter ligation method".

[00106] A two ended-adapter ligation method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure.

The probes and samples of the present disclosure are described in further detail herein. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain.

[00107] Following contacting the at least on target analyte with the at least one probe, a two-ended adapter ligation method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00108] Following release of the identifier oligonucleotide, a two-ended adapter ligation method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location, thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00109] Following collection of the released identifier oligonucleotide, a two-ended adapter ligation method can further comprise: (4) ligating to the released identifier oligonucleotide collected in step (3) at least one nucleic acid adapter.

[00110] The nucleic acid adapter can comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", the nucleic acid adapter would comprise a nucleic acid sequence that corresponds to "ROI #1".

[00111] The nucleic acid adapter can also comprise a unique molecular identifier.

[00112] The nucleic acid adapter can also comprise a first amplification primer binding site. In other aspects, the nucleic acid adapter can also comprise a second amplification primer binding site.

[00113] In some aspects, the nucleic acid adapter can also comprise a constant nucleic acid sequence to minimize ligation bias caused by differences in sequences of particular identifier oligonucleotides.

[00114] The nucleic acid adapter can be a partially double-stranded nucleic acid molecule. In aspects in which the nucleic acid adapter is partially double-stranded, the nucleic acid adapter comprises a double-stranded annealed region, a first single-stranded mismatched region, and a second single-stranded mismatched region. The first single-stranded mismatched region and the second single stranded mismatched region can be present on opposing sides of the double-stranded annealed region.

[00115] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of the nucleic acid adapter.

[00116] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a constant nucleic acid sequence to minimize ligation bias, the constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of the nucleic acid adapter.

[00117] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a unique molecular identifier, the unique molecular identifier can be present in at least one of the first or second single-stranded mismatched regions of the nucleic acid adapter.

[00118] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a first and a second amplification primer binding site, the first amplification primer binding site can be present in the first single-stranded mismatched region of the nucleic acid adapter and the second amplification primer binding site can be present in the second single-stranded mismatched region of the nucleic acid adapter.

[00119] After ligation of the at least one nucleic acid adapter, a two-ended adapter ligation method can further comprise: (5) amplifying the ligation product produced in step (4) using

amplification primers that bind to the first and second amplification primer binding sites; and (6) identifying the released oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the sample.

[00120] A two-ended adapter ligation method of the present disclosure can further comprise, prior to step (4), performing an end repair reaction using methods known in the art. The method can also further comprise, prior to step (4), performing a tailing reaction to attach a single nucleotide overhang to the 3' ends of the identifier oligonucleotide using methods known in the art. In aspects, the end repair reaction and the tailing reaction can be performed sequentially or concurrently.

[00121] In preferred aspects of a two-ended adapter ligation method, a nucleic acid adapter is ligated to both ends of the released and collected identifier oligonucleotide.

[00122] In other aspects of a two-ended adapter ligation method, at least one of the amplification primers used in step (5) to amplify the ligation product produced in step (4) comprises a nucleic acid sequence which identifies the specific location of the tissue sample form which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", at least one of the amplification primers would comprise a nucleic acid sequence that corresponds to "ROI #1".

[00123] Figure 1 shows a schematic of a preferred aspect of a two-ended adapter ligation method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising an antibody that binds to a target protein. In the upper left panel, the probe binds to the target protein. In the upper right panel, a UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target protein bound to the target binding domain. In the bottom panel, a nucleic acid adapter is ligated to both ends of the identifier oligonucleotide. In this non-limiting example, the nucleic acid adapter is partially double-stranded and comprises a double-stranded annealed region, a first single-stranded mismatched region, and a second single-stranded mismatched region. Present in the double-stranded annealed region is a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. Present in the first single-stranded mismatched region is a first amplification primer binding site. Present in the second single-stranded mismatched region

is a unique molecular identifier and a second amplification primer binding site. Following ligation of the nucleic acid adapters to the identifier oligonucleotide, the product is amplified using amplification primers that bind the first and the second amplification primer binding sites and sequenced to identify the target protein bound by the probe.

[00124] In one aspect, the present disclosure provides a composition of an identifier oligonucleotide dually ligated to two nucleic acid adapters for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide dually ligated to two nucleic acid adapters comprises an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which is capable of identifying a target analyte in a sample. Each end of the identifier oligonucleotide is attached to a nucleic acid adapter molecule, wherein the nucleic acid adapter molecule is partially double-stranded and comprises a double-stranded annealed region, a first single-stranded mismatched region and a second single-stranded mismatched region. The first single-stranded mismatched region and the second single stranded mismatched region are present on opposing sides of the double-stranded annealed region. The double-stranded mismatch region comprises a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence nucleic acid sequence which is capable of identifying a specific location of a sample. The first single-stranded mismatched region comprises a first amplification primer binding site. The second single-stranded mismatched region comprises a second amplification primer binding site and a nucleic acid sequence comprising a unique molecular identifier. A schematic of an identifier oligonucleotide dually ligated to two nucleic acid adapters is shown in the bottom panel of Figure 1.

[00125] In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "one-ended adapter ligation method".

[00126] A one-ended adapter ligation method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid

of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a first amplification primer binding site. In some aspects, the identifier oligonucleotide also comprises at least one 3' end with a single nucleotide overhang.

[00127] Following contacting the at least one target analyte with the at least one probe, a one-ended adapter ligation method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00128] Following release of the identifier oligonucleotide, a one-ended adapter ligation method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00129] Following collection of the released identifier oligonucleotide, a one-ended adapter ligation method can further comprise: (4) ligating to the released oligonucleotide collected in step (3) at least one nucleic acid adapter;

[00130] The nucleic acid adapter can comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", the nucleic acid adapter would comprise a nucleic acid sequence that corresponds to "ROI #1". The nucleic acid adapter can also comprise a unique molecular identifier. The nucleic acid adapter can also comprise a second amplification primer binding site.

[00131] In some aspects, the nucleic acid adapter can also comprise a constant nucleic acid sequence to minimize ligation bias caused by differences in sequences of particular identifier oligonucleotides. The constant nucleic acid sequence can comprise a cleavable moiety. The cleavable moiety can be enzymatically cleavable. In a non-limiting example, the enzymatically

cleavable moiety can be a USER sequence, wherein the USER sequence comprises the sequence GUGUATUG.

[00132] The nucleic acid adapter can comprise any combination of the features described above.

[00133] The nucleic acid adapter can be a partially double-stranded nucleic acid molecule. In aspects in which the nucleic acid adapter is partially double-stranded, the nucleic acid adapter comprises a double-stranded annealed region and a single-stranded mismatched region.

[00134] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of the nucleic acid adapter.

[00135] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a constant nucleic acid sequence to minimize ligation bias, the constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of the nucleic acid adapter.

[00136] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a unique molecular identifier, the unique molecular identifier can be present in the single-stranded mismatched region.

[00137] In aspects in which the nucleic acid adapter is partially double-stranded and comprises a second amplification primer binding site, the second amplification primer binding site can be present in the single-stranded mismatched region of the nucleic acid adapter.

[00138] After ligation of the at least one nucleic acid adapter, a two-ended adapter ligation method can further comprise: (5) amplifying the ligation product produced in step (4); and (6) identifying the released oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the sample.

[00139] In other aspects of a one-ended adapter ligation method of the present disclosure, at least one of the amplification primers used in step (5) to amplify the ligation product produced in step (4) comprises a nucleic acid sequence which identifies the specific location of the tissue sample form which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated “ROI #1”, at least one of the amplification primers would comprise a nucleic acid sequence that corresponds to “ROI #1”.

[00140] Figure 2 shows a schematic of a preferred aspect of a one-ended adapter ligation method of the present disclosure. In this aspect, the probe comprises a target binding domain that is an antibody that binds to a target protein. In the upper left panel, the probe binds to the target protein. In upper right panel, the UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target protein and a first amplification primer binding site. In this non-limiting example, the identifier oligonucleotide is double-stranded with one strand that comprises three separate nucleic acid molecules. The identifier oligonucleotide also comprises one 3' end with a single nucleotide overhang.

[00141] In the bottom panel of Figure 2, a nucleic acid adapter is ligated to the end of the identifier oligonucleotide that comprises the 3' single nucleotide overhang. In this non-limiting example, the nucleic acid adapter is partially double-stranded and comprises a double-stranded annealed region and a single-stranded mismatched region. Present in the double-stranded annealed region is a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. Present in the single-stranded mismatched region is a unique molecular identifier and a second amplification primer binding site. Following ligation of the nucleic acid adapter to the identifier oligonucleotide, the product is amplified using amplification primers that bind to the first and the second amplification primer binding sites and sequenced to identify the target protein bound by the probe.

[00142] In one aspect, the present disclosure provides a composition of an identifier oligonucleotide ligated to one nucleic acid adapter for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide ligated to one nucleic acid adapter comprises an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which is capable of identifying a target analyte in a sample and a first amplification primer binding site. One end of the identifier oligonucleotide is attached to a nucleic acid adapter molecule, wherein the nucleic acid adapter molecule is partially double-stranded and comprises a double-stranded annealed region and a single-stranded mismatched region and a second single-stranded mismatched region. The double-stranded mismatch region comprises a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence nucleic acid sequence which is

capable of identifying a specific location of a sample. The single-stranded mismatched region comprises a second amplification primer binding site. A schematic of an identifier oligonucleotide ligated to one nucleic acid adapter is shown in the bottom panel of Figure 2.

[00143] In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "templated-primer extension method".

[00144] A templated-primer extension method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a first amplification primer binding site.

[00145] Following contacting the at least one target analyte with the at least one probe, a templated-primer extension method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00146] Following release of the identifier oligonucleotide, a templated-primer extension method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00147] Following collection of the released identifier oligonucleotide, a templated-primer extension method can further comprise: (4) hybridizing to the released identifier oligonucleotide collected in step (3) a single stranded nucleic acid template.

[00148] The single stranded nucleic acid template can comprise a region complementary to the unique nucleic acid sequence of the identifier oligonucleotide, thereby allowing for the hybridization of the single stranded nucleic acid template and the collected identifier oligonucleotide.

[00149] The single stranded nucleic acid template can also comprise a nucleic acid sequence comprising a unique molecular identifier.

[00150] The single stranded nucleic acid template can also comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released.

[00151] The single stranded nucleic acid template can also comprise a nucleic acid sequence that is complementary to a second amplification primer binding site.

[00152] The single stranded nucleic acid template can comprise any combination of the features described above.

[00153] Following hybridization of the identifier oligonucleotide to the single stranded nucleic acid template, a templated-primer extension method can further comprise: (5) extending the identifier oligonucleotide of step (4) to form an extension produce complementary to the single stranded nucleic acid template, wherein the extension product comprises the identifier oligonucleotide and the sequence complementary to the single stranded nucleic acid template; (6) amplifying the extension product of step (6) using amplification primers that hybridize to the first and second amplification primer binding sites; and (7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the sample.

[00154] In some aspects, the single stranded nucleic acid template can comprise an affinity molecule. In aspects in which the single stranded nucleic acid template comprises an affinity molecule, a templated-primer extension method can further comprise an affinity purification step between steps (4) and (5).

[00155] Figure 3 shows a schematic of a preferred aspect of a templated-primer extension method of the present disclosure. In this aspect, the probe comprises a target binding domain that is an

antibody that binds to a target protein. In the upper left panel, the probe binds to the target protein. In upper right panel, the UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target protein and a first amplification primer binding site. In the lower right panel, the identifier oligonucleotide is hybridized to a single stranded nucleic acid template. In this non-limiting example, the single-stranded nucleic acid template comprises an affinity molecule, a nucleic acid sequence complementary to the unique nucleic acid sequence of the identifier oligonucleotide, a first unique molecular identifier, and a sequence complementary to a second amplification primer biding site. The identifier oligonucleotide is extended to form an extension product complementary to the single stranded nucleic acid template. As shown in the lower left panel, the extension product comprises the identifier oligonucleotide, the nucleic acid sequence complementary to the first unique molecular identifier, and the second amplification primer binding site. Following the extension reaction, the primer extension product is amplified using amplification primers that bind to the first and the second amplification primer binding sites. In this non-limiting example, one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. The amplified product is then sequenced to identify the target protein bound by the probe.

[00156] In one aspect, the present disclosure provides a composition of a template-extended identifier oligonucleotide for spatially detecting at least one target analyte in a sample. The template-extended identifier oligonucleotide comprises a first flow cell adapter sequence suitable for sequencing, followed by a first unique molecular identifier, followed by an identifier oligonucleotide, followed by a second unique molecular identifier, followed by a second amplification primer binding site, followed by a third unique molecular identifier, followed by a second flow cell adapter sequence suitable for sequencing. The identifier oligonucleotide comprises a first amplification primer binding site and a unique nucleic acid sequence which is capable of identifying a target analyte in a sample. A schematic of a template-extended identifier oligonucleotide is shown in the bottom panel of Figure 4.

[00157] In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "short probe hybridization method".

[00158] A short probe hybridization method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain.

[00159] Following contacting the at least one target analyte with the at least one probe, a short probe hybridization method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00160] Following release of the identifier oligonucleotide, a short probe hybridization method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00161] Following collection of the released identifier oligonucleotide, a short probe hybridization method can further comprise: (4) hybridizing to the release identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe.

[00162] The first or the second nucleic acid probe can comprise a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The first or the second nucleic acid

probe can also comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. The first or the second nucleic acid probe can also comprise a nucleic acid sequence comprising unique molecular identifier. The first nucleic acid probe can comprise a first amplification primer binding site. The second nucleic acid probe can comprise a second amplification primer binding site.

[00163] The first or the second nucleic acid probe can comprise any combination of the features described above. In a preferred aspect depicted in Figure 5, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released is located 5' to the first amplification primer binding site and the unique molecular identifier is located 3' to the second amplification primer binding site.

[00164] In another preferred aspect depicted in Figure 6, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and the unique molecular identifier are located 5' to the first amplification primer binding site.

[00165] In another preferred aspect depicted in Figure 7, the first nucleic acid probe comprises a first amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, a

nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and the unique molecular identifier are located 3' to the second amplification primer binding site.

[00166] In another preferred aspect depicted in Figure 15, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence comprising a unique molecular identifier is located 3' to the first amplification binding site.

[00167] In another preferred aspect depicted in Figure 17, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence comprising a unique molecular identifier is located 5' to the first amplification binding site.

[00168] The first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Alternatively, the first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping.

[00169] Following hybridization of the first and the second nucleic acid probe to the identifier oligonucleotide, a short probe hybridization method can further comprise: (5) in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping, ligating the first and the second nucleic acid probes together, for example, by performing a nick repair reaction. Alternatively, in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes

are not adjacent and are not overlapping, the method comprises ligating the first and the second nucleic acid probes together, for example, by performing a gap extension reaction and a nick repair reaction, such that the first and the second nucleic acid probes are ligated together.

[00170] Following ligation of the first and the second nucleic acid probe, a short probe hybridization method can further comprise: (6) amplifying the ligation product produced in step (5) using amplification primers that hybridize to the first and second amplification primer binding sites; and (7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the sample.

[00171] In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-short nucleic acid probe complex for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide-short nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target analyte in a sample. The first nucleic acid probe comprises a first amplification primer binding site, followed by a unique nucleic acid sequence capable of identifying a specific location in a sample, followed by a region complementary to the identifier oligonucleotide. The second nucleic acid probe comprises a second amplification primer binding site, followed by a nucleic acid sequence comprising a unique molecular identifier, followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 5.

[00172] In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-short nucleic acid probe complex for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide-short nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target analyte in a sample. The first nucleic acid probe comprises a first amplification primer binding site, followed by a nucleic acid sequence comprising a unique molecular identifier, followed by a region complementary to the identifier oligonucleotide, wherein the nucleic acid sequence comprising a unique molecular identifier is located 3' to the first amplification primer binding site. The second nucleic acid probe comprises a second

amplification primer binding site followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 15.

[00173] In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-short nucleic acid probe complex for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide-short nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target analyte in a sample. The first nucleic acid probe comprises a first amplification primer binding site, followed by a nucleic acid sequence comprising a unique molecular identifier, followed by a region complementary to the identifier oligonucleotide, wherein the nucleic acid sequence comprising a unique molecular identifier is located 5' to the first amplification primer binding site. The second nucleic acid probe comprises a second amplification primer binding site followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 17.

[00174] Figure 16 shows a schematic overview of an exemplary short probe hybridization method of the present disclosure. First at least one target analyte in a sample is contacted with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target domain.

[00175] Following contacting the at least one target analyte with the at least one probe, a force is then provided to a location of the sample sufficient to release the identifier oligonucleotide. The identifier oligonucleotide is collected following release, as shown in the top panel of Figure 16.

[00176] As shown in the second panel from the top of Figure 16, the released identifier oligonucleotide is then hybridized to a first nucleic acid probe and a second nucleic acid probe.

In this non-limiting example, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The nucleic acid sequence comprising the unique molecular identifier is located 3' to the first amplification primer binding site. The second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this non-limiting example, the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Following hybridization to the identifier oligonucleotide, the first and second probe are ligated together, for example, by performing a nick repair reaction.

[00177] Following ligation of the first and second nucleic acid probes, the ligation product is amplified via PCR using amplification primers that hybridize to the first and second amplification primer binding sites. As shown in the second panel from the bottom of Figure 16, the amplification primer that hybridizes to the second amplification primer binding site comprises a first flow cell binding site, a first nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the second amplification primer binding site. The amplification primer that hybridizes to the first amplification primer binding site comprises a second flow cell binding site, a second nucleic acid sequence which identifies the specific location from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the first amplification primer binding site. The PCR product shown in the bottom panel of Figure 16 is then sequenced to identify the released oligonucleotide, thereby spatially detecting the at least one target analyte in the sample.

[00178] Figure 18 shows a schematic overview of an exemplary short probe hybridization method of the present disclosure. First, at least one target analyte in a sample is contacted with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier

oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target domain.

[00179] Following contacting the at least one target analyte with the at least one probe, a force is then provided to a location of the sample sufficient to release the identifier oligonucleotide. The identifier oligonucleotide is collected following release, as shown in the top panel of Figure 18.

[00180] As shown in the second panel from the top of Figure 18, the released identifier oligonucleotide is then hybridized to a first nucleic acid probe and a second nucleic acid probe. In this non-limiting example, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The nucleic acid sequence comprising the unique molecular identifier is located 5' to the first amplification primer binding site. The second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this non-limiting example, the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Following hybridization to the identifier oligonucleotide, the first and second probe are ligated together, for example, by performing a nick repair reaction.

[00181] Following ligation of the first and second nucleic acid probes, the ligation product is amplified via PCR using amplification primers that hybridize to the first and second amplification primer binding sites. As shown in the second panel from the bottom of Figure 18, the amplification primer that hybridizes to the second amplification primer binding site comprises a first flow cell binding site, a first nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the second amplification primer binding site. The amplification primer that hybridizes to the first amplification primer binding site comprises a second flow cell binding site, a second nucleic acid sequence which identifies the specific location from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the first amplification primer binding site. The PCR product shown in the bottom panel Figure 18 is then sequenced to identify the released oligonucleotide, thereby spatially detecting the at least one target analyte in the sample.

[00182] In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "long probe hybridization method".

[00183] A long probe hybridization method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain.

[00184] Following contacting the at least one target analyte with the at least one probe, a long probe hybridization method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00185] Following release of the identifier oligonucleotide, a long probe hybridization method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00186] Following collection of the released identifier oligonucleotide(s), a long probe hybridization method can further comprise: (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe.

[00187] The first or the second nucleic acid probe can comprise a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The first or the second nucleic acid

probe can also comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released.

[00188] The first or the second nucleic acid probe can also comprise a first unique molecular identifier. The first or the second nucleic acid probe can also comprise a second unique molecular identifier. The first or the second nucleic acid probe can also comprise a third unique molecular identifier.

[00189] The first nucleic acid probe can comprise a first amplification primer binding site.

[00190] The first nucleic acid probe can also comprise a first flow cell binding site. The second nucleic acid probe can comprise a second flow cell binding site.

[00191] The first and the second nucleic acid probes can comprise any combination of the features described above. In a preferred aspect depicted in Figure 8, the first nucleic acid probe comprises a first flow cell binding site, a first unique molecular identifier, a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second flow cell binding site, a second unique molecular identifier, a third unique molecular identifier and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and the first unique molecular identifier are located 5' to the first flow cell binding site and the second and the third unique molecular identifiers are located 3' to the second flow cell binding site.

[00192] In another preferred aspect depicted in Figure 9, the first nucleic acid probe comprises a first flow cell binding site, a first unique molecular identifier, a second unique molecular identifier, a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second flow cell binding site, a third unique molecular identifier and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the first unique molecular identifier, the second unique molecular identifier and the nucleic acid sequence which identifies the specific location of the sample from

which the identifier oligonucleotide was released are located 5' to the first flow cell binding site and the third unique molecular identifier is located 3' to the second flow cell binding site.

[00193] In another preferred aspect depicted in Figure 10, the first nucleic acid probe comprises a first flow cell binding site, a first unique molecular identifier, a first amplification primer binding site and a nucleic acid sequence complementary to the identifier oligonucleotide. In this same preferred aspect, the second nucleic acid probe comprises a second flow cell binding site, a second unique molecular identifier, a third unique molecular identifier, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the first unique molecular identifier is located 5' to the first flow cell binding site and the second unique molecular identifier, the third unique molecular identifier, and the nucleic acid sequence which identifies the specific location of the sample form which the identifier oligonucleotide was released are located 3' to the second amplification primer binding site.

[00194] The first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Alternatively, the first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping.

[00195] Following hybridization of the first and the second nucleic acid probe to the identifier oligonucleotide, a long probe hybridization method can further comprise: (5) in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping, performing a nick repair reaction such that the first and the second nucleic acid probes are ligated together. Alternatively, in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping, the method comprises performing a gap extension and a nick repair reaction such that the first and the second nucleic acid probes are ligated together.

[00196] The name method can further comprise: (6) amplifying the ligation product produced in step (5) using amplification primers that hybridize to the first and second amplification primer binding sites; and (7) identifying the released identifier oligonucleotide by sequencing the

amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the sample.

[00197] In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-long nucleic acid probe complex for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide-long nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target analyte in a sample. The first nucleic acid probe comprises a first flow cell binding site suitable for sequencing, followed by a first unique molecular identifier, followed by a first amplification primer binding site, followed by a unique nucleic acid sequence capable of identifying a specific location in a sample, followed by a region complementary to the identifier oligonucleotide. The second nucleic acid probe comprises a second flow cell binding site, followed by a second unique molecular identifier, followed by a third unique molecular identifier, followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 8.

[00198] In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "direct PCR method".

[00199] A direct PCR method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a first amplification primer binding site, a second amplification primer binding site, or a unique molecular identifier. The identifier oligonucleotide can comprise any combination of these features. Any of these features

can also be flanked by regions comprising constant nucleic acid sequences of about 1 nucleotide to about 10 nucleotides.

[00200] Following contacting the at least one analyte with the at least one probe, a direct PCR method further comprises: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00201] A direct PCR method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00202] Following release of the identifier oligonucleotide, a direct PCR method can further comprise: (4) amplifying the released identifier oligonucleotide using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site. In some aspects, at least one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", at least one of the amplification primers would comprise a nucleic acid sequence that corresponds to "ROI #1". In still other aspects, at least one of the amplification primers comprises a unique molecular identifier.

[00203] Following amplification, a direct PCR method of the present disclosure can further comprise: (5) identifying the released oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the sample.

[00204] Figure 11 shows a schematic of a preferred aspect of a direct PCR method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising a nucleic acid sequence that is complementary to a target nucleic acid. In the upper panel, the probe hybridizes to the target nucleic acid. In the lower panel, a UV photo-cleavable linker located between the

target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a first amplification primer binding site, a second amplification primer binding site, a unique molecular identifier, and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. Located between these four features are constant spacer regions that are 3 nucleotides in length. The identifier oligonucleotide is double-stranded and comprises a strand that comprises 3 separate nucleic acid molecules. After release, the identifier oligonucleotide is amplified using a first amplification primer that hybridizes to the first amplification primer binding site and comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a second amplification primer that hybridizes to the second amplification primer binding site. The amplified product is then sequenced to identify the target nucleic acid bound by the probe.

[00205] Figure 19 shows a schematic of a preferred aspect of a direct PCR method of the present disclosure. In this aspect, the identifier oligonucleotide comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain and a second amplification primer binding site, as shown in the top panel of Figure 19. The identifier oligonucleotide is amplified using amplification primers that hybridize to the first and second amplification primer binding sites. As shown in the middle panel of Figure 19, the amplification primer that hybridizes to the second amplification primer binding site comprises a first flow cell binding site, a first nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the second amplification primer binding site. The amplification primer that hybridizes to the first amplification primer binding site comprises a second flow cell binding site, a second nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the first amplification primer binding site. The PCR product shown in the bottom panel of Figure 19 is sequenced to identify the released oligonucleotide, thereby spatially detecting the at least one target analyte in the sample.

[00206] In one aspect, the present disclosure provides a composition of a direct-PCR compatible identifier oligonucleotide for spatially detecting at least one target analyte in a sample. A direct-

PCR compatible identifier oligonucleotide comprises a first amplification primer binding site, followed a unique nucleic acid sequence which is capable of identifying a target analyte in a sample, followed by a unique molecular identifier, followed by a second amplification primer binding site. A schematic of a direct-PCR compatible identifier oligonucleotide is depicted in the lower panel of Figure 11.

[00207] In one aspect, the present disclosure provides a composition of a direct-PCR compatible identifier oligonucleotide for spatially detecting at least one target analyte in a sample. A direct-PCR compatible identifier oligonucleotide comprises a first amplification primer binding site, followed a nucleic acid sequence comprising a unique molecular identifier, followed by a unique nucleic acid sequence which is capable of identifying a target analyte in a sample, followed by a second amplification primer binding site. A schematic of a direct-PCR compatible identifier oligonucleotide is shown in the top panel of Figure 19.

[00208] In some aspects of the methods of the present disclosure at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten, or at least eleven, or at least twelve, or at least thirteen, or at least fourteen, or at least fifteen, or at least sixteen, or at least seventeen, or at least eighteen, or at least nineteen, or at least twenty, or at least thirty, or at least forty, or at least fifty, or at least sixty, or at least seventy, or at least eighty, or at least ninety, or at least one hundred probes can been to a single target analyte. As used herein, the term "tiling" is used to describe when more than one probe of the present disclosure is bound to a target analyte. The top panel of Figure 27 shows the tiling of probes onto a target RNA. Tiling multiple probes onto a target analyte means that each target analyte will be individually detected multiple times, increasing the overall accuracy of the measurement. In a non-limiting example, as shown in the bottom panel of Figure 27, in the case where 10 probes are tiled onto a single target RNA, one of the probes may be incorrectly detected too many times (outlier high count probe), while another probe may be incorrectly detected too few times (outlier low count probe). However, the other 8 probes may be detected at a similar level, indicating that the two outliers should be discarded during analysis and the signals from the 8 probes used to generate a more accurate measurement of the abundance of the target RNA.

[00209] The present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as an "enzyme free method".

[00210] An enzyme free method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. An identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. An identifier oligonucleotide can also comprise a capture probe binding site.

[00211] Following contacting at least one target analyte with at least one probe, an enzyme free method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00212] Following release of an identifier oligonucleotide, an enzyme free method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00213] Following collection of a released identifier oligonucleotide, an enzyme free method can further comprise: (4) hybridizing to a released identifier oligonucleotide a capture probe.

[00214] A capture probe can comprise a region complementary to the capture probe binding site. A capture probe can also comprise an affinity molecule.

[00215] Following hybridization of a capture probe, an enzyme free method can further comprise: (5) Identifying a released identifier oligonucleotide by sequencing the hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[00216] A hybridized product produced in step (4) can be sequenced using an enzyme free method of sequencing. Enzyme-free methods of sequencing have been described in, e.g., US2014946386 and US15/819,151, each of which is incorporated herein by reference in its entirety.

[00217] Figure 12 shows a schematic of a preferred aspect of an enzyme free method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising a nucleic acid sequence that is complementary to a target nucleic acid. In the top panel, the probe hybridizes to the target nucleic acid. In the middle panel, a UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target nucleic acid bound to the target binding domain and a capture probe binding site. After release, the identifier oligonucleotide is hybridized to a capture probe, as depicted in the bottom panel. The capture probe comprises a nucleic acid sequence complementary to the capture probe binding site and an affinity molecule. The hybridized product is then sequenced using enzyme free sequencing methods to identify the target nucleic acid bound by the probe.

[00218] In one aspect, the present disclosure provides a composition of a hybridized identifier oligonucleotide-capture probe complex for spatially detecting at least one target analyte in a sample. A hybridized identifier oligonucleotide-capture probe complex comprises an identifier oligonucleotide hybridized to a capture probe. An identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a specific a target analyte in a sample and a capture probe binding site. A capture probe comprises an affinity molecule and a region complementary to the capture probe binding site. A schematic of a hybridized identifier oligonucleotide-capture probe complex is depicted in the bottom of panel Figure 12.

[00219] The present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "multiplexed enzyme free method".

[00220] A multiplexed enzyme free method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a capture probe binding site. The identifier oligonucleotide can also comprise a multiplexing probe binding site.

[00221] Following contacting the at least one target analyte with the at least one probe, a multiplexed enzyme free method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

[00222] Following release of the identifier oligonucleotide, a multiplexed enzyme free method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

[00223] Following collection of the released identifier oligonucleotide, a multiplexed enzyme free method can further comprise: (4) hybridizing to the released identifier oligonucleotide a capture probe and a multiplexing probe.

[00224] A capture probe can comprise a region complementary to the capture probe binding site. A capture probe can also comprise an affinity molecule.

[00225] A multiplexing probe can comprise a region complementary to the multiplexing probe binding site. A multiplexing probe can also comprise a nucleic acid sequence which identifies

the specific location of the sample from which the identifier oligonucleotide was released and a region complementary to the multiplexing probe binding site.

[00226] Following hybridization of a capture probe and a multiplexing probe, a multiplexed enzyme free method can further comprise: (5) Identifying the released identifier oligonucleotide by sequencing the hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

[00227] A hybridized product produced in step (4) can be sequenced using an enzyme free method of sequencing. Enzyme-free methods of sequencing have been described in, e.g., US2014946386 and US15/819,151, each of which is incorporated herein by reference in its entirety.

[00228] Figure 13 shows a schematic of a preferred aspect of a multiplexed enzyme free method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising a nucleic acid sequence that is complementary to a target nucleic acid. In the top panel, the probe hybridizes to the target nucleic acid. In the middle panel, a UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a capture probe binding site, and a multiplexing probe binding site, as shown in the middle panel. After release, the identifier oligonucleotide is hybridized to a capture probe and a multiplexing probe as shown in the lower panel. The capture probe comprises a nucleic acid sequence complementary to the capture probe binding site and an affinity molecule. The multiplexing probe comprises a nucleic acid sequence complementary to the multiplexing probe binding site and a nucleic acid sequence which identifies the specific location of a sample from which the identifier oligonucleotide was released. The hybridized product is then sequenced using enzyme free sequencing methods to identify the target nucleic acid bound by the probe.

[00229] In one aspect, the present disclosure provides a composition of a hybridized identifier oligonucleotide-capture probe-multiplex probe complex for spatially detecting at least one target analyte in a sample. A hybridized identifier oligonucleotide-capture probe-multiplex probe complex comprises an identifier oligonucleotide hybridized to a capture probe and a multiplexing probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a specific a target analyte in a sample, a capture probe binding site and a

multiplexing probe binding site. The capture probe comprises an affinity molecule and a region complementary to the capture probe binding site. The multiplexing probe comprises a nucleic acid sequence which identifies the specific location of a sample from which the identifier oligonucleotide was released and a region complementary to the multiplexing probe binding site. A schematic of a hybridized identifier oligonucleotide-capture probe-multiplex probe complex is depicted in the bottom panel of Figure 13.

[00230] Figure 20 is an exemplary schematic of overview of the methods of the present disclosure. First, a sample on a microscope slide is contacted with a plurality of probes of the present disclosure (step 1 in Figure 20). The slide is then imaged and particular regions of interest (ROIs) are selected (step 2 in Figure 20). A specific ROI is then illuminated by UV light to release identifier oligonucleotides from probes bound within the ROI. The released identifier oligonucleotides are then collected via aspiration with a microcapillary. Following aspiration, the identifier oligonucleotides are transferred to a particular well within a 96 well plate. Steps 4 and 5 are then repeated for each ROI identified in step 2. After all ROIs have been illuminated and all released identifier oligonucleotides collected, the identifier oligonucleotides are sequenced using next generation sequencing methods to spatially detect at least one target analyte in the sample.

[00231] As described in the preceding, the present disclosure provides probes for the compositions and methods of spatially detecting at least one target analyte in a sample. The present disclosure provides probes comprising a target binding domain and an identifier oligonucleotide. The target binding domain is a region of the probe that specifically binds to at least one target analyte in a sample.

[00232] Probes of the present disclosure can be used for spatially detecting a target nucleic acid. In this aspect, the target binding domain can be a target nucleic acid-binding region. The target nucleic acid-binding region is preferably at least 15 nucleotides in length, and more preferably is at least 20 nucleotides in length. In specific aspects, the target nucleic acid-binding region is approximately 10 to 500, 20 to 400, 25, 30 to 300, 35, 40 to 200, or 50 to 100 nucleotides in length. Probes and methods for binding and identifying a target nucleic acid have been described in, *e.g.*, US2003/0013091, US2007/0166708, US2010/0015607, US2010/0261026, US2010/0262374, US2010/0112710, US2010/0047924, and US2014/0371088, each of which is incorporated herein by reference in its entirety.

[00233] The target nucleic acid-binding region can directly hybridize to a target nucleic acid present in a sample. Alternatively, the probes of the present disclosure can indirectly hybridize to a target nucleic acid present in a sample (via an intermediary oligonucleotide). Figure 14 illustrates a probe (or composition) of this aspect. The probe includes a target nucleic-acid binding domain which binds to a synthetic oligonucleotide (the intermediary oligonucleotide) that in turn binds to a target nucleic acid in a biological sample. It could be said that the intermediary oligonucleotide is a probe, as defined herein, since it comprises a nucleic acid backbone and is capable of binding a target nucleic acid. In these aspects, a probe's target nucleic acid-binding region hybridizes to a region of an intermediary oligonucleotide (*i.e.*, a synthetic oligonucleotide) which is different from the target nucleic acid present in a sample. Thus, the probe's target binding region is independent of the ultimate target nucleic acid in the sample. This allows economical and rapid flexibility in an assay design, as the target (present in a sample)-specific components of the assay are included in inexpensive and widely-available synthetic DNA oligonucleotides rather than the more expensive probes. Such synthetic oligonucleotides are simply designed by including a region that hybridizes to the target nucleic acid present in a sample and a region that hybridizes to a probe. Therefore, a single set of indirectly-binding probes can be used to detect an infinite variety of target nucleic acids (present in a sample) in different experiments simply by replacing the target-specific (synthetic) oligonucleotide portion of the assay.

[00234] A target nucleic acid may be DNA or RNA and preferably messenger RNA (mRNA) or miRNA.

[00235] Probes of the present disclosure can be used for detecting a target protein. In this aspect, the target binding domain can be a target protein-binding region. A target protein-binding region includes molecules or assembles that are designed to bind to at least one target protein, at least one target protein surrogate, or both and can, under appropriate conditions, form a molecular complex comprising the probe and the target protein. The target-protein binding region can include an antibody, a peptide, an aptamer, or a peptoid. The antibody can be obtained from a variety of sources, including but not limited to polyclonal antibody, monoclonal antibody, monospecific antibody, recombinantly expressed antibody, humanized antibody, plantibodies, and the like. The terms protein, polypeptide, peptide, and amino acid sequence are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be

linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids or synthetic amino acids. The terms also encompass an amino acid polymer that has been modified, for example, by disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. As used herein the term amino acid refers to either natural and/or unnatural or synthetic amino acids, including but not limited to glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. Probes and methods for binding and identifying a target protein have been described, *e.g.*, in US2011/0086774, the contents of which is incorporated herein by reference in its entirety.

[00236] An identifier oligonucleotide is a nucleic acid molecule that identifies the target analyte bound to the target binding domain. The identifier oligonucleotide comprises a unique nucleic acid sequence that identifies the target analyte bound to the target binding domain of the probe. In a non-limiting example, a probe with a target binding domain that binds to the protein P53 comprises an identifier oligonucleotide with a unique nucleic acid sequence that corresponds to P53, while a probe with a target binding domain that binds to the protein P97 comprises an identifier oligonucleotide with a unique nucleic acid sequence that corresponds to P97.

[00237] An identifier oligonucleotide can be DNA, RNA, or a combination of DNA and RNA.

[00238] In some aspects, an identifier oligonucleotide can comprise at least one amplification primer binding site. An amplification primer binding site is a nucleic acid sequence capable of binding to an amplification primer. An amplification primer can be used to amplify the nucleic molecule to which it is bound using methods known in the art, including, but not limited to, polymerase chain reaction (PCR).

[00239] In some aspects, an identifier oligonucleotide can comprise at least one unique molecular identifier.

[00240] An identifier oligonucleotide can be a single-stranded, a double-stranded, or a partially double-stranded nucleic acid molecule. In the aspects in which an identifier oligonucleotide is double-stranded or partially double-stranded, at least one of the two strands can comprise at least two separate nucleic acid molecules which, without being bound by theory, allows for denaturing of the identifier oligonucleotide at lower temperatures.

[00241] An identifier oligonucleotide can also comprise at least one 3' end that comprises a single nucleotide overhang.

[00242] An identifier oligonucleotide can also comprise a capture probe binding site. A capture probe binding site is a nucleic acid sequence to which a capture probe can bind.

[00243] A capture probe of the present disclosure can comprise a nucleic acid sequence complementary to a capture probe binding site. A capture probe can also comprise an affinity molecule.

[00244] An identifier oligonucleotide can also comprise a multiplexing probe binding site. A multiplexing probe binding site is a nucleic acid sequence to which a multiplexing probe can bind.

[00245] A multiplexing probe of the present disclosure can comprise a nucleic acid sequence complementary to a multiplexing probe binding site. A multiplexing probe can also comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released.

[00246] A probe of the present disclosure can include a region which permits the release of an identifier oligonucleotide following the application of a suitable force. In one non-limited example, the region is a cleavable motif (*e.g.*, a restriction enzyme site or cleavable linker). The cleavable motif allows release of an identifier oligonucleotide from a bound target nucleic acid or protein and the identifier oligonucleotide can then be collected and detected. The region which permits the release of an identifier oligonucleotide can be positioned between the target-binding domain and the identifier oligonucleotide, allowing for the release of the identifier oligonucleotide from the target binding domain. An identifier oligonucleotide is said to be releasable when it can be separated (*i.e.*, cleaved and released) from the remainder of the probe. Examples of cleavable motives include but are not limited to photo-cleavable linkers. Photo-cleavable linkers can be cleaved by light provided by a suitable coherent light source (*e.g.*, a laser and a UV light source) or a suitable incoherent light source (*e.g.*, an arc-lamp and a light-emitting diode (LED)).

[00247] In some aspects, the identifier oligonucleotide is collected from a solution proximal to, e.g., at least immediately above or surrounding, the point at which the identifier oligonucleotide is released or the at least one cell. The proximal solution may be collected by aspirating, e.g., via a pipette, a capillary tube, a microarray pin, a flow cell comprising holes, or another suitable aspirating system known in the art or any combination thereof. The capillary tube may comprise an optical device capable of transmitting a light force, e.g., UV light, to the at least one cell. The

pipette or a microarray pin may be attached to an array comprising a plurality of pipettes or microarray pins. The proximal solution may comprise an anionic polymer, e.g., dextran sulfate, and/or salmon sperm DNA and/or the collected signal oligonucleotide may be added to a solution comprising an anionic polymer, e.g., dextran sulfate, and/or salmon sperm DNA. Other non-specific blocking agents known in the art in addition to or instead of salmon sperm DNA may be used.

[00248] In some aspects, the identifier oligonucleotide is collected from a tissue, at least one cell or proximal to the point at which the identifier oligonucleotide is released via liquid laminar, turbulent, or transitional flow. The flow may be via a channel, e.g., having 25 to 500 m depth between the tissue and a fluidic device or impermeable barrier placed over the tissue.

[00249] In aspects where the target-binding domain of a probe is an antibody, the probe can be prepared using a cysteine bioconjugation method that is stable, site-specific to, preferably, the antibody's hinge-region heavy-chain. This preparation method provides relatively controllable identifier oligonucleotides to antibody stoichiometric ratios. A probe can comprise a plurality (i.e., more than one, e.g., 2, 3, 4, 5, or more) identifier oligonucleotides per antibody. Generally, "heavier" probes, which comprise 3 or 4 identifier oligonucleotides per antibody, are significantly less sensitive than antibodies lacking an identifier oligonucleotide or "lighter" probes, which comprise 1 or 2 identifier oligonucleotides per antibody.

[00250] In aspects, probes are provided to a sample at concentrations typically less than that used for immunohistochemistry (IHC) or for in situ hybridization (ISH). Alternately, the concentration may be significantly less than that used for IHC or ISH. For example, the probe concentration may be 2 fold less, 5 fold less, 10 fold less, 20 fold less, 25 fold less, 30 fold less, 50 fold less, 60 fold less, 70 fold less, 80 fold less, 90 fold less, 100 fold less, 200 fold less, 300 fold less, 400 fold less, 500 fold less, 600 fold less, 700 fold less, 800 fold less, 900 fold less, 1000 fold less, 2000 fold less, or less and any number in between. In aspects, probes are provided at a concentration of 100 nM, 70 nM, 60 nM, 50 nM, 40 nM, 30 nM, 20 nM, 10 nM, 9 nM, 8 nM, 7 nM, 6 nM, 5 nM, 4 nM, 3 nM, 2 nM, 1 nM, 0.9 nM, 0.8 nM, 0.7 nM, 0.6 nM, 0.5 nM, 0.4 nM, 0.3 nM, 0.2 nM, 0.1 nM, 0.09 nM, 0.08 nM, 0.07 nM, 0.06 nM, 0.05 nM, 0.04 nM, 0.03 nM, 0.02 nM, 0.01 nM, and less and any concentration in between.

[00251] Background noise, during protein detection, can be reduced by performing a negative purification of the intact probe molecule. This can be done by conducting an affinity purification

of the antibody or photo-cleavable linker after collection of eluate from a region of interest. Normally, released signal oligonucleotides will not be pulled out of solution. A protein-G or -O mechanism in a pipet tip, tube, or plate can be employed for this step. Such devices and reagents commercially available.

[00252] Background noise, during nucleic acid detection, can be reduced by performing a negative purification of the intact probe molecule. This can be done by conducting an affinity purification of the target binding domain or photo-cleavable linker after collection of eluate from a region of interest. Normally, released signal oligonucleotides will not be pulled out of solution. To assist in the negative purification, a universal purification sequence may be included in a probe, e.g., in the target binding domain.

[00253] Protein-targeting probes and nucleic acid-targeting probes may be applied simultaneously as long as conditions allow for binding of both a protein target and a nucleic acid target. Alternately, protein-targeting probes and nucleic acid-targeting probes may be applied sequentially when conditions allowing for binding of both a protein target and a nucleic acid target are not possible.

[00254] A set of probes is synonymous with a composition of probes. A set of probes includes at least one species of probes, *i.e.*, directed to one target. A set of probes preferably includes at least two, e.g., 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more species of probes. A probe set may include one or multiple copies of each species of probe.

[00255] A first set of probes only may be applied to a sample. Alternately, a second set (or higher number) of probes may be later applied to the sample. The first set and second (or higher number) may target only nucleic acids, only proteins, or a combination thereof.

[00256] In the present disclosure, two or more targets (*i.e.*, proteins, nucleic acids, or a combination thereof) are detected; 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more targets, and any number there between, are detected.

[00257] A set of probes may be pre-defined based upon the cell type or tissue type to be targeted. For example, if the tissue is a breast cancer, then the set of probes will include probes directed to proteins relevant to breast cancer cells (e.g., Her2, EGFR, and PR) and/or probes directed to proteins relevant to normal breast tissues. Additionally, the set of probes may be pre-defined

based upon developmental status of a cell or tissue to be targeted. Alternately, the set of probes may be pre-defined based upon subcellular localizations of interest, e.g., nucleus, cytoplasm, and membrane. For example, antibodies directed to Foxp3, Histone H3, or P-S6 label the nucleus, antibodies directed to CD3, CD4, PD-1, or CD45RO label the cytoplasm, and antibodies directed to PD-L1 label membranes.

[00258] A probe may be chemically synthesized or may be produced biologically using a vector into which a nucleic acid encoding the probe has been cloned.

[00259] Any probe or set of probes described herein may be used in methods and kits of the present disclosure.

[00260] For the herein-described probes, association of a unique nucleic acid sequence to a specific target nucleic acid or target protein is not fixed.

[00261] As described in the preceding, probes of the present disclosure can be used to detect a target nucleic acid or target protein present in any sample, *e.g.*, a biological sample. As will be appreciated by those in the art, the sample may comprise any number of things, including, but not limited to: cells (including both primary cells and cultured cell lines) and tissues (including cultured or explanted). In aspects, a tissue sample (fixed or unfixed) is embedded, serially sectioned, and immobilized onto a microscope slide. As is well known, a pair of serial sections will include at least one cell that is present in both serial sections. Structures and cell types, located on a first serial section will have a similar location on an adjacent serial section. The sample can be cultured cells or dissociated cells (fixed or unfixed) that have been immobilized onto a slide. A sample can be a formalin-fixed paraffin-embedded (FFPE) tissue sample.

[00262] In aspects, a tissue sample is a biopsied tumor or a portion thereof, i.e., a clinically-relevant tissue sample. For example, the tumor may be from a breast cancer. The sample may be an excised lymph node.

[00263] The sample can be obtained from virtually any organism including multicellular organisms, *e.g.*, of the plant, fungus, and animal kingdoms; preferably, the sample is obtained from an animal, *e.g.*, a mammal. Human samples are particularly preferred.

[00264] In some aspects, the probes, compositions, methods, and kits described herein are used in the diagnosis of a condition. As used herein the term diagnose or diagnosis of a condition includes predicting or diagnosing the condition, determining predisposition to the condition, monitoring treatment of the condition, diagnosing a therapeutic response of the disease, and

prognosis of the condition, condition progression, and response to particular treatment of the condition. For example, a tissue sample can be assayed according to any of the probes, methods, or kits described herein to determine the presence and/or quantity of markers of a disease or malignant cell type in the sample (relative to the non-diseased condition), thereby diagnosing or staging a disease or a cancer.

[00265] In general, samples attached to a slide can be first imaged using fluorescence (e.g., fluorescent antibodies or fluorescent stains (e.g., DAPI)) to identify morphology, regions of interest, cell types of interest, and single cells and then expression of proteins and/or nucleic acids can be digitally counted from the sample on the same slide.

[00266] Compositions and kits of the present disclosure can include probes and other reagents, for example, buffers and other reagents known in the art to facilitate binding of a protein and/or a nucleic acid in a sample, *i.e.*, for performing hybridization reactions.

[00267] A kit also will include instructions for using the components of the kit, including, but not limited to, information necessary to hybridize labeled oligonucleotides to a probe, to hybridize a probe to a target-specific oligonucleotide, to hybridize a target-specific oligonucleotide to a target nucleic acid and/or to hybridize a probe to target protein.

[00268] A region of interest may be a tissue type present in a sample, a cell type, a cell, or a subcellular structure within a cell.

[00269] Together, a comparison of the identity and abundance of the target proteins and/or target nucleic acids present in a first region of interest (*e.g.*, tissue type, a cell type (including normal and abnormal cells), and a subcellular structure within a cell) and the identity and abundance of the target proteins and/or target nucleic acids present in second region of interest or more regions of interest can be made using the methods of the present disclosure.

[00270] As described in the preceding, the products produced by the methods of the present disclosure can be used for nucleic acid amplification. In a preferred aspect, the nucleic acid amplification can be solid-phase nucleic acid amplification. Thus, in further aspects the invention provides a method of solid-phase nucleic acid amplification of template polynucleotide molecules which comprises: preparing a library of template polynucleotide molecules which have common sequences at their 5' and 3' ends using the methods of the present disclosure and carrying out a solid-phase nucleic acid amplification reaction wherein said template polynucleotide molecules are amplified. Compositions and methods for nucleic acid

amplification and sequencing have been described in, e.g., US9376678, which is incorporated herein by reference in its entirety.

[00271] The term "solid-phase amplification" as used herein refers to any nucleic acid amplification reaction carried out on or in association with a solid support such that all or a portion of the amplified products are immobilized on the solid support as they are formed. In particular, the term encompasses solid-phase polymerase chain reaction (solid-phase PCR), which is a reaction analogous to standard solution phase PCR, except that one or both of the forward and reverse amplification primers is/are immobilized on the solid support.

[00272] Although the invention encompasses "solid-phase" amplification methods in which only one amplification primer is immobilized (the other primer usually being present in free solution), it is preferred for the solid support to be provided with both the forward and the reverse primers immobilized. In practice, there will be a "plurality" of identical forward primers and/or a "plurality" of identical reverse primers immobilized on the solid support, since the PCR process requires an excess of primers to sustain amplification. References herein to forward and reverse primers are to be interpreted accordingly as encompassing a "plurality" of such primers unless the context indicates otherwise.

[00273] As will be appreciated by the skilled reader, any given PCR reaction requires at least one type of forward primer and at least one type of reverse primer specific for the template to be amplified. However, in certain aspects the forward and reverse primers may comprise template-specific portions of identical sequence, and may have entirely identical nucleotide sequence and structure (including any non-nucleotide modifications). In other words, it is possible to carry out solid-phase amplification using only one type of primer, and such single-primer methods are encompassed within the scope of the invention. Other aspects may use forward and reverse primers which contain identical template-specific sequences but which differ in some other structural features. For example one type of primer may contain a non-nucleotide modification which is not present in the other.

[00274] In other aspects of the invention the forward and reverse primers may contain template-specific portions of different sequence.

[00275] Amplification primers for solid-phase PCR are preferably immobilized by covalent attachment to the solid support at or near the 5' end of the primer, leaving the template-specific portion of the primer free for annealing to its cognate template and the 3' hydroxyl group free for

primer extension. Any suitable covalent attachment means known in the art may be used for this purpose. The chosen attachment chemistry will depend on the nature of the solid support, and any derivatization or functionalization applied to it. The primer itself may include a moiety, which may be a non-nucleotide chemical modification, to facilitate attachment. In one particularly preferred aspect the primer may include a sulphur-containing nucleophile, such as phosphorothioate or thiophosphate, at the 5' end. In the case of solid-supported polyacrylamide hydrogels (as described below), this nucleophile will bind to a "C" group present in the hydrogel. The most preferred means of attaching primers and templates to a solid support is via 5' phosphorothioate attachment to a hydrogel comprised of polymerised acrylamide and N-(5-bromoacetamidylpentyl)acrylamide (BRAPA).

[00276] The terms "cluster" and "colony" are used interchangeably herein to refer to a discrete site on a solid support comprised of a plurality of identical immobilized nucleic acid strands and a plurality of identical immobilized complementary nucleic acid strands. The term "clustered array" refers to an array formed from such clusters or colonies. In this context the term "array" is not to be understood as requiring an ordered arrangement of clusters.

[00277] The invention also encompasses methods of sequencing the amplified nucleic acids generated by solid-phase amplification. Thus, the invention provides a method of nucleic acid sequencing comprising amplifying a library of nucleic acid templates by the methods of the present disclosure described above, using solid-phase amplification as described above to amplify this library on a solid support, and carrying out a nucleic acid sequencing reaction to determine the sequence of the whole or a part of at least one amplified nucleic acid strand produced in the solid-phase amplification reaction.

[00278] Sequencing, as referred to herein, can be carried out using any suitable "sequencing-by-synthesis" technique, wherein nucleotides are added successively to a free 3' hydroxyl group, resulting in synthesis of a polynucleotide chain in the 5' to 3' direction. The nature of the nucleotide added is preferably determined after each nucleotide addition.

[00279] The initiation point for the sequencing reaction may be provided by annealing of a sequencing primer to a product of the whole genome or solid-phase amplification reaction. In this connection, one or both of the adapters added during formation of the template library may include a nucleotide sequence which permits annealing of a sequencing primer to amplified products derived by whole genome or solid-phase amplification of the template library.

[00280] The products of solid-phase amplification reactions wherein both forward and reverse amplification primers are covalently immobilized on the solid surface are so-called "bridged" structures formed by annealing of pairs of immobilized polynucleotide strands and immobilized complementary strands, both strands being attached to the solid support at the 5' end. Arrays comprised of such bridged structures provide inefficient templates for nucleic acid sequencing, since hybridization of a conventional sequencing primer to one of the immobilized strands is not favored compared to annealing of this strand to its immobilized complementary strand under standard conditions for hybridization.

[00281] In order to provide more suitable templates for nucleic acid sequencing it is preferred to remove substantially all or at least a portion of one of the immobilized strands in the "bridged" structure in order to generate a template which is at least partially single-stranded. The portion of the template which is single-stranded will thus be available for hybridization to a sequencing primer. The process of removing all or a portion of one immobilized strand in a "bridged" double-stranded nucleic acid structure may be referred to herein as "linearization".

[00282] Bridged template structures may be linearized by cleavage of one or both strands with a restriction endonuclease or by cleavage of one strand with a nicking endonuclease. Other methods of cleavage can be used as an alternative to restriction enzymes or nicking enzymes, including inter alfa chemical cleavage (e.g. cleavage of a diol linkage with periodate), cleavage of abasic sites by cleavage with endonuclease, or by exposure to heat or alkali, cleavage of ribonucleotides incorporated into amplification products otherwise comprised of deoxyribonucleotides, photochemical cleavage or cleavage of a peptide linker.

[00283] It will be appreciated that a linearization step may not be essential if the solid-phase amplification reaction is performed with only one primer covalently immobilized and the other in free solution.

[00284] In order to generate a linearized template suitable for sequencing it is necessary to remove "unequal" amounts of the complementary strands in the bridged structure formed by amplification so as to leave behind a linearized template for sequencing which is fully or partially single stranded. Most preferably one strand of the bridged structure is substantially or completely removed.

[00285] Following the cleavage step, regardless of the method used for cleavage, the product of the cleavage reaction may be subjected to denaturing conditions in order to remove the portion(s)

of the cleaved strand(s) that are not attached to the solid support. Suitable denaturing conditions will be apparent to the skilled reader with reference to standard molecular biology protocols.

[00286] Denaturation (and subsequent re-annealing of the cleaved strands) results in the production of a sequencing template which is partially or substantially single-stranded. A sequencing reaction may then be initiated by hybridization of a sequencing primer to the single-stranded portion of the template.

[00287] Thus, the nucleic acid sequencing reaction may comprise hybridizing a sequencing primer to a single-stranded region of a linearized amplification product, sequentially incorporating one or more nucleotides into a polynucleotide strand complementary to the region of amplified template strand to be sequenced, identifying the base present in one or more of the incorporated nucleotide(s) and thereby determining the sequence of a region of the template strand.

[00288] One preferred sequencing method which can be used in accordance with the invention relies on the use of modified nucleotides that can act as chain terminators. Once the modified nucleotide has been incorporated into the growing polynucleotide chain complementary to the region of the template being sequenced there is no free 3'-OH group available to direct further sequence extension and therefore the polymerase cannot add further nucleotides. Once the nature of the base incorporated into the growing chain has been determined, the 3' block may be removed to allow addition of the next successive nucleotide. By ordering the products derived using these modified nucleotides it is possible to deduce the DNA sequence of the DNA template. Such reactions can be done in a single experiment if each of the modified nucleotides has attached a different label, known to correspond to the particular base, to facilitate discrimination between the bases added at each incorporation step. Alternatively, a separate reaction may be carried out containing each of the modified nucleotides separately.

[00289] The modified nucleotides may carry a label to facilitate their detection. Preferably this is a fluorescent label. Each nucleotide type may carry a different fluorescent label. However the detectable label need not be a fluorescent label. Any label can be used which allows the detection of an incorporated nucleotide.

[00290] One method for detecting fluorescently labelled nucleotides comprises using laser light of a wavelength specific for the labelled nucleotides, or the use of other suitable sources of

illumination. The fluorescence from the label on the nucleotide may be detected by a CCD camera or other suitable detection means.

[00291] The invention is not intended to be limited to use of the sequencing method outlined above, as essentially any sequencing methodology which relies on successive incorporation of nucleotides into a polynucleotide chain can be used. Suitable alternative techniques include, for example, Pyrosequencing, FISSEQ (fluorescent in situ sequencing), MPSS (massively parallel signature sequencing) and sequencing by ligation-based methods.

[00292] In methods of the present disclosure, the unique nucleic acid sequence present in the identifier oligonucleotide of a probe which identifies the target analyte bound to the target binding domain of the probe can comprise between about 5 nucleotides and about 50 nucleotides. Preferably, the sequence comprises between about 20 nucleotides and about 40 nucleotides. Even more preferably, the sequence comprises about 35 nucleotides. In some preferred aspects, the sequence comprises 10 nucleotides.

[00293] In methods of the present disclosure, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released comprises between about 6 nucleotides and about 15 nucleotides. Preferably, the sequence comprises about 12 nucleotides.

[00294] In methods of the present disclosure, an amplification primer binding sites comprises between about 18 nucleotides and about 40 nucleotides. Preferably, an amplification primer binding sites comprises about 32 nucleotides.

[00295] In some aspects of the methods of the present disclosure, an amplification primer binding site can comprise an i7 sequence, wherein the i7 sequence comprises the sequence set forth in SEQ ID NO: 1.

[00296] In some aspects of the methods of the present disclosure, an amplification primer binding site can comprise an i5 sequence, wherein the i5 sequence comprises the sequence set forth in SEQ ID NO: 2.

[00297] In some aspects of the methods of the present disclosure, an amplification primer can comprise a flow cell adapter sequence, wherein the flow cell adapter sequence is suitable for sequencing. Preferably, at least one amplification primer used in the methods of the present disclosure comprises a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3. Preferably still, at least one amplification

primer used in the methods of the present disclosure comprises a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

[00298] In methods of the present disclosure, a unique molecular identifier can comprise between about 6 nucleotides and about 30 nucleotides. Preferably, a unique molecular identifier can comprise about 15 nucleotides. The terms unique molecular identifier and random molecular tags are used interchangeably herein. Using methods known in that art, unique molecular identifiers are random sequences that can be used to correct for biases in amplification prior to sequencing.

[00299] In methods of the present disclosure, a constant nucleic acid sequence to minimize ligation bias comprises between about 1 nucleotide and about 15 nucleotides. Preferably, the constant sequence comprises about 8 nucleotides.

[00300] In some aspects, a flow cell binding site can comprise between about 15 to about 40 nucleotides. A flow cell binding site can comprise about 29 nucleotides. A flow cell binding site can comprise about 24 nucleotides.

[00301] In some aspects, a target binding domain can comprise between about 10 to about 70 nucleotides. A target binding domain can comprise between about 30 to about 55 nucleotides. A target binding domain can comprise between about 35 to about 50 nucleotides.

[00302] In some aspects, a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain can comprise between about 20 to about 40. A unique nucleic acid sequence which identifies the target analyte bound to the target binding domain can comprise about 25 nucleotides, or about 35 nucleotides, or about 12 nucleotides.

[00303] In some aspects, an amplification primer binding site can comprise between about 20 to about 50 nucleotides. An amplification primer binding site can comprise about 33 nucleotides, or about 34 nucleotides.

[00304] In some aspects, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can comprise between about 1 to about 20 nucleotides. A nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can comprise about 8 nucleotides.

[00305] In some aspects, a nucleic acid sequence comprising a unique molecular identifier can comprise between about 5 to about 20 nucleotides. A nucleic acid sequence comprising a unique molecular identifier can comprise about 14 nucleotides.

[00306] As used herein, the terms "region of interest" and "ROI" are used in their broadest sense to refer to a specific location within a sample that is to be analyzed using the methods of the present disclosure.

[00307] As used herein, the term "adjacent" can mean within about 1 nucleotide, or within about 2 nucleotides, or within about 3 nucleotides, or within about 4 nucleotides, or within about 5 nucleotides, or within about 6 nucleotides, or within about 7 nucleotides, or within about 8 nucleotides, or within about 9 nucleotides, or within about 10 nucleotides, or within about 11 nucleotides, or within about 12 nucleotides, or within about 13 nucleotides, or within about 14 nucleotides, or within about 15 nucleotides, or within about 16 nucleotides, or within about 17 nucleotides, or within about 18 nucleotides, or within about 19 nucleotides, or within about 20 nucleotides, or within about 21 nucleotides, or within about 22 nucleotides, or within about 23 nucleotides, or within about 24 nucleotides, or within about 25 nucleotides, or within about 26 nucleotides, or within about 27 nucleotides, or within about 28 nucleotides, or within about 29 nucleotides, or within about 30 nucleotides, or within about 40 nucleotides, or within about nucleotides, or within about 50 nucleotides, or within about 60 nucleotides, or within about 70 nucleotides, or within about 80 nucleotides, or within about 90 nucleotides, or within about 100 nucleotides.

[00308] As used herein, the term "spatially detecting" is used in its broadest sense to refer to the identification of the presence of a specific target analyte within a specific region of interest in a sample. Spatially detecting can comprise quantifying the amount of a specific target analyte present within a specific region of interest in a sample. Spatially detecting can further comprise quantifying the relative amount of a first target analyte within a specific region of interest in a sample as compared to the amount of at least a second target analyte within a specific region of interest in a sample. Spatially detecting can also comprise quantifying the relative amount of a specific target analyte within a first region of interest in a sample compared to the amount of the same target analyte in at least a second region of interest in the same sample or different sample.

[00309] In some aspects of the methods and compositions of the present disclosure, a target analyte can be any molecule within a sample that is to be spatially detected. Target analytes include, but are not limited to, nucleic acid molecules and protein molecules. When the target analyte is a protein, the protein can be referred to as a target protein. When the target analyte is a nucleic acid, the nucleic acid can be referred to as a target nucleic acid. Target nucleic acids can include, but

are not limited to, mRNA molecules, micro RNA (miRNA) molecules, tRNA molecules, rRNA molecules, gDNA or any other nucleic acid present within a sample.

[00310] In some aspects of the methods and compositions of the present disclosure, the term target binding domain is used in its broadest sense to refer to a portion of a probe of the present disclosure that binds to, either directly or indirectly, a target analyte located in a sample. A target binding domain can comprise nucleic acid, protein, at least one antibody, an aptamer, or any combination thereof. A target binding domain can comprise DNA, RNA or any combination thereof. A target binding domain can comprise any number of modified nucleotides and/or nucleic acid analogues.

[00311] In the aspect that the target analyte to be spatially detected is a target protein, a target binding domain can be a protein-target binding domain. A protein-target binding domain can comprise an antibody or antibody fragment that binds to the target protein.

[00312] In the aspect that the target analyte to be spatially detected is a target nucleic acid, a target binding domain can be a target nucleic acid-binding region. A target nucleic acid-binding region can comprise a nucleic acid that is complementary to the target nucleic acid to be spatially detected. A target nucleic acid-binding region can comprise a nucleic acid that hybridizes to the target nucleic acid to be detected.

[00313] As used herein, the term "hybridize" is used in its broadest sense to mean the formation of a stable nucleic acid duplex. In one aspect, "stable duplex" means that a duplex structure is not destroyed by a stringent wash under conditions such as, for example, a temperature of either about 5 °C below or about 5 °C above the Tm of a strand of the duplex and low monovalent salt concentration, *e.g.,* less than 0.2 M, or less than 0.1 M or salt concentrations known to those of skill in the art. A duplex can be "perfectly matched", such that the polynucleotide and/or oligonucleotide strands making up the duplex form a double stranded structure with one another such that every nucleotide in each strand undergoes Watson-Crick base pairing with a nucleotide in the other strand. The term "duplex" comprises, but is not limited to, the pairing of nucleoside analogs, such as deoxyinosine, nucleosides with 2-aminopurine bases, PNAs, and the like, that can be employed. A duplex can comprise at least one mismatch, wherein the term "mismatch" means that a pair of nucleotides in the duplex fail to undergo Watson-Crick bonding.

[00314] As used herein, the term "hybridization conditions," will typically include salt concentrations of less than about 1 M, more usually less than about 500 mM and even more usually less than about 200 mM. Hybridization temperatures can be as low as 5 °C, but are typically greater

than 22 °C, more typically greater than about 30 °C, and often in excess of about 37 °C. Hybridizations are usually performed under stringent conditions, *e.g.,* conditions under which a probe will specifically hybridize to its target analyte. Stringent conditions are sequence-dependent and are different in different circumstances. Longer fragments can require higher hybridization temperatures for specific hybridization. As other factors can affect the stringency of hybridization, including base composition and length of the complementary strands, presence of organic solvents and extent of base mismatching, the combination of parameters is more important than the absolute measure of any one alone. Certain hybridization conditions will promote the formation of a duplex between the entire length of a target binding domain and the target analyte. Other hybridization conditions will promote the formation of a duplex only along certain portions of the target binding domain.

[00315] In some aspects of the methods and compositions of the present disclosure, a probe can comprise a target binding domain directly, or indirectly linked to an identifier oligonucleotide. In the context of a probe, an identifier oligonucleotide is a polynucleotide that comprises a nucleic acid sequence that identifies the target analyte bound to the target binding domain of that probe. That is to say, the identifier oligonucleotide comprises a specific nucleic acid sequence that is *a priori* assigned to the specific target analyte bound to the target binding to which the identifier oligonucleotide is attached. In a non-limiting example, a probe designated as "probe *X*" designed to spatially detect "target analyte *X*" comprises a target binding domain designated "target binding domain *X*" linked to an identifier oligonucleotide designated "identifier oligonucleotide *X*". Target binding domain *X* binds to target analyte *X* and identifier oligonucleotide *X* comprises a nucleic acid sequence, designated as "nucleic acid sequence *X*", which corresponds to target analyte *X*. Thus, if a skilled artisan practicing the methods of the present disclosure were to collect identifier oligonucleotides released from a region of interest in sample and obtain nucleic acid sequence *X* after sequencing, the skilled artisan would understand that to indicate that target analyte *X* was present in that region of interest. The amount, or number of sequencing reads, of nucleic acid sequence *X* can be used to determine the quantify, in absolute or relative terms, the amount of target analyte *X* within the region of interest.

[00316] As used herein, the term "amplification primer binding site" is used in its broadest sense to refer to a nucleic acid sequence that is complementary to, or at least partially complementary to at least one amplification primer, wherein the amplification primer is a short single-stranded or

partially single-stranded oligonucleotide that is sufficient to prime DNA and/or RNA synthesis, for example, by PCR.

[00317] In some aspects of the methods and compositions of the present disclosure, a target binding domain can be linked to an identifier oligonucleotide by a cleavable linker. Suitable cleavable linkers include, but are not limited to, chemically cleavable linkers (*e.g.* a linker that is cleaved when exposed to a particular chemical, combination of chemicals or reaction conditions), a photo-cleavable linker (*e.g.* a linker that is cleaved when exposed to light of a sufficient wavelength or light comprising a sufficient range of wavelengths), or an enzymatically cleavable linker (*e.g.* a linker that is cleaved by a specific enzyme or class of enzymes). Thus, as used herein the phrase "providing a force to a location of the sample sufficient to release an identifier oligonucleotide" is used in its broadest sense to describe changing the conditions within a certain region of interest in a sample such that, for any probe bound to a target analyte within that region of interest, the linker between the target binding domain of the probe and the identifier oligonucleotide of the probe is cleaved, thereby separating the identifier oligonucleotide from the target binding domain so that the identifier oligonucleotide can be subsequently collected from solution. For example, in aspects wherein a probe comprises a chemically cleavable linker between the target binding domain and the identifier oligonucleotide, providing a force to a location of the sample sufficient to release an identifier oligonucleotide can comprise exposing that location of the sample to the specific chemical, combination of chemicals or reaction conditions that catalyze the cleavage of the linker. In another non-limiting example, in aspects wherein a probe comprises a photo-cleavable linker between the target binding domain and the identifier oligonucleotide, providing a force to a location of the sample sufficient to release an identifier oligonucleotide can comprise exposing/exciting that location of the sample with light of a sufficient wavelength capable of cleaving the photo-cleavable linker. In another non-limiting example, in aspects wherein a probe comprises an enzymatically cleavable linker between the target binding domain and the identifier oligonucleotide, providing a force to a location of the sample sufficient to release an identifier oligonucleotide can comprise exposing that location of the sample to an amount of enzyme sufficient to catalyze the cleavage of the linker.

[00318] Providing a force to a location of the sample sufficient to release an identifier oligonucleotide can result in at least about 10%, or at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least

about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 99% of probes bound to target analytes within that location of the sample to undergo cleavage of the linker connecting the target binding domain and the identifier oligonucleotide.

[00319] As would be appreciated by one skilled in the art, the term "unique molecular identifier" or "UMI" refer to short nucleic acid sequences that are used to quantify and reduce quantitative bias caused by nucleic acid amplification prior to sequencing reactions.

[00320] In some aspects of the methods and compositions of the present disclosure, an affinity moiety can comprise biotin, avidin, streptavidin, nucleic acid, or any combination thereof.

[00321] In some aspects of the methods and compositions of the present disclosure, a probe can comprise at least at least about 5, about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

[00322] In some aspects of the methods and compositions of the present disclosure, a target binding domain can comprise at least at least about 5, about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

[00323] In some aspects of the methods and compositions of the present disclosure, an identifier oligonucleotide can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

[00324] In some aspects of the methods and compositions of the present disclosure, an amplification primer can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

[00325] In some aspects of the methods and compositions of the present disclosure, a nucleic acid probe can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

[00326] In some aspects of the methods and compositions of the present disclosure, a nucleic acid complementary to a portion of a identifier oligonucleotide can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95 or at least about 100 nucleotides.

[00327] In some aspects of the methods and compositions of the present disclosure, a nucleic acid sequence comprising a molecular identifier can comprise at least about 5, or at least about 10 nucleotides, or at least about 15, or at least about 20, or at least about 25, or at least about 30, or at least about 35, or at least about 40, or at least about 45, or at least about 50 nucleotides.

[00328] In some aspects of the methods and compositions of the present disclosure, an amplification primer binding site can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65 or at least about 70 nucleotides.

[00329] In some aspects of the methods and compositions of the present disclosure, a flow cell adapter sequence suitable for sequencing can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95 or at least about 100 nucleotides.

[00330] In some aspects of the methods and compositions of the present disclosure, a flow cell binding site can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 90, at least about 95 or at least about 100 nucleotides.

[00331] In some aspects of the methods and compositions of the present disclosure, a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released can comprise at least about 5, or at least about 10 nucleotides, or at least about 15, or at least about 20, or at least about 25, or at least about 30, or at least about 35, or at least about 40, or at least about 45, or at least about 50 nucleotides.

[00332] In some aspects of the methods and compositions of the present disclosure, a unique nucleic acid sequence which identifies the target analyte bound to a target binding domain can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 90, at least about 95 or at least about 100 nucleotides.

[00333] In some aspects of the methods and compositions of the present disclosure, a probe, a target binding domain, an identifier oligonucleotide, an amplification primer, a nucleic acid probe, a nucleic acid complementary to a portion of a identifier oligonucleotide, a nucleic acid sequence comprising a molecular identifier, an amplification primer binding site, a flow cell adapter sequence, a flow cell binding site, a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released, a unique nucleic acid sequence which identifies the target analyte bound to a target binding domain or any combination thereof can comprise at least one natural base, can comprise no natural bases, can comprise at least one modified nucleotide or nucleic acid analog, can comprise no modified

nucleotides or nucleic acid analogs, can comprise at least one universal base, can comprise no universal bases, can comprise at least one degenerate base or can comprise no degenerate bases.

[00334] In some aspects of the methods and compositions of the present disclosure, a probe, a target binding domain, an identifier oligonucleotide, an amplification primer, a nucleic acid probe, a nucleic acid complementary to a portion of a identifier oligonucleotide, a nucleic acid sequence comprising a molecular identifier, an amplification primer binding site, a flow cell adapter sequence, a flow cell binding site, a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released, a unique nucleic acid sequence which identifies the target analyte bound to a target binding domain or any combination thereof can comprise any combination natural bases (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more natural bases), modified nucleotides or nucleic acid analogs (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more modified or analog nucleotides), universal bases (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more universal bases), or degenerate bases (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more degenerative bases). When present in a combination, the natural bases, modified nucleotides or nucleic acid analogs, universal bases and degenerate bases can be arranged in any order.

[00335] The terms "modified nucleotides" or "nucleic acid analogues" include, but are not limited to, locked nucleic acids (LNA), bridged nucleic acids (BNA), propyne-modified nucleic acids, zip nucleic acids (ZNA®), isoguanine and isocytosine. Preferably, the modified nucleotides or nucleic acid analogues are locked nucleic acids (LNAs).

[00336] The term "locked nucleic acids (LNA)" as used herein includes, but is not limited to, a modified RNA nucleotide in which the ribose moiety comprises a methylene bridge connecting the 2' oxygen and the 4' carbon. This methylene bridge locks the ribose in the 3'-endo confirmation, also known as the north confirmation, that is found in A-form RNA duplexes. The term inaccessible RNA can be used interchangeably with LNA. The term "bridged nucleic acids (BNA)" as used herein includes, but is not limited to, modified RNA molecules that comprise a five-membered or six-membered bridged structure with a fixed 3'-endo confirmation, also known as the north confirmation. The bridged structure connects the 2' oxygen of the ribose to the 4' carbon of the ribose. Various different bridge structures are possible containing carbon, nitrogen, and hydrogen atoms. The term "propyne-modified nucleic acids" as used herein includes, but is not limited to, pyrimidines, namely cytosine and thymine/uracil, that comprise a propyne modification at the C5 position of the nucleic acid base. The term "zip nucleic acids

(ZNA®)" as used herein includes, but is not limited to, oligonucleotides that are conjugated with cationic spermine moieties.

[00337] The term "universal base" as used herein includes, but is not limited to, a nucleotide base does not follow Watson-Crick base pair rules but rather can bind to any of the four canonical bases (A, T/U, C, G) located on the target nucleic acid. The term "degenerate base" as used herein includes, but is not limited to, a nucleotide base that does not follow Watson-Crick base pair rules but rather can bind to at least two of the four canonical bases A, T/U, C, G), but not all four. A degenerate base can also be termed a Wobble base; these terms are used interchangeably herein.

[00338] As used in this Specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

[00339] Unless specifically stated or obvious from context, as used herein, the term "or" is understood to be inclusive and covers both "or" and "and".

[00340] Unless specifically stated or obvious from context, as used herein, the term "about" is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term "about."

[00341] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although other probes, compositions, methods, and kits similar, or equivalent, to those described herein can be used in the practice of the present disclosure, the preferred materials and methods are described herein. It is to be understood that the terminology used herein is for the purpose of describing particular aspects only, and is not intended to be limiting.

EXAMPLES

[00342] Example 1—two-ended adapter ligation method for 96 multiplexed samples

[00343] In this example, a two-ended adapter ligation method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. The nucleic acid adapters used in this experiment were partially double-stranded. The nucleic acid adapters comprised a first strand and a second strand. The first strand comprised a 5' phosphate moiety for ligation. The first strand also comprised a constant nucleic acid sequence to minimize ligation

bias (GCGTAGTG), a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a first amplification primer binding site (SEQ ID NO: 2). The second strand comprised a single overhanging thymine nucleotide at the 3' end, a sequence complementary to the constant nucleic acid sequence to minimize ligation bias present in the first strand, a sequence complementary to the unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released present in the first strand and a second amplification primer binding site (SEQ ID NO: 1).

[00344] To form the partially double stranded nucleic acid adapters, first strand oligonucleotides and second strand oligonucleotides were combined in equimolar proportion for a final total oligonucleotide concentration of 28 μM in buffer comprising 50 mM NaCl. The oligonucleotide mixture was heated at 95°C for 2 minutes and cooled at ambient temperature for 30 minutes, thereby annealing the first stand and second strand oligonucleotides together to form the partially double-stranded nucleic acid adapters. Annealed nucleic acid adaptors were diluted to final concentration ranging between 0.02 μM to 0.002 μM in a solution of 10 mM Tris pH 8 and 0.05% Tween20.

[00345] Collected identifier oligonucleotides were end repaired and A-tailed using NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) with a modified protocol. End repair/A-tail master mix was prepared by combining the following: 627.8 μL of PCR-grade H2O, 143.9 μL of NEBNext Ultra II End Prep Reaction Buffer, and 61.7 μL of NEBNext Ultra II End Prep Enzyme Mix. 8.3 μL of end repair/A-tail master mix was added to 4 μL of each sample of identifier oligonucleotides. The reaction was incubated for 30 minutes at 20°C with a heated lid of >75°C, followed by a second incubation for 30 minutes at 65°C. The repaired/A-tailed identifier oligonucleotide mixtures were then stored at 4°C.

[00346] Following end repair and A-tailing, the nucleic acid adaptors were ligated to the repaired/A-tailed identifier oligonucleotides by adding 6.4 μL of NEBNext Ultra II Ligation Master Mix, 0.2 μL of NEBNext Ligation Enhancer, and 1 μL of the nucleic acid adapter dilution to each repaired/A-tailed identifier oligonucleotide mixture. These reactions were incubated for 15 minutes at 20°C with the heated lid off and subsequently quenched with 1 μL

0.5M EDTA. All of the reactions were then pooled into a single 15 mL conical tube to form a pooled adapter-ligated sample.

[00347] The pooled adaptor-ligated sample was purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350 µL of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). AMPure XP bead cleanup was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. AMPure XP bead cleanup was then repeated with 400 µL of AMPure XP beads and eluted in 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20 to obtain purified adapter-ligated samples.

[00348] Following AMPure XP cleanup, PCR reactions with purified adaptor-ligated sample were prepared to amplify the adapter-ligated identifier oligonucleotides. To 6 µL of the purified adapter-ligated sample, 10 µL of NEBNext Ultra II Q5 Master Mix, 0.2 µL of 100 µM forward and reverse primers, and 3.6 µL of PCR-grade H2O was added. The forward primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the first strand of the nucleic acid adapter. Table 1 provides the sequences of the forward primers used.

Table 1. Forward primers for two-ended adapter ligation.

Primer Sequence	SEQ ID NO
AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	5
AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	6
AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	7
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	8
AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	9
AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	10
AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	11
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	12

[00349] The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second

amplification primer binding site located on the second strand of the nucleic acid adapter. Table 2 provides the sequences of the reverse primers used.

Table 2. Reverse primers for two-ended adapter ligation

Primer Sequence	SEQ ID NO
CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	13
CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	14
CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	15
CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	16
CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	17
CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	18
CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	19
CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	20
CAAGCAGAAGACGGCATACGAGATCATAGCCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	21
CAAGCAGAAGACGGCATACGAGATTTCGCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	22
CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	23
CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	24

[00350] The optimal number of PCR cycles was determined empirically with triplicate PCR reactions. Alternately, the optimal number of PCR cycles could have been determined using real-time/qPCR. The PCR program used comprised the following steps:

(1) 30 seconds at 98°C

(2) 10 seconds at 98°C

(3) 1 minute at 65°C

(4) Repeating steps (2) and (3) nine times

(5) 5 minutes at 65°C

[00351] The amplified products were purified using 18 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00352] The amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with a custom spike-in primer comprising the nucleotide sequence ACACTCTTTAAGACGACGTCGCTATGGCCTCTCC (SEQ ID NO: 25).

[00353] Example 2—one-ended adapter ligation method for 96 multiplexed samples

[00354] In this example, a one-ended adapter ligation method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. The nucleic acid adapters used in this experiment were partially double-stranded. The nucleic acid adapters comprised a first strand and a second strand. The first strand comprised a 5' phosphate moiety for ligation. The first strand also comprised a constant nucleic acid sequence to minimize ligation bias (CACTACGC), a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a first amplification primer binding site (SEQ ID NO: 1). The second strand comprised a single overhanging thymine nucleotide at the 3' end, a sequence complementary to the constant nucleic acid sequence to minimize ligation bias present in the first strand and a sequence complementary to the unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released present in the first strand.

[00355] To form the partially double stranded nucleic acid adapters, first strand oligonucleotides and second strand oligonucleotides were combined in equimolar proportion for a final total oligonucleotide concentration of 28µM in 50 mM NaCl. The oligonucleotide mixture was heated to 95°C for 2 minutes and cooled at ambient temperature for 30 minutes, thereby annealing the first stand and second strand oligonucleotides together to form the partially double-stranded nucleic acid adapters. Annealed nucleic acid adaptors were diluted to a final concentration ranging between 0.02 µM to 0.002 µM in a solution of 10 mM Tris pH 8 and 0.05% Tween20.

[00356] Nucleic acid adapters were ligated to the collected identifier oligonucleotides by addition of 10 µL of 2X rapid ligation buffer (Enzymatics), 1 µL of T4 DNA Rapid Ligase (Enzymatics),

and 1 µL of annealed nucleic acid adapter dilutions to each sample of collected identifier oligonucleotides. Samples were incubated for 15 minutes at 20°C with the heated lid off and subsequently quenched with 1 µL 0.5M EDTA. All of the reactions were then pooled into a single 15 mL conical tube to form a pooled adapter-ligated sample.

[00357] The pooled adaptor-ligated sample was purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350 µL of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). AMPure XP bead cleanup was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. AMPure XP bead cleanup was then repeated with 400 µL of AMPure XP beads and eluted in 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20 to obtain purified adapter-ligated samples.

[00358] Following AMPure XP cleanup, PCR reactions with purified adaptor-ligated sample were prepared to amplify the adapter-ligated identifier oligonucleotides. To 6 µL of the purified adapter-ligated sample, 10 µL of NEBNext Ultra II Q5 Master Mix, 0.2 µL of 100 µM forward and reverse primers, and 3.6 µL of PCR-grade H_2O was added. The forward primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the first strand of the nucleic acid adapter. Table 3 provides the sequences of the forward primers used.

Table 3. Forward primers for two-ended adapter ligation.

Primer Sequence	SEQ ID NO
AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	5
AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	6
AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	7
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	8
AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	9
AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	10
AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	11
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	12

[00359] The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second amplification primer binding site located on the second strand of the nucleic acid adapter. Table 4 provides the sequences of the reverse primers used.

Table 4. Reverse primers for two-ended adapter ligation.

Primer Sequence	SEQ ID NO
CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	13
CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	14
CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	15
CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	16
CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	17
CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	18
CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	19
CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	20
CAAGCAGAAGACGGCATACGAGATCATAGCCGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	21
CAAGCAGAAGACGGCATACGAGATTTCGCGGAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	22
CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	23
CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	24

[00360] The optimal number of PCR cycles was determined empirically with triplicate PCR reactions. Alternately, the optimal number of PCR cycles could have been determined using real-time/qPCR. The PCR program used comprised the following steps:

(1) 30 seconds at 98°C

(2) 10 seconds at 98°C

(3) 1 minute at 65°C

(4) Repeating steps (2) and (3) nine times

(5) 5 minutes at 65°C

[00361] The amplified products were purified using 18 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00362] The amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with a custom spike-in primer comprising the nucleotide sequence ACACTCTTTAAGACGACGTCGCTATGGCCTCTCC (SEQ ID NO: 25).

[00363] Example 3—templated-primer extension method for 96 multiplexed samples

[00364] In this example, a templated-primer extension method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. The single stranded nucleic acid templates used in this example comprised a 3' biotin moiety, a region complementary to the unique nucleic acid sequences present in the collected identifier oligonucleotides, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding sequence (GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT, SEQ ID NO: 26). Table 5 provides the sequences of the single-stranded nucleic acid templates used in this example.

Table 5. Single stranded nucleic acid templates for templated-primer extension method

Single-stranded nucleic acid templates	SEQ ID NO
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTTGAAGCACACCGTTTTTCTTTCTTCTTTCACGG	27
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNACCCACAGGTTATACGGGATTATCCGGTTATCCA	28
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCGACACCGAGTTCGACCGTTATGTTGGTAGGATC	29
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCGGTGTGTAAGCGTAACGATGTTGGTGTCGCTCT	30
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCAGACACTGCGACAACTCACGATCATGACACAGA	31
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNATATTCTGTACTCAGTGCCTATCCACCTAATAGG	32
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTTCAGTTATAATGTGTCCAGCAGAAGCAGGAATT	33
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNGTCCTTTGTTGGGCGGACCGTAATGAGGAATTTG	34
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNGATGAGACTTCTACATGTCCGATGTTTTTGTGCT	35

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNACTC ACACATAGTACTGACACGTAAGATAGGATG	36
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNTTAC CCTATCTCGTCTATGTACGTCAGGCTGAAT	37
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNATCA ACGTAGGGTAAGGTCATATTTTTACCTTAC	38
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNTTCC CTCTTTCTCCGCTTATGGATGAAAGGACAG	39
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCCTG CACAGTGAGTTTCTTTCACTCTAACTCTCT	40
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNTGTC GCTCTAGTGTGACTTTTCCACCTCGCATCT	41
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNATAT CTTTCTCGGGTAAAGATTAGGCGTCCGATA	42
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCGAT TAGCCGTAGACGCAACTCATTGCCGAAGAT	43
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNTGTG AGCATTTCAGTACGAGTGATGCAGATAAAC	44
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNTATA GTTACCAAGTACTATGGGTTGGTGGAAGCC	45
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCCAA TTATACTGTCTGTTATGTTCTCGGATAAGC	46
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNTCAG GTGCTTGTAGGCTCATGATAGGGGTAATGC	47
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCTCT GCTGTAATCTCAGCTCCACTTGTTTCTAAG	48
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNGTGC ATATTGCAGCTGAGCCAGCTCAATTTGAAG	49
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCCGT TGATTTACGCAACAGCGGCTTATATAGCTC	50
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCATC ATCGACAGTTCGCAGCCCTATAACATGATA	51
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNATCG CAGGATGGTACAGCATCATACATGATGAGC	52
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCTGA TAAGTCGTAGGAATGTCGCTTAATACGGAT	53
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNATGG CGGTTTCGGGTCCTGCACTATTCCTAATAA	54
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNCCAG TACGGGTACTAATAAGTGTCATATCTATTG	55
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNTGTT GGAGAGGTTAGAGGTGAGGAGGCGAAGATA	56

[00365] Single stranded nucleic acid templates were ordered from Integrated DNA Technologies, Inc. and quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific). Individual single stranded nucleic acid templates were normalized to a standard concentration

and then pooled to be equimolar. The pool of single stranded nucleic acid templates was diluted to 0.83 nM in a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00366] The collected identifier oligonucleotides were hybridized to the single stranded nucleic acid templates and extended by addition of 10 µL of NEBNext Ultra II Q5 Master Mix (New England Biolabs), 4 µL of the diluted single stranded nucleic acid template pool and 4 µL of H_2O to 2 µL of each sample of identifier oligonucleotides. The following PCR program was used to extend the identifier oligonucleotides:

(1) 30 seconds at 98°C, 10×

(2) 1 minutes at 98°C,

(3) 1 minutes at 68°C

(4) 1 minutes at 72°C

(5) Repeating steps (2)-(4) ten times

(6) 2 minutes at 72°C

[00367] The extension products were stored at 4°C. Magnetic streptavidin beads (MyOne Streptavidin C1 beads, Thermo Fisher Scientific) were washed in 1X Binding and Washing Buffer (5 mM Tris-HCl, 0.5 mM EDTA, 1M NaCl), and 5 µL of streptavidin beads were added to each extension product sample. The extension product samples were incubated with the beads on an orbital mixer for a minimum of 15 minutes. Following incubation, the samples were heated to 95°C for 3 minutes and transferred to a magnetic plate. Supernatant was extracted immediately after sufficient bead pelleting to yield the purified extension product samples.

[00368] The purified extension product samples were amplified by adding to 7.5 µL of each purified extension product sample, 12.5 µL of NEBNext Ultra II Q5 Master Mix, 0.25 µL of 100 µM forward primer, 1 µL of 25 µM reverse primer and 3.8 µL of PCR-grade H_2O. The forward primer comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the identifier oligonucleotide. Table 6 provides the sequences of the forward primers used in this example.

Table 6. Forward primers for templated-primer extension method.

Primer Sequence	SEQ ID NO
AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTTA AGACGACGTCGCTATGGCCTCTCC	5
AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	6
AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTTA AGACGACGTCGCTATGGCCTCTCC	7
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTTA AGACGACGTCGCTATGGCCTCTCC	8
AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	9
AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTTA AGACGACGTCGCTATGGCCTCTCC	10
AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTTTA AGACGACGTCGCTATGGCCTCTCC	11
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTTA AGACGACGTCGCTATGGCCTCTCC	12

[00369] The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second amplification primer binding site located single-stranded nucleic acid template. Table 7 provides the sequences of the reverse primers used.

Table 7. Reverse primers for templated-primer extension method.

Primer Sequence	SEQ ID NO
CAAGCAGAAGACGGCATACGAGATGTCGGTAAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	57
CAAGCAGAAGACGGCATACGAGATAGGTCACTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	58
CAAGCAGAAGACGGCATACGAGATGAATCCGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	59
CAAGCAGAAGACGGCATACGAGATGTACCTTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	60
CAAGCAGAAGACGGCATACGAGATCATGAGGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	61
CAAGCAGAAGACGGCATACGAGATTGACTGACGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	62
CAAGCAGAAGACGGCATACGAGATCGTATTCGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	63
CAAGCAGAAGACGGCATACGAGATCTCCTAGAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	64
CAAGCAGAAGACGGCATACGAGATTAGTTGCGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	65

CAAGCAGAAGACGGCATACGAGATGAGATACGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	66
CAAGCAGAAGACGGCATACGAGATAGGTGTACGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	67
CAAGCAGAAGACGGCATACGAGATTAATGCCGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	68
CAAGCAGAAGACGGCATACGAGATTCAGACGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	69
CAAGCAGAAGACGGCATACGAGATGATAGGCTGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	70
CAAGCAGAAGACGGCATACGAGATTGGTACAGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	71
CAAGCAGAAGACGGCATACGAGATCAAGGTCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	72
CAAGCAGAAGACGGCATACGAGATGCTATCCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	73
CAAGCAGAAGACGGCATACGAGATATGGAAGGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	74
CAAGCAGAAGACGGCATACGAGATTCAAGGACGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	75
CAAGCAGAAGACGGCATACGAGATGTTACGCAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	76
CAAGCAGAAGACGGCATACGAGATAGTCTGTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	77
CAAGCAGAAGACGGCATACGAGATGCACGTAAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	78
CAAGCAGAAGACGGCATACGAGATAACCTTGGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	79
CAAGCAGAAGACGGCATACGAGATATTGCGTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	80
CAAGCAGAAGACGGCATACGAGATACCTGGAAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	81
CAAGCAGAAGACGGCATACGAGATGGAGATGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	82
CAAGCAGAAGACGGCATACGAGATGTACTCTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	83
CAAGCAGAAGACGGCATACGAGATGTAACGACGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	84
CAAGCAGAAGACGGCATACGAGATATTCCTCCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	85
CAAGCAGAAGACGGCATACGAGATGTGTTCCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	86
CAAGCAGAAGACGGCATACGAGATAAGCACTGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	87
CAAGCAGAAGACGGCATACGAGATCTAGCAAGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	88
CAAGCAGAAGACGGCATACGAGATTGCTTCCAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	89
CAAGCAGAAGACGGCATACGAGATGCTTAGCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	90

CAAGCAGAAGACGGCATACGAGATAACCGTTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	91
CAAGCAGAAGACGGCATACGAGATGACATTCCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	92
CAAGCAGAAGACGGCATACGAGATAGACCGTAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	93
CAAGCAGAAGACGGCATACGAGATGATACTGGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	94
CAAGCAGAAGACGGCATACGAGATTGCGTAGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	95
CAAGCAGAAGACGGCATACGAGATTCGGTTACGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	96
CAAGCAGAAGACGGCATACGAGATATGACGTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	97
CAAGCAGAAGACGGCATACGAGATGCTGTAAGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	98
CAAGCAGAAGACGGCATACGAGATGCAATGGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	99
CAAGCAGAAGACGGCATACGAGATATCTCGCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	100
CAAGCAGAAGACGGCATACGAGATGGCTATTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	101
CAAGCAGAAGACGGCATACGAGATGGTGTCTTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	102
CAAGCAGAAGACGGCATACGAGATTCAACTGGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	103
CAAGCAGAAGACGGCATACGAGATCTTCACCAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	104
CAAGCAGAAGACGGCATACGAGATACGGTCTTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	105
CAAGCAGAAGACGGCATACGAGATTCTCGCAAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	106
CAAGCAGAAGACGGCATACGAGATGGAATTGCGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	107
CAAGCAGAAGACGGCATACGAGATACGGATTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	108
CAAGCAGAAGACGGCATACGAGATTTAAGCGGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	109
CAAGCAGAAGACGGCATACGAGATTGCAGGTAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	110
CAAGCAGAAGACGGCATACGAGATCAATCGACGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	111
CAAGCAGAAGACGGCATACGAGATGTGCCATAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	112
CAAGCAGAAGACGGCATACGAGATTGTTCGAGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	113
CAAGCAGAAGACGGCATACGAGATTGGAGTTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	114
CAAGCAGAAGACGGCATACGAGATACGATGACGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	115

CAAGCAGAAGACGGCATACGAGATTGATGTCCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	116
CAAGCAGAAGACGGCATACGAGATTGAACCTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	117
CAAGCAGAAGACGGCATACGAGATCTTCGTTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	118
CAAGCAGAAGACGGCATACGAGATCTTCTGAGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	119
CAAGCAGAAGACGGCATACGAGATTGCTCATGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	120
CAAGCAGAAGACGGCATACGAGATAGTTCGTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	121
CAAGCAGAAGACGGCATACGAGATTAGCGTCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	122
CAAGCAGAAGACGGCATACGAGATGGCGTTATGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	123
CAAGCAGAAGACGGCATACGAGATGGTGATTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	124
CAAGCAGAAGACGGCATACGAGATAACTTGCCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	125
CAAGCAGAAGACGGCATACGAGATGCAAGATCGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	126
CAAGCAGAAGACGGCATACGAGATTCGCATTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	127
CAAGCAGAAGACGGCATACGAGATTGTACACCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	128
CAAGCAGAAGACGGCATACGAGATAGCTCCTAGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	129
CAAGCAGAAGACGGCATACGAGATGCAATTCGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	130
CAAGCAGAAGACGGCATACGAGATCTTAGGACGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	131
CAAGCAGAAGACGGCATACGAGATGTCCTAAGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	132
CAAGCAGAAGACGGCATACGAGATAACGTGGAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	133
CAAGCAGAAGACGGCATACGAGATCTGTGTTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	134
CAAGCAGAAGACGGCATACGAGATGTTAAGGCGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	135
CAAGCAGAAGACGGCATACGAGATCACCTTACGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	136
CAAGCAGAAGACGGCATACGAGATTGGTAGCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	137
CAAGCAGAAGACGGCATACGAGATCAGTGAAGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	138
CAAGCAGAAGACGGCATACGAGATGTTCAACCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	139
CAAGCAGAAGACGGCATACGAGATTGGCTATCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	140

CAAGCAGAAGACGGCATACGAGATCTGGAGTAGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	141
CAAGCAGAAGACGGCATACGAGATTCTCTTCCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	142
CAAGCAGAAGACGGCATACGAGATTCTAACGCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	143
CAAGCAGAAGACGGCATACGAGATGGTCAGATGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	144
CAAGCAGAAGACGGCATACGAGATCTCTGGTTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	145
CAAGCAGAAGACGGCATACGAGATTGTGGTACGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	146
CAAGCAGAAGACGGCATACGAGATCCTATACCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	147
CAAGCAGAAGACGGCATACGAGATTTCTCTCGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	148
CAAGCAGAAGACGGCATACGAGATGTATGCTGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	149
CAAGCAGAAGACGGCATACGAGATAAGTCGAGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	150
CAAGCAGAAGACGGCATACGAGATAACCGAAGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	151
CAAGCAGAAGACGGCATACGAGATTGTTGTGGGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	152

[00370] The PCR program used to amplify the purified extension products comprised the following steps:

(1) 30 seconds at 98°C

(2) 10 seconds at 98°C

(3) 30 seconds at 65°C

(4) 30 seconds at 72°C

(5) Repeat steps (2)-(4) eighteen times

(6) 2 minutes at 72°C

[00371] The amplified extension products were stored at 4°C. 4 µL of each PCR reaction were combined into four pools, for 24 samples per pool.

[00372] The pooled PCR reactions were purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 100 µL of AMPure XP beads and 400 µL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). Purification was performed with 76.8 µL of diluted AMPure XP beads and eluted in 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. Beads were retained and cleanup process was repeated

with 24 μL of AMPure XP buffer and eluted in 20 μL of buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00373] The purified PCR products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The purified PCR products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with a custom spike-in primer comprising the nucleotide sequence ACACTCTTTAAGACGACGTCGCTATGGCCTCTCC (SEQ ID NO: 25).

[00374] Example 4—long probe hybridization method for 96 multiplexed samples

[00375] In this example, a long probe hybridization method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. In this example, the first nucleic acid probe comprises a 5' phosphate moiety, a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a first amplification primer binding site comprising an i7 sequence (SEQ ID NO: 1), unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a P7 flow cell adapter sequence (SEQ ID NO: 4). The second nucleic acid probe comprises a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier, a second amplification primer binding site comprising and i5 sequence (SEQ ID NO: 2) and a P5 flow cell adapter sequence (SEQ ID NO: 3).

[00376] The first and second nucleic acid probes were ordered from Integrated DNA Technologies, Inc. and quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific). Individual nucleic acid probes were normalized to a standard concentration, pooled to be equimolar, and diluted to 0.83 nM in a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The nucleic acid probes and the identifier oligonucleotides were hybridized by combining 0.5 μL of diluted nucleic acid probe pool with 2 μL of a mixture of identifier oligonucleotides collected from a sample solution in a buffer comprising 50 mM NaCl. This mixture was heated for 2 minutes at 95°C and cooled for 30 minutes at ambient temperature to yield an annealed identifier oligonucleotide-nucleic acid probe mixture.

[00377] In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were not adjacent and were not overlapping, a gap extension reaction was performed. To 2.5 μL of each annealed

identifier oligonucleotide-nucleic acid probe mixture, 3.8 μL of NEBNext Ultra II Q5 Master mix and 1.3 μL of PCR-grade H_2O was added. The mixture was then subjected to the following Gap extension temperature cycle:

(1) 30 seconds at 98°C

(2) 1 minute at 98°C

(3) 1 minute at 68°C

(4) 1 minute at 72°C,

(5) Repeat steps (2)-(4) ten times

(6) 2 minutes at 72°C

[00378] The gap extension products were then stored at 4°C. The first and second nucleic acid probes were then ligated together by adding to 1 μL of the gap extension product, 10 μL of 2X rapid ligation buffer (Enzymatics), 1 μL of T4 DNA Rapid Ligase (Enzymatics), and 8 μL of PCR-grade H_2O. These ligation reactions were incubated for 15 minutes 20°C, subsequently quenched with 1 μL 0.5M EDTA, and pooled into a single 15 mL conical tube.

[00379] In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were adjacent and were not overlapping, nick repair ligation reaction was performed. To 2.5 μL of each annealed identifier oligonucleotide-nucleic acid probe mixture, 10 μL of 2X rapid ligation buffer (Enzymatics), 1 μL of T4 DNA Rapid Ligase (Enzymatics), and 1 μL of PCR-grade H_2O was added. These ligation reactions were incubated for 15 minutes at 20°C, subsequently quenched with 1 μL of 0.5M EDTA, and pooled into a single 15 mL conical tube.

[00380] The pools of quenched ligation reactions were then purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350 μL of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). The purification was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200 μL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The AMPure XP bead cleanup was then repeated with 400 μL of AMPure XP beads and eluted in 20 μL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00381] To amplify the purified ligation products, PCR reactions with purified ligation products and primers were prepared. To 6 μL of purified ligation product, 10 μL of NEBNext Ultra II Q5 Master Mix, 0.2 μL of 100 μM forward primer (CAAGCAGAAGACGGCATACGA, SEQ ID

NO: 153) and reverse primer (AATGATACGGCGACCACCGA, SEQ ID NO: 154) and 3.6 µL of PCR-grade H_2O was added. The PCR program used to amplify the purified extension products comprised the following steps:

(1) 30 seconds at 98°C

(2) 10 seconds at 98°C

(3) 30 seconds at 65°C

(4) 30 seconds at 72°C

(5) Repeat steps (2)-(4) eighteen times

(6) 2 minutes at 72°C

[00382] The amplified products were stored at 4°C. 4 µL of each PCR reaction were combined into six pools, for 16 samples per pool. The amplified products were further purified using an AMPure XP bead cleanup with 64 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00383] The purified amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The purified amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with either standard sequencing primers or a custom spike-in Read1 primer (SEQ ID NO: 25).

[00384] Example 5—short probe hybridization method for 96 multiplexed samples

[00385] In this example, a short probe hybridization method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. In this example, the first nucleic acid probe comprises a 5' phosphate moiety, a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a first amplification primer binding site comprising an i7 sequence (SEQ ID NO: 1) and a unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. The second nucleic acid probe comprises a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding site comprising an i5 sequence (SEQ ID NO: 2).

[00386] The first and second nucleic acid probes were ordered from Integrated DNA Technologies, Inc. and quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher

Scientific). Individual nucleic acid probes were normalized to a standard concentration, pooled to be equimolar, and diluted to 0.83 nM in a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The nucleic acid probes and the identifier oligonucleotides were hybridized by combining 0.5 μL of diluted nucleic acid probe pool with 2 μL of a mixture of identifier oligonucleotides collected from a sample solution in a buffer comprising 50 mM NaCl. This mixture was heated for 2 minutes at 95°C and cooled for 30 minutes at ambient temperature to yield an annealed identifier oligonucleotide-nucleic acid probe mixture.

[00387] In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were not adjacent and were not overlapping, a gap extension reaction was performed. To 2.5 μL of each annealed identifier oligonucleotide-nucleic acid probe mixture, 3.8 μL of NEBNext Ultra II Q5 Master mix and 1.3 μL of PCR-grade H_2O was added. The mixture was then subjected to the following Gap extension temperature cycle:

(1) 30 seconds at 98°C

(2) 1 minute at 98°C

(3) 1 minute at 68°C

(4) 1 minute at 72°C,

(5) Repeat steps (2)-(4) ten times

(6) 2 minutes at 72°C

[00388] The gap extension products were then stored at 4°C. The first and second nucleic acid probes were then ligated together by adding to 1 μL of the gap extension product, 10 μL of 2X rapid ligation buffer (Enzymatics), 1 μL of T4 DNA Rapid Ligase (Enzymatics), and 8 μL of PCR-grade H_2O. These ligation reactions were incubated 15 min 20°C, quenched with 1 μL 0.5M EDTA, and pooled into a single 15 mL conical tube.

[00389] In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were adjacent and were not overlapping, nick repair ligation reaction was performed. To 2.5 μL of each annealed identifier oligonucleotide-nucleic acid probe mixture, 10 μL of 2X rapid ligation buffer (Enzymatics), 1 μL of T4 DNA Rapid Ligase (Enzymatics), and 1 μL of PCR-grade H_2O was added. These ligation reactions were incubated for 15 minutes at 20°C, subsequently quenched with 1 μL of 0.5M EDTA, and pooled into a single 15 mL conical tube.

[00390] The pools of quenched ligation reactions were then purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350 µL of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). The purification was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The AMPure XP bead cleanup was then repeated with 400 µL of AMPure XP beads and eluted in 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00391] To amplify the purified ligation products, PCR reactions with purified ligation products and primers were prepared. To 6 µL of purified ligation product, 10 µL of NEBNext Ultra II Q5 Master Mix, 0.2 µL of 100 µM forward primer and reverse primer and 3.6 µL of PCR-grade H_2O was added. . The forward primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the first strand of the nucleic acid adapter. Table 8 provides the sequences of the forward primers used.

Table 8. Forward primers for short probe hybridization method.

Primer Sequence	SEQ ID NO
CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	13
CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	14
CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	15
CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	16
CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	17
CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	18
CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	19
CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	20
CAAGCAGAAGACGGCATACGAGATCATAGCCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	21
CAAGCAGAAGACGGCATACGAGATTTCGCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	22
CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	23

CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATCT	24

[00392] The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second amplification primer binding site located on the second strand of the nucleic acid adapter. Table 9 provides the sequences of the reverse primers used.

Table 9. Reverse primers for short probe hybridization method.

Primer Sequence	SEQ ID NO
AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCTACA CGACGCTCTTCCGATCT	155
AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCCTAC ACGACGCTCTTCCGATCT	156
AATGATACGGCGACCACCGAGATCTACACCCTATCCTACACTCTTTCCCTACA CGACGCTCTTCCGATCT	157
AATGATACGGCGACCACCGAGATCTACACGGCTCTGAACACTCTTTCCCTACA CGACGCTCTTCCGATCT	158
AATGATACGGCGACCACCGAGATCTACACAGGCGAAGACACTCTTTCCCTAC ACGACGCTCTTCCGATCT	159
AATGATACGGCGACCACCGAGATCTACACTAATCTTAACACTCTTTCCCTACA CGACGCTCTTCCGATCT	160
AATGATACGGCGACCACCGAGATCTACACCAGGACGTACACTCTTTCCCTACA CGACGCTCTTCCGATCT	161
AATGATACGGCGACCACCGAGATCTACACGTACTGACACACTCTTTCCCTACA CGACGCTCTTCCGATCT	162

[00393] The PCR program used to amplify the purified extension products comprised the following steps:

(1) 30 seconds at 98°C

(2) 10 seconds at 98°C

(3) 30 seconds at 65°C

(4) 30 seconds at 72°C

(5) Repeat steps (2)-(4) eighteen times

(6) 2 minutes at 72°C

[00394] The amplified products were stored at 4°C. 4 µL of each PCR reaction were combined into six pools, for 16 samples per pool. The amplified products were further purified using an AMPure XP bead cleanup with 64 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00395] The purified amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The purified amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with either standard sequencing primers or a custom spike-in Read1 primer (SEQ ID NO: 25).

[00396] Example 6—direct PCR method for 96 multiplexed samples

[00397] In this example, a direct PCR method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. In this example, 8 species of forward amplification primers and 12 species of reverse amplification primers were used. The forward primers comprised a P5 flow cell adapter (SEQ ID NO: 3), a nucleic acid sequence comprising a unique molecular identifier and a region complementary to a first amplification primer binding site present on the identifier oligonucleotide. Table 10 provides the sequences of the forward amplification primers used.

Table 10. Forward amplification primers for direct PCR method.

Primer Sequence	SEQ ID NO
AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	5
AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	6
AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	7
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	8
AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	9
AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	10
AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTT TAAGACGACGTCGCTATGGCCTCTCC	11
AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTT AAGACGACGTCGCTATGGCCTCTCC	12

[00398] The reverse primers comprised a P7 flow cell adapter (SEQ ID NO: 4), a nucleic acid sequence comprising a unique molecular identifier and a region complementary to a second amplification primer binding site present on the identifier oligonucleotide. Table 11 provides the sequences of the reverse amplification primers used.

Table 11. Reverse amplification primers for direct PCR method.

Primer Sequence	SEQ ID NO
CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	163
CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	164
CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	165
CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	166
CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	167
CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	168
CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	169
CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	170
CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	171
CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	172
CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	173
CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT	174

[00399] In the case of 8 forward amplification primers and 12 reverse amplification primers, when the unique molecular identifiers from a pair of forward and reverse primers are combined, a total of 96 unique combinations can be obtained, allowing for the multiplexing of 96 samples.

[00400] To amplify the collected identifier oligonucleotides for sequencing, PCR reactions with collected identifier oligonucleotides and forward and reverse amplification primers were prepared on a 96-well plate with 2 μL of each identifier oligonucleotide sample, 10 μL of NEBNext Ultra II Q5 Master Mix, 2 μL of 10 μM forward amplification primer, 2 μL of reverse amplification primer, and 4 μL of PCR-grade H_2O. Each well in the 96-well plate contained an identifier oligonucleotide sample and a unique combination of forward and reverse amplification primers. The PCR program used comprised the following steps:

(1) 30 seconds at 98°C

(2) 10 seconds at 98°C

(3) 30 seconds at 65°C

(4) 30 seconds at 72°C

(5) Repeat steps (2)-(4) six to ten times

(6) 2 minutes at 72°C

[00401] The amplified products were stored at 4°C. 10 µL of each PCR reaction was combined into a single 15 mL conical tube.

[00402] The pooled PCR reactions were purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 115.2 µL of AMPure XP beads and 1036.8 µL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). Purification was performed with 1152 µL of diluted AMPure XP beads and eluted in 60 µL of a buffer comprising 10 mM Tris pH 8. The purification process was repeated with 60 µL of AMPure XP beads and eluted in 70 µL of a buffer comprising 10 mM Tris pH 8.

[00403] Following AMPure XP cleanup, PCR reactions with universal primers were prepared with 9 µL of pooled direct PCR product, 15 µL of NEBNext Ultra II Q5 Master Mix, 3 µL of 10 µM universal P7 primer (SEQ ID NO: 153) and 2 µL of 10 µM universal P5 primer (SEQ ID NO: 154). The PCR program used was:

(1) 30 seconds at 98°C,

(2) 10 seconds at 98°C

(3) 30 seconds at 65°C

(4) 30 seconds at 72°C

(5) Repeat steps (2)-(4) 15 to 24 times

(6) 2 minutes at 72°C

[00404] Two rounds of AMPure XP bead cleanup was performed. The first round was performed with 30 µL of beads and eluted with 20 µL of a buffer comprising 10 mM Tris pH 8 and the second round was performed with 20 µL beads and eluted with 11 µL of a buffer comprising 10 mM Tris pH 8.

[00405] These purified PCR products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics). The purified PCR products were also diluted for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with a custom spike-in primer (SEQ ID NO: 25).

[00406] Example 7—spatially detecting target analytes in a FFPE sample

[00407] The methods of the present invention were used to spatially detect a plurality of different target analytes, including target proteins and target RNAs, in a sample of inflamed human tonsil tissue FFPE section.

[00408] In one experiment, 30 different target proteins were spatially detected using the methods of the present disclosure in two serial sections cut from the inflamed human tonsil tissue FFPE section. The 30 target proteins are put forth in Table 12. The 30 target proteins included IgG Rabbit isotype and IgG Mouse isotype as negative controls, as these target proteins should not have been present in the inflamed human tonsil sample and therefore should not have been detected.

Table 12. Target Proteins

Target Protein	Target Protein
AKT	FOXP3
B7-H3	GZMB
Bcl-2	Histone H3
Beta-2-microglobulin	Ki67
Beta-catenin	CD20
CD14	P-AKT
CD19	PanCK
CD3	PD1
CD4	PD-L1
CD44	S6
CD45	STAT3
CD45RO	P-STAT3
CD56	VISTA
CD68	IgG Rabbit isotype (negative control)
CD8A	IgG Mouse isotype (control)

[00409] To spatially detect the 30 target proteins, 30 different probes of the present disclosure were used. Each probe comprised a target binding domain comprising an antibody that specifically binds to one of the 30 target proteins in Table 12. The two serial sections were contacted with a plurality of the 30 different probes. Ninety-six regions of interest (ROI) were then identified. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a short probe hybridization method of the present disclosure, thereby spatially detecting the 30 target proteins in the two serial sections. As shown

in Figure 21, the number of reads per target protein in each ROI for the two serial sections were well correlated, demonstrating that the method yields reproducible results.

[00410] In a second experiment, 20 different target RNAs were spatially detected using the methods of the present disclosure in two different serial sections cut from the inflamed human tonsil tissue FFPE section. The 20 different target RNAs are put forth in Table 13. The 20 target RNAs included 6 negative controls(Negative Probe) that should not have been detected in the sample.

Table 13. Target RNAs

Target RNA	Target RNA	Target RNA
CD3E	CD40	CTLA4
CD3G	CD45	GAPDH
CD4	CD74	KRT13
CD20	CD79A	PD1
PSA	RP56	Negative Probe #1
Negative Probe #2	Negative Probe #3	Negative Probe #4
Negative Probe #5	Negative Probe #6	

[00411] To spatially detect the 20 Target RNAs, 20 different probes of the present disclosure were used. Each probe comprised a target binding domain comprising a nucleic acid sequence complementary to at least one portion of one of the 20 target RNAs. The two serial sections were contacted with a plurality of the 20 different probes. Ninety-six regions of interest (ROI) were then identified. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a direct PCR method of the present disclosure, thereby spatially detecting the 20 target RNAs in the two serial sections. As shown in Figure 22, the number of reads per target RNA in each ROI for the two serial sections were well correlated, demonstrating that the method yields reproducible results.

[00412] Example 8—spatially detecting target proteins in a fluorescently stained FFPE sample

[00413] In another experiment, a 5 μm FFPE section of inflamed human tonsil tissue was stained with 4 fluorescent visualization markers: (1) CD3E, a T-cell marker; (2) PanCK, an epithelial cell marker; (3) Ki-67, a proliferation marker; and (4) SYTO83, a DNA stain, as shown in the left panel of Figure 23. The stained FFPE section was then contacted with the probes directed

against 30 target proteins, as described in Example 7. As shown in the left panel Figure 23, 96 regions of interest (ROIs) were selected. Each ROI was a circle with a 500 μm diameter. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a short probe hybridization method of the present disclosure, thereby spatially detecting the 30 target proteins in the FFPE section. As shown in the right panel of Figure 23, PanCK, CD3E and Ki67 were spatially detected in ROIs that correlated with their fluorescent visualization markers. Thus, the results generated by the methods of the present disclosure correlate with the results generated using established immunohistochemical method.

[00414] Example 9—spatially detecting target RNAs in a FFPE sample

[00415] In another experiment, a 5 μm section from an inflamed human tonsil tissue FFPE block was contacted with probes directed against 20 target RNAs, as described in Example 7. 96 regions of interest (ROIs) were then selected. Each ROI was a circle with a 500 μm diameter. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a direct PCR method of the present disclosure, thereby spatially detecting the 20 target RNAs in the two serial sections. The total RNA from a 20 μm section from the same inflamed human tonsil tissue FFPE block was then isolated. The total RNA was analyzed using the NanoString nCounter® system. Figure 24 shows that the average number of counts for 11 different RNA targets recorded using the methods of the present disclosure were well correlated with the average number of counts for the same 11 different RNA targets recorded using the nCounter® system. Thus, the results generated using the methods of the present disclosure correlate with the results generated using established direct detection methods.

[00416] Example 10—spatially detecting target RNAs in specific sub-regions of an ROI

[00417] In another experiment, a 5 μm section from an inflamed human tonsil tissue FFPE block was contacted with probes directed against 30 target proteins, as described in Example 7. The same 5 μm section was also stained with 4 fluorescent visualization markers: (1) CD3E, a T-cell marker; (2) PanCK, an epithelial cell marker; (3) Ki-67, a proliferation marker; and (4) SYTO83, a DNA stain. As shown in Figure 25, 48 regions of interest (ROIs) were identified. For each ROI, two sub-regions were then identified based on the fluorescent staining. Areas of an ROI that were fluorescently stained positive for PanCK (PanCK+) were designated a "tumor" sub-

region and the areas of an ROI that lacked PanCK fluorescent staining were designated a "micro-environment" sub-region, as shown in Figure 25. For each ROI, the tumor sub-region and the micro-environment sub-region were separately illuminated with UV light to release the identifier oligonucleotides from the probes bound within each sub-region by creating a custom mask based on the intensity of PanCK fluorescent staining. The released identifier oligonucleotides were also separately collected. The collected identifier oligonucleotides were then analyzed using the short-probe hybridization method of the present disclosure and the NanoString nCounter® system. As shown in the bottom panel of Figure 25, the results using the NanoString nCounter® system and the short-probe hybridization method of the present disclosure were well correlated. Furthermore, in the tumor sub-regions, PanCK was detected at a significantly higher level as compared to the micro-environment sub-regions. Thus, the spatial detection results provided by the methods of the present disclosure are consistent with established fluorescent immunohistochemical methods and allows for the spatial detection within highly specific regions of a sample.

[00418] Example 11—96-plex human immuno-oncology panel

[00419] A 96-plex human immuno-oncology panel was designed for use in the direct-PCR methods of the present invention. The panel comprised a plurality of probes that could be used to spatially detect 96 different human target RNAs using the direct-PCR methods of the present disclosure. The 96 target RNAs are shown in Table 14.

Table 14. Target RNAs

Target	Target	Target	Target	Target	Target	Target	Target
AKT1	CD3E	CEACAM8	FOXP3	IFNG	KRT1	NCAM1	SOD2
ARG1	CD4	CMKLR1	GZMB	IFNGR1	KRT10	NKG7	SOX10
B2M	CD40	CSF1R	H3F3A	IL10	KRT14	NT5E	STAT1
BATF3	CD40LG	CTLA4	HAVCR2	IL12B	KRT17	PDCD1	STAT2
BCL2	CD44	CTNNB1	HIF1A	IL15	KRT18	PDCD1LG2	STAT3
BCL2L1	CD47	CXCL10	HLA_DQA1	IL1B	KRT19	PECAM1	TBX21
CCL5	CD68	CXCL9	HLA_DRB	IL6	KRT6A	PMEL	TIGIT
CCND1	CD74	CSCR6	HLA_E	ITGAM	KRT7	PSMB10	TNF
CD14	CD86	DKK2	ICAM1	ITGAV	LAG3	PTEN	TNFRSF9
CD27	CD8A	EPCAM	ICOSLG	ITGAX	LY6E	PRTPRC	TNFSF4

CD274	CEACAM1	FAS	IDO1	ITGB2	MKI67	RPS6	VEGFA
CD276	CEACAM6	FASLG	IFNAR1	ITGB8	MS4A1	S100B	VSIR

[00420] In total, the panel comprised 928 different probes. Each of the probes comprised an identifier oligonucleotide comprising a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identified the target RNA bound to the target binding domain and a second amplification primer binding site. Figure 26 shows a schematic of the probes used in the panel. For each of the 96 target RNAs, there was at least one probe within the 928 probe set comprising a target binding domain that directly or indirectly hybridized to that target RNA. For most of the 96 target RNAs, there were 10 different probes that directly or indirectly hybridized to the specific target RNA. These 10 different probes directly or indirectly hybridized to different locations on the target RNA to create a "tiling" effect, as shown in the top panel of Figure 27. Tiling multiple probes onto a target RNA means that each target RNA will be individually detected multiple times, increasing the overall accuracy of the measurement. For example, as shown in the bottom panel of Figure 27, in the case where 10 probes are tiled onto a single target RNA, one of the probes may be incorrectly detected too many times (outlier high count probe), while another probe may be incorrectly detected too few times (outlier low count probe). However, the other 8 probes may be detected at a similar level, indicating that the two outliers should be discarded during analysis and the signals from the 8 probes used to generate a more accurate measurement of the abundance of the target RNA.

[00421] The set of 928 probes also comprised 80 negative control probes. Each of the 80 negative control probes comprised a target binding domain comprising a scrambled, non-specific nucleic acid sequence that was designed using guidelines from the External RNA Controls Consortium such that the target binding domain should not be complementary to RNA molecules present within a human sample. Thus, these 80 negative control probes should not be detected during analysis.

[00422] The 96-plex human immune-oncology panel was used to analyze a tissue microarray. The tissue microarray comprised FFPE samples of 22 common human cell lines, including normal and cancerous cell types. Some of the cell lines are shown in Table 15.

[00423] Table 15. Cell lines

Cell Line	Cell Line	Cell Line	Cell Line
CCRF-CEM	DAUDI	H596	H2228
HT29	HUT78	HUH7	JURKAT
M14	MDA-MB-468	MOLT4	RAJI
SKBR3	SUDHL1	SUDHL4	

[00424] The tissue microarray also comprised one mouse cell line (3T3) as a negative control. Each of the FFPE samples on the microarray was contacted with a plurality of the 928 different probes in the imunno-onocology panel. As shown in Figure 28, for each of the FFPE samples, at least three circular regions of interest (ROIs) with a diameter of 300 μm were selected. As a negative control, ROIs were also selected on regions of the microarray that did not comprise a FFPE sample (glass negative control). For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a direct PCR method of the present disclosure thereby spatially detecting the 96 target RNAs in each of the FFPE samples on the tissue microarray.

[00425] Figure 29 shows that a sufficient read depth was achieved using a MiSeq v3 flowcell. The top panel of Figure 30 shows that none of the target RNAs were spatially detected for the glass negative control ROIs. Likewise, the bottom panel of Figure 30 shows that nearly none of the target RNAs were detected in the negative control mouse 3T3 FFPE sample. Conversely, as shown in Figures 31, 33 and 34, specific target RNAs were successfully detected in the HEK293 (human embryonic kidney) FFPE sample and the Jurkat (human T-cell lymphocyte) FFPE sample. Figures 31, 33 and 34 show that clusters of "tiled" probes were detected for particular target RNAs, including AKT1, B2M, CD3E, HIF1A, PTEN, RPS6, STAT1, STAT2, STAT3, VEGF, PTPRC (CD45), and KRT1/10/18/19. These results indicated that there are certain target RNAs that are differentially transcribed in the two different cells lines. The results of this experiment were also verified using the NanoString nCounter system to identify the collected identifier oligonucleotides. As shown in Figure 27, the results using the direct-PCR method of the present disclosure were consistent with the results obtained using the NanoString nCounter system.

What is claimed is:

1. A method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising:

(1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain;

(2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide;

(3) collecting the released identifier oligonucleotide;

(4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe,

wherein the first nucleic acid probe comprises:

a nucleic acid complementary to a portion of the identifier oligonucleotide,

a nucleic acid sequence comprising a unique molecular identifier,

a first amplification primer binding site, and

wherein the second nucleic acid probe comprises:

a nucleic acid complementary to a portion of the identifier oligonucleotide, and

a second amplification primer binding site, and

wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping;

(5) ligating the hybridized first and second nucleic acid probes together;

(6) amplifying the ligation product produced in step (5); and

(7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

2. The method of claim 1, wherein amplifying the ligation product comprises performing PCR.

3. The method of claim 2, wherein performing PCR comprises a first amplification primer and second amplification primer,

wherein the first amplification primer comprises:

a first flow cell binding site;

a first nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; and

a nucleic acid sequence complementary to the second amplification primer binding site; and

wherein the second amplification primer comprises:

a second flow cell binding site;

a second nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released; and

a nucleic acid sequence complementary to the first amplification primer binding site.

4. The method of claim 3, wherein at least one of the amplification primers comprises an affinity molecule.

5. The method of claim 3, wherein at least one of the flow cell binding sites comprises a flow cell adapter sequence suitable for sequencing.

6. The method of claim 3, wherein at least one of the flow cell binding sites comprises a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3.

7. The method of claim 3, wherein at least one of the flow cell binding sites comprises a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

8. The method of claim 3, wherein at least one of the flow cell binding sites comprises about 15 to about 40 nucleotides

9. The method of claim 1, wherein the unique nucleic acid sequence which identifies the target analyte bound to the target binding domain comprises between about 10 nucleotides and about 50 nucleotides

10. The method of claim 1, wherein at least one of the amplification primer binding sites comprises about 20 to about 50 nucleotides.

11. The method of claim 1, wherein at least one of the amplification primer binding sites comprises an i7 sequence, wherein the i7 sequence comprises the sequence set forth in SEQ ID NO: 1.

12. The method of claim 1, wherein at least one of the amplification primer binding sites comprises an i5 sequence, wherein the i5 sequence comprises the sequence set forth in SEQ ID NO: 2.

13. The method of claim 3, wherein at least one of the nucleic acid sequences which identify the specific location of the tissue sample from which the identifier oligonucleotide was released comprises about 1 to about 20 nucleotides.

14. The method of claim 1, wherein the nucleic acid sequence comprising a unique molecular identifier comprises about 5 to about 20 nucleotides.

15. A method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising:

(1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises:

a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain,

a nucleic acid sequence comprising a unique molecular identifier,

a first amplification primer binding site, and

a second amplification primer binding site;

(2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide;

(3) collecting the released identifier oligonucleotide;

(4) amplifying the collected identifier oligonucleotide;

(5) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

16. The method of claim 15, wherein amplifying the collected identifier oligonucleotide comprises performing PCR.

17. The method of claim 16, wherein performing PCR comprises a first amplification primer and second amplification primer,

wherein the first amplification primer comprises:

a first flow cell binding site;

a first nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; and

a nucleic acid sequence complementary to the second amplification primer binding site; and

wherein the second amplification primer comprises:

a second flow cell binding site;

a second nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released; and

a nucleic acid sequence complementary to the first amplification primer binding site.

18. The method of claim 17, wherein at least one of the flow cell binding sites comprises a flow cell adapter sequence suitable for sequencing.

19. The method of claim 17, wherein at least one of the flow cell binding sites comprises a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3.

20. The method of claim 17, wherein at least one of the flow cell binding sites comprises a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

21. The method of claim 17, wherein at least one of the flow cell binding sites comprises about 15 to about 40 nucleotides

22. The method of claim 15, wherein the unique nucleic acid sequence which identifies the target analyte bound to the target binding domain comprises between about 5 nucleotides and about 25 nucleotides

23. The method of claim 15, wherein at least one of the amplification primer binding sites comprises about 10 to about 40 nucleotides.

24. The method of claim 17, wherein at least one of the amplification primer binding sites comprises an i7 sequence, wherein the i7 sequence comprises the sequence set forth in SEQ ID NO: 1.

25. The method of claim 17, wherein at least one of the amplification primer binding sites comprises an i5 sequence, wherein the i5 sequence comprises the sequence set forth in SEQ ID NO: 2.

26. The method of claim 17, wherein at least one of the nucleic acid sequences which identify the specific location of the tissue sample from which the identifier oligonucleotide was released comprises about 1 to about 20 nucleotides.

27. The method of claim 15, wherein the nucleic acid sequence comprising a unique molecular identifier comprises about 5 to about 20 nucleotides.

28. The method of claim 15, wherein the target binding domain comprises between about 10 nucleotides to about 70 nucleotides.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of:	Joseph M. Beechem *et al.*	Confirmation No.:	TBD
Serial No.:	TBD	Group Art Unit:	TBD
Filed:	Herewith	Examiner:	TBD

FOR: **BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION**

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

PETITION UNDER 37 CFR 1.84(a)(2) TO ACCEPT COLOR DRAWINGS

Color drawings have been submitted in this application. The color drawings are necessary as the only practical medium by which aspects of the claimed subject matter may be accurately conveyed and are necessary to illustrate features of the claims. No new matter has been added.

The first paragraph of the BRIEF DESCRIPTION OF THE DRAWINGS section of the Specification as filed herewith includes the following language:

> The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

The color drawings are being filed electronically *via* EFS-Web herewith, on February 11, 2019.

The requisite fee set forth in 37 CFR §1.17(h) has been submitted herewith. Applicants authorize the payment of any deficiency to be charged to Deposit Account No. 50-1283, Reference No. NATE-037/001US (321329-2523).

Dated: February 11, 2019

Respectfully submitted,
COOLEY LLP

COOLEY LLP
ATTN: Patent Group
1299 Pennsylvania Avenue NW, Suite 700
Washington, DC 20004

Tel: (617) 937-2340
Fax: (202) 842-7899

By: /Matthew Pavao/
Matthew Pavao
Reg. No. 50,572

199026549 v1

Electronic Patent Application Fee Transmittal

Application Number:	
Filing Date:	
Title of Invention:	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION
First Named Inventor/Applicant Name:	Joseph M. BEECHEM
Filer:	Matthew Pavao
Attorney Docket Number:	NATE-037001US 321329-2523

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
UTILITY APPLICATION FILING	1011	1	300	300
UTILITY SEARCH FEE	1111	1	660	660
UTILITY EXAMINATION FEE	1311	1	760	760
Pages:				
UTILITY APPL SIZE FEE PER 50 SHEETS >100	1081	1	400	400
Claims:				
CLAIMS IN EXCESS OF 20	1202	8	100	800
Miscellaneous-Filing:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
LATE FILING FEE FOR OATH OR DECLARATION	1051	1	160	160
Petition:				
PETITION FEE- 37 CFR 1.17(H) (GROUP III)	1464	1	140	140
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**3220**

Electronic Acknowledgement Receipt

EFS ID:	35107740
Application Number:	16272487
International Application Number:	
Confirmation Number:	6864
Title of Invention:	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION
First Named Inventor/Applicant Name:	Joseph M. BEECHEM
Customer Number:	141429
Filer:	Matthew Pavao
Filer Authorized By:	
Attorney Docket Number:	NATE-037001US 321329-2523
Receipt Date:	11-FEB-2019
Filing Date:	
Time Stamp:	16:07:08
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	$3220
RAM confirmation Number	021219INTEFSW00002877501283
Deposit Account	501283
Authorized User	Lisa Romano

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.16 (National application filing, search, and examination fees)

37 CFR 1.17 (Patent application and reexamination processing fees)

37 CFR 1.19 (Document supply fees)

37 CFR 1.20 (Post Issuance fees)

37 CFR 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal of New Application	NATE_037001US_Transmtital.pdf	103879 6812b812e2e70e9341439600c5236df6a29aee3d	no	1
Warnings:					
Information:					
2	Power of Attorney	NATE_037001US_POA.pdf	1461770 dfb59fa7dfbfdd688908414104d420e03f1e0c42	no	1
Warnings:					
Information:					
3	CRF Statement Paper and CRF are the same	NATE_037001US_Statement.pdf	105015 73ef574a7d3ad9b582bf7850a16326efadbcdc11	no	1
Warnings:					
Information:					
4	Drawings-other than black and white line drawings	NATE_037001US_Drawings.pdf	6828085 767e1673fb6dbceb5473f39e91916be3c29186d1	no	34
Warnings:					
Information:					
5	Sequence Listing (Text File)	NATE_037001US_Sequence_Listing.txt	49664	no	-
Warnings:					
Information:					
6	Application Data Sheet	NATE_037001US_ADS.pdf	1256104 bab30f62ddb5c51e755525912c24f63a9d6d2f86	no	10
Warnings:					
Information:					

7		NATE_037001US_Specification.pdf	469517 5ba0c083087243221be93f3836fd6d9ac279243f	yes	109

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Abstract	109	109
Claims	104	108
Specification	1	103

Warnings:					
Information:					
8	A petition from Applicant to accept Color Drawings or Color Photographs	NATE_037001US_Petition_Colored_Drawings.pdf	95766 6341bff28a02217fec82f614c6c35b560732c9b1	no	1
Warnings:					
Information:					
9	Fee Worksheet (SB06)	fee-info.pdf	41724 2398fafdcebc03da88da83564985ed09ec4a87c2	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	10411524		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

==========
Sequence Listing was accepted.
See attached Validation Report.
If you need help call the Patent Electronic Business Center at (866) 217-9197 (toll free).
Reviewer: Wheat Jr, Scott (ASRC)
Timestamp: [year=2019; month=2; day=14; hr=8; min=45; sec=1; ms=639;]
==========

Validated By CRFValidator v 1.0.5

Application No: 16272487 Version No: 1.0

Input Set:

Output Set:

Started: 2019-02-11 16:08:14.407
Finished: 2019-02-11 16:08:17.250
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 843 ms
Total Warnings: 792
Total Errors: 1
No. of SeqIDs Defined: 174
Actual SeqID Count: 174

Error code	Error Description
E 201	Mandatory field data missing in <140>
W 213	Artificial or Unknown found in <213> in SEQ ID (1)
W 213	Artificial or Unknown found in <213> in SEQ ID (2)
W 213	Artificial or Unknown found in <213> in SEQ ID (3)
W 213	Artificial or Unknown found in <213> in SEQ ID (4)
W 213	Artificial or Unknown found in <213> in SEQ ID (5)
W 213	Artificial or Unknown found in <213> in SEQ ID (6)
W 213	Artificial or Unknown found in <213> in SEQ ID (7)
W 213	Artificial or Unknown found in <213> in SEQ ID (8)
W 213	Artificial or Unknown found in <213> in SEQ ID (9)
W 213	Artificial or Unknown found in <213> in SEQ ID (10)
W 213	Artificial or Unknown found in <213> in SEQ ID (11)
W 213	Artificial or Unknown found in <213> in SEQ ID (12)
W 213	Artificial or Unknown found in <213> in SEQ ID (13)
W 213	Artificial or Unknown found in <213> in SEQ ID (14)
W 213	Artificial or Unknown found in <213> in SEQ ID (15)
W 213	Artificial or Unknown found in <213> in SEQ ID (16)
W 213	Artificial or Unknown found in <213> in SEQ ID (17)
W 213	Artificial or Unknown found in <213> in SEQ ID (18)
W 213	Artificial or Unknown found in <213> in SEQ ID (19)

Input Set:

Output Set:

Started: 2019-02-11 16:08:14.407
Finished: 2019-02-11 16:08:17.250
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 843 ms
Total Warnings: 792
Total Errors: 1
No. of SeqIDs Defined: 174
Actual SeqID Count: 174

Error code	Error Description
W 213	Artificial or Unknown found in <213> in SEQ ID (20) This error has occured more than 20 times, will not be displayed
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(27) on line number 393
W 447	n or Xaa used, for: SEQID(28) on line number 413
W 447	n or Xaa used, for: SEQID(28) on line number 413
W 447	n or Xaa used, for: SEQID(28) on line number 413
W 447	n or Xaa used, for: SEQID(28) on line number 413
W 447	n or Xaa used, for: SEQID(28) on line number 413 This error has occured more than 20 times, will not be displayed

SEQUENCE LISTING

<110> NanoString Technologies, Inc.
Beechem, Joseph
Kim, Dae
Hoang, Margaret
Gregory, Mark
Piazza, Erin
Zhou, Denise

<120> BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

<130> NATE-037/001US (321329-2523)

<140> US 16/272,487
<141> 2019-02-11

<150> US 62/629,180
<151> 2018-02-12

<150> US 62/771,212
<151> 2018-11-26

<160> 174

<170> PatentIn version 3.5

<210> 1
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> i7 sequence

<400> 1
agatcggaag agcacacgtc tgaactccag tcac 34

<210> 2
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> i5 sequence

<400> 2
acactctttc cctacacgac gctcttccga 30

<210> 3
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> P5 flow cell adapter

<400> 3
aatgatacgg cgaccaccga gatctacac 29

<210> 4
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> P7 flow cell adapter

<400> 4
atctcgtatg ccgtcttctg cttg 24

<210> 5
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i501 Primer

<400> 5
aatgatacgg cgaccaccga gatctacacg ctcagatata gcctacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 6
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i502 Primer

<400> 6
aatgatacgg cgaccaccga gatctacacg ctcagaatag aggcacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 7
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i503 Primer

<400> 7
aatgatacgg cgaccaccga gatctacacg ctcagaccta tcctacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 8
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i504 Primer

<400> 8
aatgatacgg cgaccaccga gatctacacg ctcagaggct ctgaacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 9
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i505 Primer

<400> 9
aatgatacgg cgaccaccga gatctacacg ctcagaaggc gaagacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 10
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i506 Primer

<400> 10
aatgatacgg cgaccaccga gatctacacg ctcagataat cttaacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 11
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i507 Primer

<400> 11
aatgatacgg cgaccaccga gatctacacg ctcagacagg acgtacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 12
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> DSP_NGS_i508 Primer

<400> 12
aatgatacgg cgaccaccga gatctacacg ctcagagtac tgacacactc tttaagacga 60

cgtcgctatg gcctctcc 78

<210> 13
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i701 Primer

<400> 13
caagcagaag acggcatacg agatcgagta atgtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 14
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i702 Primer

<400> 14
caagcagaag acggcatacg agattctccg gagtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 15
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i703 Primer

<400> 15
caagcagaag acggcatacg agataatgag cggtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 16
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i704 Primer

<400> 16
caagcagaag acggcatacg agatggaatc tcgtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 17
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i705 Primer

<400> 17
caagcagaag acggcatacg agatttctga atgtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 18
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i706 Primer

<400> 18
caagcagaag acggcatacg agatacgaat tcgtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 19
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i707 Primer

<400> 19
caagcagaag acggcatacg agatagcttc aggtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 20
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i708 Primer

<400> 20
caagcagaag acggcatacg agatgcgcat tagtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 21
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i709 Primer

<400> 21
caagcagaag acggcatacg agatcatagc cggtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 22
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i710 Primer

<400> 22
caagcagaag acggcatacg agatttcgcg gagtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 23
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i711 Primer

<400> 23
caagcagaag acggcatacg agatgcgcga gagtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 24
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> NEBNext i712 Primer

<400> 24
caagcagaag acggcatacg agatctatcg ctgtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 25
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Spike-in primer

<400> 25
acactcttta agacgacgtc gctatggcct ctcc 34

<210> 26
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Amplification primer binding site

<400> 26
gtgactggag ttcagacgtg tgctcttccg atct 34

<210> 27
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P01_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 27
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt tgaagcacac 60

cgtttttctt tcttctttca cgg 83

<210> 28
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P03_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 28
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna cccacaggtt 60

atacgggatt atccggttat cca 83

<210> 29
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P04_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 29
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc gacaccgagt 60

tcgaccgtta tgttggtagg atc 83

<210> 30
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P05_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 30
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc ggtgtgtaag 60

cgtaacgatg ttggtgtcgc tct 83

<210> 31
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P06_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 31
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc agacactgcg 60

acaactcacg atcatgacac aga 83

<210> 32
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P07_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 32
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna tattctgtac 60

tcagtgccta tccacctaat agg 83

<210> 33
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P09_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 33
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt tcagttataa 60

tgtgtccagc agaagcagga att 83

<210> 34
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P11_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 34
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnng tcctttgttg 60

ggcggaccgt aatgaggaat ttg 83

<210> 35
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P13_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 35
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnng atgagacttc 60

tacatgtccg atgtttttgt gct 83

<210> 36
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P15_PriExtUMI15_tpt

<220>

<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 36
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna ctcacacata 60

gtactgacac gtaagatagg atg 83

<210> 37
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P16_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 37
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt taccctatct 60

cgtctatgta cgtcaggctg aat 83

<210> 38
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P17_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 38
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna tcaacgtagg 60

gtaaggtcat atttttacct tac 83

<210> 39
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P18_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 39
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt tccctctttc 60

tccgcttatg gatgaaagga cag 83

<210> 40
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P19_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 40
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc ctgcacagtg 60

agtttctttc actctaactc tct 83

<210> 41
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P20_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 41
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt gtcgctctag 60

tgtgactttt ccacctcgca tct 83

<210> 42
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P21_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 42
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna tatctttctc 60

gggtaaagat taggcgtccg ata 83

<210> 43
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P23_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 43
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc gattagccgt 60

agacgcaact cattgccgaa gat 83

<210> 44
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P24_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 44
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt gtgagcattt 60

cagtacgagt gatgcagata aac 83

<210> 45

<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P25_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 45
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt atagttacca 60

agtactatgg gttggtggaa gcc 83

<210> 46
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P26_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 46
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc caattatact 60

gtctgttatg ttctcggata agc 83

<210> 47
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P27_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 47
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt caggtgcttg 60

taggctcatg ataggggtaa tgc 83

<210> 48
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P28_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 48
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc tctgctgtaa 60

tctcagctcc acttgtttct aag 83

<210> 49
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P29_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 49
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnng tgcatattgc 60

agctgagcca gctcaatttg aag 83

<210> 50
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P30_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 50

gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc cgttgattta 60

cgcaacagcg gcttatatag ctc 83

<210> 51
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P31_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 51
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc atcatcgaca 60

gttcgcagcc ctataacatg ata 83

<210> 52
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P32_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 52
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna tcgcaggatg 60

gtacagcatc atacatgatg agc 83

<210> 53
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P33_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)

<223> n is a, c, g, or t

<400> 53
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna tcgcaggatg 60

gtacagcatc atacatgatg agc 83

<210> 54
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P34_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 54
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnna tggcggtttc 60

gggtcctgca ctattcctaa taa 83

<210> 55
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P35_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 55
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnc cagtacgggt 60

actaataagt gtcatatcta ttg 83

<210> 56
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> P36_PriExtUMI15_tpt

<220>
<221> misc_feature
<222> (35)..(49)
<223> n is a, c, g, or t

<400> 56
gtgactggag ttcagacgtg tgctcttccg atctnnnnnn nnnnnnnnnt gttggagagg 60

ttagaggtga ggaggcgaag ata 83

<210> 57
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> Reverse primer

<400> 57
caagcagaag acggcatacg agatgtcggt aagtgactgg agttcagacg tgtgctcttc 60

cgatct 66

<210> 58
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> Reverse primer

<400> 58
caagcagaag acggcatacg agataggtca ctgtgactgg agttcagacg tgtgctcttc

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
16/272,487	02/11/2019	1634	3080	NATE-037001US 321329-2523	28	2

CONFIRMATION NO. 6864

141429
Cooley LLP/NanoString Technologies, Inc
1299 Pennsylvania Avenue NW
Suite 700
Washington, DC 20004

FILING RECEIPT

OC000000106074483

Date Mailed: 02/22/2019

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Joseph M. Beechem, Eugene, OR;
Dae Kim, Bellevue, WA;
Margaret Hoang, Seattle, WA;
Mark Gregory, Seattle, WA;
Erin Piazza, Edmonds, WA;
Denise Zhou, Seattle, WA;

Applicant(s)

NanoString Technologies, Inc., Seattle, WA

Power of Attorney: The patent practitioners associated with Customer Number 141429

Domestic Priority data as claimed by applicant

This appln claims benefit of 62/771,212 11/26/2018
and claims benefit of 62/629,180 02/12/2018

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.) - None.
Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

Permission to Access Application via Priority Document Exchange: Yes

Permission to Access Search Results: No

Applicant may provide or rescind an authorization for access using Form PTO/SB/39 or Form PTO/SB/69 as appropriate.

If Required, Foreign Filing License Granted: 02/21/2019

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 16/272,487**

Projected Publication Date: 08/15/2019

Non-Publication Request: No

Early Publication Request: No

Title

BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

Preliminary Class

435

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications: No

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific

countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop

technology, manufacture products, deliver services, and grow your business, visit http://www.SelectUSA.gov or call +1-202-482-6800.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 16/272,487

APPLICATION AS FILED - PART I

	(Column 1)	(Column 2)	SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
FOR	NUMBER FILED	NUMBER EXTRA	RATE($)	FEE($)		RATE($)	FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	300
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	660
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	760
TOTAL CLAIMS (37 CFR 1.16(i))	28 minus 20 =	* 8			OR	x 100 =	800
INDEPENDENT CLAIMS (37 CFR 1.16(h))	2 minus 3 =	*				x 460 =	0.00
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						400
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							0.00
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	2920

APPLICATION AS AMENDED - PART II

		(Column 1)		(Column 2)	(Column 3)	SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)		RATE($)	ADDITIONAL FEE($)
AMENDMENT A	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

		(Column 1)		(Column 2)	(Column 3)					
		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)		RATE($)	ADDITIONAL FEE($)
AMENDMENT B	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
16/272,487	02/11/2019	Joseph M. Beechem	NATE-037001US 321329-2523

CONFIRMATION NO. 6864

141429
Cooley LLP/NanoString Technologies, Inc
1299 Pennsylvania Avenue NW
Suite 700
Washington, DC 20004

INFORMAL NOTICE

OC000000106074484

Date Mailed: 02/22/2019

INFORMATIONAL NOTICE TO APPLICANT

Applicant is notified that the above-identified application contains the deficiencies noted below. No period for reply is set forth in this notice for correction of these deficiencies. However, if a deficiency relates to the inventor's oath or declaration, the applicant must file an oath or declaration in compliance with 37 CFR 1.63, or a substitute statement in compliance with 37 CFR 1.64, executed by or with respect to each actual inventor no later than the expiration of the time period set in the "Notice of Allowability" to avoid abandonment. See 37 CFR 1.53(f).

The item(s) indicated below are also required and should be submitted with any reply to this notice to avoid further processing delays.

- A properly executed inventor's oath or declaration has not been received for the following inventor(s):
  Joseph M. Beechem
  Dae Kim
  Margaret Hoang
  Mark Gregory
  Erin Piazza
  Denise Zhou

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at **(571) 272-4000** or **(571) 272-4200** or **1-888-786-0101**.

/afessehaye/

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of:	Joseph M. Beechem *et al.*	Confirmation No.:	6864
Application No.:	16/272,487	Group Art Unit:	1634
Filed:	February 11, 2019	Examiner:	Not Yet Assigned
FOR:	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION		

VIA EFS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

COMMUNICATION

Applicant respectfully submits six (6) executed Declaration documents (37 C.F.R. 1.63) for Utility Application Using an Application Data Sheet (37 C.F.R. 1.76) for each of the inventors named on the above-referenced application.

Applicant believes no fees are due in connection with this filing. However, the Director is hereby authorized to charge any deficiency in fees filed, asserted to be filed or which should have been filed herewith to our Deposit Account No. 50-1283.

Dated: March 7, 2019

Respectfully submitted,
COOLEY LLP

COOLEY LLP
ATTN: IP Docketing Department
1299 Pennsylvania Avenue NW, Suite 700
Washington, DC 20004

Tel: (617) 937-2340
Fax: (202) 842-7899

By: /Matthew Pavao/
Matthew Pavao, Ph.D.
Reg. No. 50,572

PTO/AIA/01 (06-12)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

As the below named inventor, I hereby declare that:

This declaration is directed to:

- [ ] The attached application, or
- [x] United States application or PCT international application number 16/272,487 filed on 02/11/2019 .

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Joseph M. BEECHEM Date (Optional): 4-May-2019

Signature:

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

PTO/AIA/01 (06-12)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

As the below named inventor, I hereby declare that:

This declaration is directed to:

- [ ] The attached application, or
- [x] United States application or PCT international application number 16/272,487 filed on 02/11/2019.

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Dae KIM Date (Optional): 3/4/2019

Signature: 

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

PTO/AIA/01 (06-12)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

As the below named inventor, I hereby declare that:

This declaration is directed to:

- [ ] The attached application, or
- [X] United States application or PCT international application number 16/272,487 filed on 02/11/2019.

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Margaret HOANG Date (Optional): 3/5/2019

Signature: Margaret Hoang

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

PTO/AIA/01 (08-12)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

As the below named inventor, I hereby declare that:

This declaration is directed to:

[ ] The attached application, or

[X] United States application or PCT international application number 16/272,487 filed on 02/11/2019 .

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

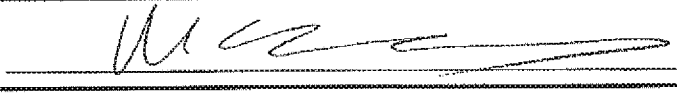
I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Mark GREGORY Date (Optional): 3/4/19

Signature: 

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

PTO/AIA/01 (06-12)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

As the below named inventor, I hereby declare that:

This declaration is directed to:

- ☐ The attached application, or
- ☒ United States application or PCT international application number 16/272,487 filed on 02/11/2019.

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Erin PIAZZA Date (Optional): 3/1/2019

Signature: Erin Piazza

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

PTO/AIA/01 (08-12)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION

As the below named inventor, I hereby declare that:

This declaration is directed to:

- [ ] The attached application, or
- [x] United States application or PCT international application number 16/272,487 filed on 02/11/2019.

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Denise ZHOU Date (Optional): March 5, 2019

Signature: Denise Zhou

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

Electronic Acknowledgement Receipt

EFS ID:	35350148
Application Number:	16272487
International Application Number:	
Confirmation Number:	6864
Title of Invention:	BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION
First Named Inventor/Applicant Name:	Joseph M. Beechem
Customer Number:	141429
Filer:	Matthew Pavao/Lisa Romano
Filer Authorized By:	Matthew Pavao
Attorney Docket Number:	NATE-037001US 321329-2523
Receipt Date:	07-MAR-2019
Filing Date:	11-FEB-2019
Time Stamp:	15:28:17
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Miscellaneous Incoming Letter	NATE_037001US_Communication.pdf	94827 130f5a53542b479697fc93a5fc631b0cba26dca9	no	1
Warnings:					

Information:					
2	Oath or Declaration filed	NATE_037001US_Declarations.pdf	2170955 aad0f5b571acdf8128070ad4ec1159b9613181d3	no	6
Warnings:					
Information:					
			Total Files Size (in bytes):	2265782	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/08a (07-09)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Complete if Known	
Application Number	16/272,487
Filing Date	02-11-2019
First Named Inventor	Joseph M. BEECHEM
Art Unit	1634
Examiner Name	Not Yet Assigned
Attorney Docket Number	NATE-037/001US 321329-2523

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	1.	5,320,814	06-14-1994	Walt et al.	
	2.	5,543,838	08-06-1996	Hosier et al.	
	3.	5,635,352	06-03-1997	Urdea et al.	
	4.	6,207,392	03-27-2001	Weiss et al.	
	5.	6,429,027	08-06-2002	Chee et al.	
	6.	6,596,257	07-22-2003	Bryan	
	7.	6,602,661	08-05-2003	Knezevic et al.	
	8.	6,690,470	02-10-2004	Baer et al.	
	9.	6,942,968	09-13-2005	Dickinson et al.	
	10.	7,060,507	06-13-2006	Akeson et al.	
	11.	7,132,519	11-07-2006	Monforte et al.	
	12.	7,214,477	05-08-2007	Emmert-Buck	
	13.	7,402,399	07-22-2008	Mukherjeei et al.	
	14.	7,569,392	08-04-2009	Levy et al.	
	15.	7,648,828	01-19-2010	Chan-Hui et al.	
	16.	7,728,287	06-01-2010	Felton et al.	
	17.	8,003,312	08-23-2011	Krutzik et al.	
	18.	8,221,972	07-17-2012	Lemaire et al.	
	19.	8,309,306	11-13-2012	Nolan et al.	
	20.	8,349,574	01-08-2013	Bates et al.	
	21.	8,362,415	01-29-2013	Felton et al.	
	22.	8,486,623	07-16-2013	Monforte et al.	

Examiner Signature:		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached. This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

PTO/SB/08a (07-09)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Complete if Known	
Application Number	16/272,487
Filing Date	02-11-2019
First Named Inventor	Joseph M. BEECHEM
Art Unit	1634
Examiner Name	Not Yet Assigned
Attorney Docket Number	NATE-037/001US 321329-2523

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	23.	8,637,650	01-28-2014	Cherkasov et al.	
	24.	8,753,824	06-17-2014	Papin et al.	
	25.	8,865,414	10-21-2014	Hennig et al.	
	26.	8,906,700	12-09-2014	Lim et al.	
	27.	9,046,477	06-02-2015	Emedcoles et al.	
	28.	9,228,948	01-05-2016	Emedcoles et al.	
	29.	9,297,762	03-29-2016	Emedcoles et al.	
	30.	9,304,084	04-05-2016	Emedcoles et al.	
	31.	9,376,678	06-28-2016	Gormley et al.	
	32.	2001/0053334	12-20-2001	Chen et al.	
	33.	2002/0028457	03-07-2002	Empedocles et al.	
	34.	2002/0034737	03-21-2002	Drmanac	
	35.	2002/0034827	03-21-2002	Singh et al.	
	36.	2002/0039728	04-04-2002	Kain et al.	
	37.	2002/0051971	05-02-2002	Stuelpnagel et al.	
	38.	2002/0068018	06-06-2002	Pepper et al.	
	39.	2002/0177141	11-28-2002	Chee et al.	
	40.	2002/0187515	12-12-2002	Chee et al.	
	41.	2003/0008323	01-09-2003	Ravkin et al.	
	42.	2003/0013091	01-16-2003	Dimitrov	
	43.	2003/0028981	02-13-2003	Chandler et al.	
	44.	2003/0186426	10-02-2003	Brewer et al.	

Examiner Signature:		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached. This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

PTO/SB/08a (07-09)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Complete if Known	
Application Number	16/272,487
Filing Date	02-11-2019
First Named Inventor	Joseph M. BEECHEM
Art Unit	1634
Examiner Name	Not Yet Assigned
Attorney Docket Number	NATE-037/001US 321329-2523

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	45.	2004/0000519	01-01-2004	Yong et al.	
	46.	2005/0048498	03-03-2005	Woudenberg et al.	
	47.	2005/0233318	10-20-2005	Chee et al.	
	48.	2007/0166708	07-19-2007	Dimitrov et al.	
	49.	2008/0038725	02-14-2008	Luo et al.	
	50.	2008/0038734	02-14-2008	Sorge et al.	
	51.	2010/0015607	01-21-2010	Geiss et al.	
	52.	2010/0047924	02-25-2010	Webster et al.	
	53.	2010/0112710	05-06-2010	Geiss et al.	
	54.	2010/0151472	06-17-2010	Nolan et al.	
	55.	2010/0209913	08-19-2010	Endress et al.	
	56.	2010/0261026	10-14-2010	Ferree et al.	
	57.	2010/0262374	10-14-2010	Hwang et al.	
	58.	2011/0086774	04-14-2011	Dunaway et al.	
	59.	2011/0151451	06-23-2011	Lemaire et al.	
	60.	2011/0223613	09-15-2011	Gut	
	61.	2011/0245111	10-06-2011	Chee	
	62.	2013/0023433	01-24-2013	Luo et al.	
	63.	2013/0178372	07-11-2013	Geiss et al.	
	64.	2014/0120532	05-01-2014	Lee et al.	
	65.	2014/0120550	05-01-2014	Baranov	
	66.	2014/0121117	05-01-2014	Tanner	

Examiner Signature:		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached. This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

PTO/SB/08a (07-09)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Complete if Known	
Application Number	16/272,487
Filing Date	02-11-2019
First Named Inventor	Joseph M. BEECHEM
Art Unit	1634
Examiner Name	Not Yet Assigned
Attorney Docket Number	NATE-037/001US 321329-2523

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	67.	2014/0371088	12-18-2014	Webster	
	68.	2015/0080233	03-19-2015	Bendall et al.	
	69.	2015/0099650	04-09-2015	Sood et al.	
	70.	2015/0132763	05-14-2015	Amorese et al.	
	71.	2015/0141297	05-21-2015	Lim et al.	
	72.	2015/0287578	10-08-2015	Bendall et al.	
	73.	2015/0329852	11-19-2015	Nolan	
	74.	2015/0368697	12-24-2015	Samusik et al.	
	75.	2016/0194701	07-07-2016	Beechem et al.	
	76.	2016/0362730	12-15-2016	Alexander et al.	
	77.	2017/0016053	01-19-2017	Beechem et al.	
	78.	2017/0016909	01-19-2017	Beechem et al.	
	79.	2017/0275669	09-28-2017	Weissleder et al.	
	80.	2018/0142286	05-24-2018	Dunaway et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]
	81.	WO 97/14028 A2	04-17-1997	Luminex Corporation		
	82.	WO 97/07245 A1	02-27-1997	Rabani, Ely		
	83.	WO 99/18434 A1	04-15-1999	Trustees of Tufts College		

Examiner Signature:		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached. This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Complete if Known	
Application Number	16/272,487
Filing Date	02-11-2019
First Named Inventor	Joseph M. BEECHEM
Art Unit	1634
Examiner Name	Not Yet Assigned
Attorney Docket Number	NATE-037/001US 321329-2523

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]
	84.	WO 00/73777 A1	12-07-2000	Bioarray Solutions, LLC		
	85.	WO 01/00875 A2	01-04-2001	Motorola, Inc		
	86.	WO 2012/106385 A2	08-09-2012	Apprise Bio, Inc.		
	87.	WO 2013/122996 A1	08-22-2013	The Johns Hopkins University		
	88.	WO 2014/200767 A1	12-18-2014	The General Hospital Corporation		
	89.	WO 2015/128272 A2	09-03-2015	Ventana Medical Systems, Inc.		
	90.	WO 2017/015099 A1	01-26-2017	Nanostring Technologies, Inc.		

NON PATENT LITERATURE DOCUMENTS

Examiner's Initials	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title ofthe item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	91.	AGASTI, S. S. et al., "Photocleavable DNA Barcode - Antibody Conjugates Allow Sensitive andn Multiplexed Protein Analysis in Single Cells," J. Am. Chem. Soc., 134:18499-18502 (2012).	
	92.	ALFANO, R. R. et al., "Optical Sensing, Imaging, and Manipulation for Biological and Biomedical Applications," SPIE-The International Society for Optical Engineering, July 2000, Vol. 4082, Taiwan, 342 pages.	
	93.	ANGELO, M. et al., "Multiplexed ion beam imaging of human breast tumors," Nature Medicine, 20(4):436-442 (2014).	
	94.	CESANO, A. et al., "Abstract 1371: Spatially-resolved, multiplexed digital characterizatio nof protein distribution and abundance in FFPE tissue sections," AACR 107th Annual Meeting, April 16-20, 2016, 5 pages.	

Examiner Signature:		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached. This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Complete if Known	
Application Number	16/272,487
Filing Date	02-11-2019
First Named Inventor	Joseph M. BEECHEM
Art Unit	1634
Examiner Name	Not Yet Assigned
Attorney Docket Number	NATE-037/001US 321329-2523

NON PATENT LITERATURE DOCUMENTS

Examiner's Initials	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title ofthe item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	95.	CHOW, S. et al., "Measurement of MAP Kinase Activation by Flow Cytometry Using Phospho-Specific Antibodies to MEK and ERK: Potential for Pharmacodynamic Monitoring of Signal Transduction Inhibitors," Cytometry (Communications in Clinical Cytometry), 46:72-78 (2001).	
	96.	CROSSETTO, N. et al., "Spatially resolved transcriptomics and beyond," Nature Reviews Genetics, 16:57-66 (2015).	
	97.	DIERCK, K. et al., "Quantitative multiplexed profiling of cellular signaling networks using phosphotyrosine-specific DNA-tagged SH2 domains," Nature Methods, 3:737-744 (2006).	
	98.	DRUMMOND, E. S. et al., "Proteomic analysis of neurons microdissected from formalinfixed, paraffin-embedded Alzheimer's disease brain tissue," Scientific Reports, 5:15456, 8 pages. doi: 10.1038/srep15456.	
	99.	FERGUSON, J. A. et al., "High-Density Fiber-Optic DNA Random Microsphere Array," Analytical Chemistry, 72(22):5618-5624 (2000).	
	100.	FREI, A. P. et al., "Highly multiplexed simultaneous detection of RNAs and proteins in single cells," Nature Methods, 13:269-275 (2016).	
	101.	GEISS, G. K. et al., "Direct multiplexed measurement of gene expression with color-coded probe pairs," Nature Biotechnology, 26(3):317-326 (2008).	
	102.	GULLBERG, M. et al., "Cytokine detection by antibody-based proximity ligation," PNAS, 101(22):8420-8424 (2004).	
	103.	LIND, K. & KUBISTA, M., "Development and evaluation of three real-time immuno-PCR assemblages for quantification of PSA," Journal of Immunological Methods, 304:107-116 (2005).	
	104.	Poster #3434, Power Point Presentation, "Digital spatial profiling platform allows for spatially-resolved, high-plex quantification of mRNA distribution and abundance on FFPE and fresh frozen tissue sections," AACR Annual Meeting, April 14-18, 2018, 1 page.	

Examiner Signature:		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached. This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

PTO/SB/08a (07-09)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Complete if Known	
Application Number	16/272,487
Filing Date	02-11-2019
First Named Inventor	Joseph M. BEECHEM
Art Unit	1634
Examiner Name	Not Yet Assigned
Attorney Docket Number	NATE-037/001US 321329-2523

NON PATENT LITERATURE DOCUMENTS

Examiner's Initials	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title ofthe item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	105.	STEEMERS, F. J. et al., "Screening unlabeled DNA targets with randomly ordered fiber-optic gene arrays," Nature Biotechnology, 18:91-94 (2000).	
	106.	THIERY-LAVENANT, G. et al., "Targeted Multiplex Imaging Mass Spectrometry in Transmission Geometry for Subcellular Spatial Resolution," Journal of the American Society for Mass Spectrometry, 24:609-614 (2013).	
	107.	ULLAL, A. V. et al., "Cancer cell profiling by barcoding allows multiplexed protein analysis in fine needle aspirates," Science Translational Medicine, 6(219):219ra9, 22 pages. doi: 10.1126/scitranslmed.3007361.	
	108.	WERNER, J. H. et al., "Current status of DNA sequencing by single molecule detection, Proc. SPIE 3602, Advances in Fluorescence Sensing Technology IV, pp. 355-366 (1999).	
	109.	ZINCHUK, V. et al., "Quantitative Colocalization Analysis of Multicolor Confocal Immunofluorescence Microscopy Images: Pushing Pixels to Explore Biological Phenomena," The Japan Society of Histochemistry and Cytochemistry, 40(4):101-111 (2007).	
	110.	ZOLLINGER, D. et al., "Abstract 3434: Digital spatial profiling platform allows for spatially resolved, high-plex quantification of mRNA distribution and abundance on FFPE and fresh frozen tissue sections," AACR Annual Meeting, April 14-18, 2018, 2 pages.	

203107908 v1

Examiner Signature:		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached. This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 15/10, 33/50, 15/14

A2

(11) International Publication Number: **WO 97/14028**

(43) International Publication Date: 17 April 1997 (17.04.97)

(21) International Application Number: PCT/US96/16198

(22) International Filing Date: 10 October 1996 (10.10.96)

(30) Priority Data:

08/540,814	11 October 1995 (11.10.95)	US
08/542,401	11 October 1995 (11.10.95)	US

(71) Applicant *(for all designated States except US)*: LUMINEX CORPORATION [US/US]; 12212 Technology Boulevard #K, Austin, TX 78727-6115 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CHANDLER, Van, S. [US/US]; 2808 McKinney Avenue #410, Dallas, TX 75204 (US). FULTON, R., Jerrold [US/US]; 305 Evergreen Trail, Cedar Hill, TX 75104 (US). CHANDLER, Mark, B. [US/US]; 4 Niles Road, Austin, TX 78700 (US).

(74) Agent: PATTERSON, Melinda, L.; P.O. Box 4433, Houston, TX 77210 (US).

(81) Designated States: AL, AM, AT, AU, BA, BB, BG, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: MULTIPLEXED ANALYSIS OF CLINICAL SPECIMENS APPARATUS AND METHOD

(57) Abstract

A method for the multiplexed diagnostic and genetic analysis of enzymes, DNA fragments, antibodies, and other biomolecules comprises the steps of constructing an appropriately labeled beadset, exposing the beadset to a clinical sample, and analyzing the combined sample/beadset by flow cytometry is disclosed. Flow cytometric measurements are used to classify, in real-time, beads within an exposed beadset and textual explanations, based on the accumulated data obtained during real-time analysis, are generated for the user. The inventive technology enables the simultaneous, and automated, detection and interpretation of multiple biomolecules or DNA sequences in real-time while also reducing the cost of performing diagnostic and genetic assays.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia
AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

MULTIPLEXED ANALYSIS OF CLINICAL SPECIMENS APPARATUS AND METHOD

Microfiche appendix A contains a listing of selected Visual Basic and C programming source code in accordance with the inventive multiplexed assay method. Microfiche appendix A, comprising 1 sheet having a total of 58 frames, contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent disclosure, as it appears in the Patent and Trademark Office patent files or records, but otherwise reserves all copyright rights whatsoever.

The invention relates generally to laboratory diagnostic and genetic analysis and, more particularly, to a flow cytometric method for the simultaneous and multiplexed diagnostic and genetic analysis of clinical specimens.

Analysis of clinical specimens is important in science and medicine. A wide variety of assays to determine qualitative and/or quantitative characteristics of a specimen are known in the art. Detection of multiple analytes, or separately identifiable characteristics of one or more analytes, through single-step assay processes are presently not possible or, to the extent possible, have provided only very limited capability and have not yielded satisfactory results. Some of the reasons for these disappointing results include the extended times typically required to enable the detection and classification of multiple analytes, the inherent limitations of known reagents, the low sensitivities achievable in prior art assays which often lead to significant analytical errors and the unwieldy collection, classification, and analysis of prior art algorithms vis à vis the large amounts of data obtained and the subsequent computational requirements to analyze that data.

Clearly, it would be an improvement in the art to have adequate apparatus and methods for reliably performing real-time multiple determinations, substantially simultaneously, through a single or limited step assay process. A capability to perform simultaneous, multiple

determinations in a single assay process is known as "multiplexing" and a process to implement such a capability is a "multiplexed assay."

Flow Cytometry

One well known prior art technique used in assay procedures for which a multiplexed assay capability would be particularly advantageous is flow cytometry. Flow cytometry is an optical technique that analyzes particular particles in a fluid mixture based on the particles' optical characteristics using an instrument known as a flow cytometer. Background information on flow cytometry may be found in Shapiro, "Practical Flow Cytometry," Third Ed. (Alan R. Liss, Inc. 1995); and Melamed et al., "Flow Cytometry and Sorting," Second Ed. (Wiley-Liss 1990), which are incorporated herein by reference.

Flow cytometers hydrodynamically focus a fluid suspension of particles into a thin stream so that the particles flow down the stream in substantially single file and pass through an examination zone. A focused light beam, such as a laser beam illuminates the particles as they flow through the examination zone. Optical detectors within the flow cytometer measure certain characteristics of the light as it interacts with the particles. Commonly used flow cytometers such as the Becton-Dickinson Immunocytometry Systems "FACSCAN" (San Jose, CA) can measure forward light scatter (generally correlated with the refractive index and size of the particle being illuminated), side light scatter (generally correlated with the particle's size), and particle fluorescence at one or more wavelengths. (Fluorescence is typically imparted by incorporating, or attaching a fluorochrome within the particle.) Flow cytometers and various techniques for their use are described in, generally, in "Practical Flow Cytometry" by Howard M. Shapiro (Alan R. Liss, Inc., 1985) and "Flow Cytometry and Sorting, Second Edition" edited by Melamed et al. (Wiley-Liss, 1990).

One skilled in the art will recognize that one type of "particle" analyzed by a flow cytometer may be man-made microspheres or beads. Microspheres or beads for use in flow cytometry are generally known in the art and may be obtained from manufacturers such as Spherotech (Libertyville, IL), and Molecular Probes (Eugene, OR).

Although a multiplexed analysis capability theoretically would provide enormous benefits in the art of flow cytometry, very little multiplexing capability has been previously achieved. Prior multiplexed assays have obtained only a limited number of determinations. A review of some of these prior art techniques is provided by McHugh, "Flow Microsphere Immunoassay for the Quantitative and Simultaneous Detection of Multiple Soluble Analytes," in Methods in Cell Biology, 42, Part B, (Academic Press, 1994). For example, McHugh et al., "Microsphere-Based Fluorescence Immunoassays Using Flow Cytometry Instrumentation," in Clinical Flow Cytometry Ed. K.D. Bauer, et al., Williams and Williams, Baltimore, MD, 1993, 535-544, describe an assay where microspheres of different sizes are used as supports and the identification of microspheres associated with different analytes was based on distinguishing a microsphere's size. Other references in this area include Lindmo, et al., "Immunometric Assay by Flow Cytometry Using Mixtures of Two Particle Types of Different Affinity," J. Immun. Meth., 126, 183-189 (1990); McHugh, "Flow Cytometry and the Application of Microsphere-Based Fluorescence Immunoassays," Immunochemica, 5, 116 (1991); Horan et al., "Fluid Phase Particle Fluorescence Analysis: Rheumatoid Factor Specificity Evaluated by Laser Flow Cytophotometry" in Immunoassays in the Clinical Laboratory, 185-198 (Liss 1979); Wilson et al., "A New Microsphere-Based Immunofluorescence Assay Using Flow Cytometry," J. Immunological Methods, 107, 225-230 (1988); and Fulwyler et al., "Flow Microsphere Immunoassay for the Quantitative and Simultaneous Detection of Multiple Soluble Analytes," Meth. Cell Biol., 33, 613-629 (1990).

The above cited methods have been unsatisfactory as applied to provide a fully multiplexed assay capable of real-time analysis of more than a few different analytes. For example, certain of the assay methods replaced a single ELISA procedure with a flow cytometer-based assay. These methods were based on only a few characteristics of the particles under analysis and enabled simultaneous determination of only a very few analytes in the assay. Also, the analytic determinations made were hampered due to software limitations including the inability to perform real-time processing of the acquired assay data. In summary, although it has been previously hypothesized that flow cytometry may possibly be adapted to operate and

provide benefit in a multiple analyte assay process, such an adaptation has not in reality been accomplished.

Analysis of Genetic Information

The availability of genetic information and association of disease with mutation(s) of critical genes has generated a rich field of clinical analysis. In particular, the use polymerase chain reaction (PCR) and its variants have facilitated genetic analysis. A major advance in this field is described in our co-pending and contemporaneously filed U.S. Application entitled "Methods and Compositions for Flow Cytometric Determination of DNA Sequences." This co-pending application describes a powerful flow cytometric assay for PCR products, which may be multiplexed in accordance with the present invention. A multiplexed flow cytometric assay for PCR reaction products would provide a significant advantage in the field of genetic analysis.

Recent advances in genetic analyses have provided a wealth of information regarding specific mutations occurring in particular genes in given disease states. Consequently, use of an individual's genetic information in diagnosis of disease is becoming increasingly prevalent. Genes responsible for disease have been cloned and characterized in a number of cases, and it has been shown that responsible genetic defects may be a gross gene alteration, a small gene alteration, or even in some cases, a point mutation. There are a number of reported examples of diseases caused by genetic mutations. Testing of gene expression by analysis of cDNA or mRNA, and testing of normal genes and alleles, as in cases of tissue typing and forensics, are becoming widespread. Other uses of DNA analysis, for example in paternity testing, etc., are also important and can be used in accordance with the invention.

Current techniques for genetic analysis have been greatly facilitated by the development and use of polymerase chain reaction (PCR) to amplify selected segments of DNA. The power and sensitivity of the PCR has prompted its application to a wide variety of analytical problems in which detection of DNA or RNA sequences is required.

PCR is capable of amplifying short fragments of DNA, providing short (20 bases or more) nucleotides are supplied as primers. The primers anneal to either end of a span of denatured DNA target and, upon renaturation, enzymes synthesize the intervening complementary sequences by extending the primer along the target strand. During denaturation, the temperature is raised to break apart the target and newly synthesized complementary sequence. Upon cooling, renaturation and annealing, primers bind to the target and the newly made opposite strand and now the primer is extended again creating the complement. The result is that in each cycle of heating and renaturation followed by primer extension, the amount of target sequence is doubled.

One major difficulty with adoption of PCR is the cumbersome nature of the methods of analyzing the reaction's amplified DNA products. Methods for detecting genetic abnormalities and PCR products have been described but they are cumbersome and time consuming. For example, U.S. Patent No. 5,429,923 issued July 4, 1995 to Seidman, et al., describes a method for detecting mutations in persons having, or suspected of having, hypertrophic cardiomyopathy. That method involves amplifying a DNA sequence suspected of containing the disease associated mutation, combining the amplified product with an RNA probe to produce an RNA-DNA hybrid and detecting the mutation by digesting unhybridized portions of the RNA strand by treating the hybridized product with an RNAse to detect mutations, and then measuring the size of the products of the RNAse reaction to determine whether cleavage of the RNA molecule has occurred.

Other methods used for detecting mutations in DNA sequences, including direct sequencing methods (Maxim and Gilbert, Proc. Natl. Acad. Sci. U.S.A., 74, 560-564, 1977); PCR amplification of specific alleles, PASA (Botttema and Sommer, Muta. Res., 288, 93-102, 1993); and reverse dot blot method (Kawasaki, et al., Methods in Enzymology, 218, 369-81, 1993) have been described. These techniques, while useful, are time consuming and cumbersome and for that reason are not readily adaptable to diagnostic assays for use on a large scale.

At least one use of flow cytometry for the assay of a PCR product has been reported but that assay has not been adapted to multiplexing. See Vlieger et al., "Quantitation of Polymerase Chain Reaction Products by Hybridization-Based Assays with Fluorescent Colorimetric, or

Chemiluminescent Detection," Anal. Biochem., 205, 1-7 (1992). In Vlieger et al. a PCR product was labeled using primers that contained biotinylated nucleotides. Unreacted primers were first removed and the amplified portion annealed with a labeled complementary probe in solution. Beaded microspheres of avidin were then attached to the annealed complementary material. The avidin beads bearing the annealed complementary material were then processed by a flow cytometer. The procedure was limited, inter alia, in that avidin beads having only a single specificity were employed. Further, real-time analysis of the assay's data was not possible.

Data Manipulation

The large volume of data typically generated during flow cytometric multiple analyte assays, combined with the limited capabilities of prior techniques to collect, sort and analyze such data have provided significant obstacles in achieving a satisfactory multiplexed assay. The computing methods used in prior art flow cytometric analyses have generally been insufficient and unsuited for accurately and timely analyzing large volumes of data such as would be generated by multiplexed assays; particularly when more than two analytes (or properties of a single analyte) are to be simultaneously determined.

The present invention enables the simultaneous determination of multiple distinct analytes to a far greater degree than existing techniques. Further, the invention provides an improved data classification and analysis methodology that enables the meaningful analysis of highly multiplexed assays in real-time. The invention is broadly applicable to multiplexed analysis of a number of analytes in a host of bioassays in which there is currently a need in the art.

The present invention provides improved methods, instrumentation, and products for detecting multiple analytes in a fluid sample by flow cytometric analysis and for analyzing and presenting the data in real-time. An advantage of the invention is that it allows one rapidly and simultaneously to detect a wide variety of analytes of interest in a single assay step.

The invention employs a pool of bead subsets. The individual subsets are prepared so that beads within a subset are relatively homogeneous but differ in at least one distinguishing characteristic from beads in any other subset. Therefore, the subset to which a bead belongs can readily be determined after beads from different subsets are pooled.

In a preferred embodiment, the beads within each subset are uniform with respect to at least three and preferably four known classification parameter values measured with a flow cytometer: e.g., forward light scatter (C_1) which generally correlates with size and refractive index; side light scatter (C_2) which generally correlates with size; and fluorescent emission in at least one wavelength (C_3), and preferably in two wavelengths (C_3 and C_4), which generally results from the presence of fluorochrome(s) in or on the beads. Because beads from different subsets differ in at least one of the above listed classification parameters, and the classification parameters for each subset are known, a bead's subset identity can be verified during flow cytometric analysis of the pool in a single assay step and in real-time.

Prior to pooling subsets of beads to form a beadset, the beads within each subset can be coupled to a reactant that will specifically react with a given analyte of interest in a fluid sample to be tested. Usually, different subsets will be coupled to different reactants so as to detect different analytes. For example, subset 1 may be labeled so as to detect analyte A (AnA); subset 2 may be labeled so as to detect analyte B (AnB); etc.

At some point prior to assay, the variously labeled subsets are pooled. The pooled beads, or beadset, are then mixed with a fluid sample to test for analytes reactive with the various reactants bound to the beads. The system is designed so that reactions between the reactants on the bead surfaces and the corresponding analytes in the fluid sample will cause changes in the intensity of at least one additional fluorescent signal (F_m) emitted from a fluorochrome that fluoresces at a wavelength distinct from the wavelengths of classification parameters C_3 or C_4. The F_m signal serves as a "measurement signal," that is, it indicates the extent to which the reactant on a given bead has undergone a reaction with its corresponding analyte. The F_m signal may result from the addition to the assay mixture of fluorescently labeled "secondary" reagent that binds to the bead surface at the site where a reactant-analyte reaction has occurred.

When the mixture (pooled beads and fluid sample) is run through a flow cytometer, each bead is individually examined. The classification parameters, e.g., C_1, C_2, C_3, and C_4, are measured and used to classify each bead into the subset to which it belongs and, therefore, identify the analyte that the bead is designed to detect. The F_m value of the bead is determined to indicate the concentration of analyte of interest in the fluid sample. Not only are many beads from each subset rapidly evaluated in a single run, multiple subsets are evaluated in a single run. Thus, in a single-pass and in real-time a sample is evaluated for multiple analytes. Measured F_m values for all beads assayed and classified as belonging to a given subset may be averaged or otherwise manipulated statistically to give a single meaningful data point, displayed in histogram format to provide information about the distribution of F_m values within the subset, or analyzed as a function of time to provide information about the rate of a reaction involving that analyte.

In a preferred embodiment, the beads will have two or more fluorochromes incorporated within or on them so that each of the beads in a given subset will possess at least four different classification parameters, e.g., C_1, C_2, C_3, and C_4. For example, the beads may be made to contain a red fluorochrome (C_3), such as nile red, and bear an orange fluorochrome (C_4), such as Cy3 or phycoerythrin. A third fluorochrome, such as fluorescein, may be used as a source of the C_n or F_m signal. As those of skill in the art will recognize, additional fluorochromes may be used to generate additional C_n signals. That is, given suitable fluorochromes and equipment, those of skill in the art may use multiple fluorochromes to measure a variety of C_n or F_m values, thus expanding the multiplexing power of the system even further.

In certain applications designed for more quantitative analysis of analyte concentrations or for kinetic studies, multiple subsets of beads may be coupled to the same reactant but at varying concentrations so as to produce subsets of beads varying in density of bound reactant rather than in the type of reactant. In such an embodiment, the reactant associated with classification parameter C_4, for example, may be incorporated directly into the reactive reagent

that is coupled to the beads, thereby allowing C_4 conveniently to serve as an indicator of density of reactant on the bead surface as well as an indicator of reactant identity.

To prepare subsets varying in reactant density one may, for example, select, isolate, or prepare a starting panel of different subsets of beads, each subset differing from the other subsets in one or more of C_1, C_2, or C_3. Each of those subsets may be further subdivided into a number of aliquots. Beads in each aliquot may be coupled with a reactant of choice that has been fluorescently labeled with a fluorochrome associated with C_4 (e.g., Analyte A labeled with Cy3) under conditions such that the concentration or density of reactant bound to the beads of each aliquot will differ from that of each other aliquot in the subset. Alternatively, an entire subset may be treated with the C_4 fluorochrome under conditions that produce a heterogeneous distribution of C_4 reactant on beads within the subset. The subset may then be sorted with a cell sorter on the basis of the intensity of C_4 to yield further subsets that differ from one another in C_4 intensity.

One limitation of the alternative embodiment of using C_4 labeled reactant as a classification agent is that one must design the system so that the value of C_4 as a classification parameter is not lost. Therefore, one must take care to assure that the C_4 intensities of all subsets carrying reagent A differs from the C_4 intensities of all subsets carrying reagents B, C, and so forth. Otherwise, C_4 would not be useful as a parameter to discriminate reactant A from reactant B, etc.

With either embodiment, the number of subsets that can be prepared and used in practice of the invention is theoretically quite high, but in practice will depend, *inter alia*, on the level of homogeneity within a subset and the precision of the measurements that are obtained with a flow cytometer. The intra-subset heterogeneity for a given parameter, e.g., forward angle light scatter C_1, correlates inversely with the number of different subsets for that parameter that can be discriminated by flow cytometric assay. It is therefore desirable to prepare subsets so that the coefficients of variation for the value of each classification parameter (C_1, C_2, C_3, and C_4) to be used in a given analysis is minimized. Doing this will maximize the number of subsets that can

be discriminated by the flow cytometer. Bead subsets may be subjected to flow cytometric sorting or other procedures at various different points in preparation or maintenance of the bead subsets to increase homogeneity within the subset. Of course, with simple assays designed to detect only a few different analytes, more heterogeneity can be allowed within a subset without compromising the reliability of the assay.

In an illustrative embodiment set forth here to explain one manner in which the invention can work in practice, the beads are used to test for a variety of antibodies in a fluid sample. A panel of bead subsets having known varying C_1, C_2, C_3, and C_4 values is first prepared or otherwise obtained. The beads within each subset are then coupled to a given antigen of interest. Each subset receives a different antigen. The subsets are then pooled to form an assay beadset and may be stored for later use and/or sold as a commercial test kit.

In the assay procedure, the beads are mixed with the fluid to be analyzed for antibodies reactive with the variety of antigens carried on the beads under conditions that will permit antigen-antibody interaction. The beads are labeled with a "secondary" reagent that binds to antibodies bound to the antigens on the beads and that also bears the measurement fluorochrome associated with parameter F_m (e.g., fluorescein). A fluoresceinated antibody specific for immunoglobulin may be used for this purpose. The beads are then run through a flow cytometer, and each bead is classified by its characteristic classification parameters as belonging to subset-1, subset-2, etc. At the same time, the presence of antibodies specific for antigen A, B, etc., can be detected by measuring green fluorescence, F_m, of each bead. The classification parameters C_1, C_2, C_3, and C_4 allow one to determine the subset to which a bead belongs, which serves as an identifier for the antigen carried on the bead. The F_m value of the bead indicates the extent to which the antibody reactive with that antigen is present in the sample.

Although assays for antibodies were used above as an illustration, those of ordinary skill in the art will recognize that the invention is not so limited in scope, but is widely applicable to detecting any of a number of analytes in a sample of interest. For example, the methods described here may be used to detect enzymes or DNA or virtually any analyte detectable by virtue of a given physical or chemical reaction. A number of suitable assay procedures for

detection and quantification of enzymes and DNA (particularly as the result of a PCR process) are described in more detail below.

The present invention also provides a significant advance in the art by providing a rapid and sensitive flow cytometric assay for analysis of genetic sequences that is widely applicable to detection of RNA, differing alleles, and any of a number of genetic abnormalities. In general, the methods of the present invention employ a competitive hybridization assay using DNA coupled microspheres and fluorescent DNA probes. Probes and microsphere-linked oligonucleotides could also include RNA, PNA, and non-natural nucleotide analogs.

In practice of the invention, oligonucleotides from a region of a gene of interest, often a polymorphic allele or a region to which a disease associated mutation has been mapped, are synthesized and coupled to a microsphere (bead) by standard techniques such as by carbodiimide coupling. A fluorescent oligonucleotide, complementary to the oligonucleotide on the bead, is also synthesized. To perform a test in accordance with the invention, DNA which is to be tested is purified and either assayed unamplified, or subjected to amplification by PCR, RT-PCR, or LCR amplification using standard techniques and PCR initiation probes directed to amplify the particular region of DNA of interest. The PCR product is then incubated with the beads under conditions sufficient to allow hybridization between the amplified DNA and the oligonucleotides present on the beads. A fluorescent DNA probe that is complementary to the oligonucleotide coupled to the beads is also added under competitive hybridization conditions. Aliquots of the beads so reacted are then run through a flow cytometer and the intensity of fluorescence on each bead is measured to detect the level of fluorescence which indicates the presence or absence of given sequences in the samples.

For example, when beads labeled with an oligonucleotide probe corresponding to a non-mutated (wild-type) DNA segment are hybridized with the PCR product from an individual who has a non-mutated wild-type DNA sequence in the genetic region of interest, the PCR product will effect a significant competitive displacement of fluorescent oligonucleotide probe from the beads and, therefore, cause a measurable decrease in fluorescence of the beads, e.g., as compared to a

control reaction that did not receive PCR reaction product. If, on the other hand, a PCR product from an individual having a mutation in the region of interest is incubated with the beads bearing the wild-type probe, a significantly lesser degree of displacement and resulting decrease in intensity of fluorescence on the beads will be observed because the mutated PCR product will be a less effective competitor for binding to the oligonucleotide coupled to the bead than the perfectly complementary fluorescent wild-type probe. Alternatively, the beads may be coupled to an oligonucleotide corresponding to a mutation known to be associated with a particular disease and similar principles applied. In the multiplexed analysis of nucleic acid sequences, bead subsets are prepared with all known, or possible, variants of the sequence of interest and then mixed to form a bead set. The reactivity of the test sample, e.g. PCR product, with the wild-type sequence and other variants can then be assayed simultaneously. The relative reactivity of the PCR product with subsets bearing the wild-type or variant sequences identifies the sequence of the PCR product. The matrix of information derived from this type of competitive hybridization in which the test sequence and the entire panel of probe sequences react simultaneously allows identification of the PCR product as wild-type, known mutant, or unknown mutant. The invention thus provides one with the ability to measure any of a number of genetic variations including point mutations, insertions, deletions, inversions, and alleles in a simple, exquisitely sensitive, and efficient format.

Figure 1 is a block diagram of an illustrative hardware system for performing a multiplex assay method in accordance with the invention.

Figure 2 is a block diagram of an illustrative software system for performing a multiplex assay method in accordance with the invention.

Figure 3 is a flow-chart for a preprocessing phase in accordance with the inventive multiplexed assay method.

Figure 4 shows an assay database in accordance with the invention.

Figure 5 shows a baseline data acquisition table for an illustrative multiple analyte assay in accordance with the invention.

Figure 6 shows an assay definition table in accordance with the invention.

Figure 7 shows a discriminant table for an illustrative multiple analyte assay in accordance with the invention.

Figure 8 shows a decision tree view of the illustrative discriminant function table of Figure 7.

Figure 9 is a flow-chart for a real-time analysis phase of a multiple analyte assay in accordance with the invention.

Figure 10 shows a results table for an illustrative multiple analyte assay in accordance with the invention.

Figure 11 shows a interpretation table for an illustrative multiple analyte assay in accordance with the invention.

Figure 12 is a flow-chart for an interpretation phase of a multiple analyte assay in accordance with the invention

Figures 13a through 13e show an assay database in accordance with the invention for a specific experimental example.

Figure 14 shows a decision tree view for an illustrative (experimental example) discriminant table.

Figures 15a, 15b, and 15c show individual inhibition assays for IgG, IgA, and IgM antibodies.

Figures 16a, 16b, and 16c show cross reactivity determinations between IgG, IgA, and IgM assay components.

Figure 17 shows the determination of human IgG concentrations by flow cytometry.

Figure 18 shows the determination of human IgA concentrations by flow cytometry.

Figure 19 shows the determination of human IgM concentrations by flow cytometry.

Figure 20 shows the simultaneous determination of human IgG, IgA, and IgM concentrations by flow cytometry.

Figure 21 shows the specificity of monoclonal antibody MAB384 binding towards bead immobilized epitope sequences.

Figure 22 shows the specificity of monoclonal antibody MAB384 binding in the presence of soluble epitope containing peptide.

Figure 23 shows the specificity of monoclonal antibody MAB384 binding in the presence of soluble biotin.

Figure 24 shows the detection of anti-Rubella IgG antibodies by a sandwich assay between rubella coated beads and a fluorescent goat anti-human IgG antibody.

Figure 25 shows a calibration assay using serial dilutions of anti-Rubella IgG antibodies in a sandwich assay between rubella coated beads and a fluorescent goat anti-human IgG antibody.

Figures 26a and 26b show the simultaneous assay for six anti-ToRCH IgG, and simultaneous assay for the six anti-ToRCH IgM antibodies.

Figure 27 shows the determination of IgG anti-grass allergen activities for six dogs.

Figure 28 shows the determination of IgE anti-grass allergen activities for six dogs.

Figure 29 shows the multiple analyte IgG and IgE screening of dog serum A96324 for sixteen grass allergens

Figure 30 shows the multiple analyte IgG and IgE screening of dog serum A96325 for sixteen grass allergens

Figure 31 shows the multiple analyte IgG and IgE screening of dog serum A96319 for sixteen grass allergens

Figure 32 shows the multiple analyte IgG and IgE screening of dog serum A96317 for sixteen grass allergens

Figure 33 shows the multiple analyte IgG and IgE screening of dog serum A96326 for sixteen grass allergens

Figure 34 shows the multiple analyte IgG and IgE screening of dog serum A96323 for sixteen grass allergens

Figure 35 shows an antibody pair analysis for use with a human chorionic gonadotropin capture assay.

Figure 36 shows the use of bead linked antibody MAB602 with fluorescently labeled antibody AB633 in a human chorionic gonadotropin capture assay.

Figure 37a and 37b show cross reactivity analyses between components of an anti-hCG capture system and an anti-AFP capture system.

Figures 38a and 38b compare the effects of eliminating wash steps in hCG and AFP capture system assays.

Figures 39a and 39b show the determination of hCG and AFP concentrations in samples and standards using a homogeneous capture assay format.

Figure 40 shows the inhibition of Anti-IgG binding to bead based IgG by soluble IgG antibodies. Inhibition was determined at five concentrations of soluble IgG, and four IgG loading levels on the beads.

Figure 41 shows the slope of the inhibition pattern across the four loading levels of IgG on the beads plotted against the soluble IgG concentration.

Figure 42 shows a five point standard curve derived from inhibition data of the 50 µg/mL IgG bead set.

Figures 43a through 43c show DNA detection using a double stranded competitor and a wild-type "B" oligonucleotideprobe.

Figures 44a and 44b show DNA detection using a single stranded competitor and a wild-type "B" oligonucleotideprobe.

Figure 45 shows the differentiation by orange and red fluorescence of fourteen bead sets.

Figure 46 shows a titration of a fluorescent oligonucleotide in the presence or absence of an inhibitor. Beads bearing complementary oligonucleotides were used in a capture assay.

Figure 47 shows the inhibition of binding between a fluorescent oligonucleotide and its complementary oligonucleotide bound to a bead. Varying concentrations of complementary and point mutant competitors were used in the determination.

Figure 48 shows the efficacy of inhibitors across fourteen DNA sequence binding sets.

Figure 49 shows the typing of four simulated alleles of DQA1.

Figure 50 shows the typing of five known, homozygous DQA1 alleles.

Figures 51a through 51f show the results of an exemplary multiplexed assay according to the invention.

According to the present invention, assay components and methods for the measurement of enzymes, DNA fragments, antibodies, and other biomolecules are provided. The inventive technology improves the speed and sensitivity of flow cytometric analysis while reducing the cost of performing diagnostic and genetic assays. Further, and of tremendous significance, a multiplexed assay in accordance with the invention enables the simultaneous automated assay of multiple (at least an order of magnitude greater than available in the prior techniques) biomolecules or DNA sequences in real-time.

As those of ordinary skill in the art will recognize, the invention has an enormous number of applications in diagnostic assay techniques. Beadsets may be prepared, for example, so as to detect or screen for any of a number of sample characteristics, pathological conditions, or reactants in fluids. Beadsets may be designed, for example, to detect antigens or antibodies associated with any of a number of infectious agents including (without limitation, bacteria, viruses, fungi, mycoplasma, rickettsia, chlamydia, and protozoa), to assay for autoantibodies associated with autoimmune disease, to assay for agents of sexually transmitted disease, or to

assay for analytes associated with pulmonary disorders, gastrointestinal disorders, cardiovascular disorders, and the like. Similarly, the beadset may be designed to detect any of a number of substances of abuse, environmental substances, or substances of veterinary importance. An advantage of the invention is that it allows one to assemble a panel of tests that may be run on an individual suspected of having a syndrome to simultaneously detect a causative agent for the syndrome.

Suitable panels may include, for example, a tumor marker panel including antigens such as prostate-specific antigen (PSA), carcinoembryonic antigen (CEA), and other suitable tumor markers; a regional allergy panel including pollen and allergens tested for by allergists of a particular region and comprising allergens known to occur in that region; a pregnancy panel comprising tests for human chorionic gonadotropin, hepatitis B surface antigen, rubella virus, alpha fetoprotein, 3′ estradiol, and other substances of interest in a pregnant individual; a hormone panel comprising tests for T4, TSH, and other hormones of interests; an autoimmune disease panel comprising tests for rheumatoid factors and antinuclear antibodies and other markers associated with autoimmune disease; a blood borne virus panel and a therapeutic drug panel comprising tests for Cyclosporin, Digoxin, and other therapeutic drugs of interest.

Bead Technology

An important feature of the flow cytometric technology and techniques described here is the fabrication and use of particles (e.g., microspheres or beads that make up a beadset). It is through the use of appropriately labeled homogeneous bead subsets, combined to produce a pooled beadset, that the instant multiplexed assay method is practiced. Beads suitable for use as a starting material in accordance with the invention are generally known in the art and may be obtained from manufacturers such as Spherotech (Libertyville, IL) and Molecular Probes (Eugene, OR). Once a homogeneous subset of beads is obtained, the beads are labeled with an appropriate reactant such as a biomolecule, DNA sequence, and/or other reactant. Known methods to incorporate such labels include polymerization, dissolving, and attachment.

A Method for the Multiplexed Assay of Clinical Samples

Development of a multiplexed assay for use in accordance with the invention can be divided into three phases: (1) preprocessing, (2) real-time analysis, and (3) interpretation. During the preprocessing phase, baseline data is collected independently, via flow cytometric techniques, for each of an assay's bead subsets. Baseline data is used to generate a set of functions that can classify any individual bead as belonging to one of the assay's subsets or to a rejection class. During the analysis phase, flow cytometric measurements are used to classify, in real-time, each bead within an exposed beadset according to the aforementioned functions. Additionally, measurements relating to each subset's analyte are accumulated. During the interpretation phase the assay's real-time numerical results are associated with textual explanations and these textual explanations are displayed to a user.

The inventive method allows the detection of a plurality of analytes simultaneously during a single flow cytometric processing step. Benefits of the inventive multiplex assay method include increased speed and reduced cost to analyze a clinical sample.

System Hardware

Figure 1 shows, in block diagram form, a system for implementing the inventive multiplexed assay method. Flow cytometer **100** output consists of a series of electrical signals indicative of one or more specified measured characteristics on each bead processed. These measurement signals are transmitted to computer **105** via data bus **110** and interface board **115**. During the preprocessing phase, the signals are used by the computer to generate an assay database. During the real-time analysis phase, the signals are processed by the computer (using the assay database) in accordance with the inventive method to produce a multiplexed/simultaneous assay of a clinical sample.

Flow cytometer **100** operates in a conventional manner. That is, beads are processed by illuminating them, essentially one at a time, with a laser beam. Measurements of the scattered laser light are obtained for each illuminated bead by a plurality of optical detectors. In addition, if a bead contains at least one appropriate fluorescing compound it will fluoresce when illuminated. A plurality of optical detectors within the flow cytometer measure fluorescence at a

plurality of wavelengths. Typical measured bead characteristics include, but are not limited to, forward light scatter, side light scatter, red fluorescence, green fluorescence, and orange fluorescence. One of ordinary skill in the use of flow cytometric techniques will recognize that the use of green fluorescent markers or labels can cause cross-channel interference between optical detectors designed to detect green and orange wavelengths (e.g., approximately 530 nanometers and approximately 585 nanometers respectively). A training set of beads, in combination with standard data manipulation, can correct for this cross-channel interference by providing the physical measurements required for mathematical correction of the fluorescence measurements.

One of ordinary skill will further recognize that many alternative flow cytometer setups are possible. For instance, additional color sensitive detectors could be used to measure the presence of other fluorescence wavelengths. Further, two or more laser beams can be used in combination to illuminate beads as they flow through the cytometer to allow excitation of fluorochromes at different wavelengths.

Computer **105** can be a conventional computer such as a personal computer or engineering workstation. In one embodiment, the computer is a personal computer having an Intel "486" processor, running Microsoft Corporation's "WINDOWS" operating system, and a number of ISA expansion slots.

Interface board **115** is designed to plug into one of the computer's **100** ISA (Industry Standard Architecture) expansion slots. While the design of an interface board is, in general, different for each specific type of flow cytometer **100**, its primary functions include (1) receiving and parsing measurement data signals generated by the flow cytometer's detectors, (2) receiving control parameter status information from the flow cytometer, and (3) sending control parameter commands to the flow cytometer. The precise manner in which these functions are carried out are dependent upon the type (make and model) of the flow cytometer used. In one embodiment, employing a Becton-Dickinson "FACSCAN" flow cytometer (San Jose, CA), the interface board uses control signals generated by the flow cytometer to distinguish measurement

data and flow cytometer parameter and control signals. Measured data include forward light scatter, side light scatter, red fluorescence, green fluorescence, and orange fluorescence. Parameter and control signals include flow cytometer amplifier gain adjustments and status information.

While the design of an interface board **115** for use with the inventive assay method would be a routine task for one skilled in the art of diagnostic medical equipment design having the benefit of this disclosure, an important aspect for any interface board is its ability to accommodate the transmission data rate generated by whatever flow cytometer is used. For example, the "FACSCAN" flow cytometer can transmit a 16-bit (2 byte) word every 4 microseconds resulting in burst data rates of 500,000 bytes per second. Microfiche appendix A provides a detailed source code embodiment of the inventive assay method for use with the "FACSCAN" flow cytometer.

Data bus **115** provides a physical communication link between the flow cytometer **100** and the interface board **110**. Its physical and electrical characteristics (e.g., data width and bandwidth) are dependent upon the capabilities of the flow cytometer. It is noted that the data bus need not be a totally digital bus. If the flow cytometer does not include analog-to-digital conversion of measured bead characteristics (e.g., light scatter and fluorescence signals), then the data bus must communicate these analog signals to the interface board. It is then necessary that digital conversion of these signals be provided by either the interface board or another peripheral device before the data is transmitted to the computer **105**.

System Software

As shown in **Figure 2**, the software architecture for the inventive assay method can be divided into two parts. A graphical user interface (GUI) **200** provides the means by which a user (1) receives assay results and (2) interacts with the flow cytometer. A dynamically linked library (DLL) **205** provides the means through which the inventive real-time assay is performed and includes routines necessary to (1) interact with interface board **115** and (2) send and receive information to the flow cytometer **100**.

An important aspect of the inventive assay method is that it performs a simultaneous analysis for multiple analytes in real-time. One of ordinary skill in the art of computer software development will realize that real-time processing can impose severe time constraints on the operational program code, i.e., the DLL **205**. For example, the "FACSCAN" flow cytometer can process, or measure, approximately 2,000 beads per second, where each bead is associated with eight 16-bit data values. Thus, to process flow cytometer data in real-time from a "FACSCAN," the DLL should be able to accept, and process, at a consistent data rate of at least 32,000 bytes per second. The need to accommodate this data rate, while also having sufficient time to perform real-time analysis based on the data, will generally necessitate that some of the DLL code be written in assembly language.

In a current embodiment, the GUI **200** is implemented in the visual basic programming language and the DLL **205** is implemented in C and assembly language programming. Microfiche appendix A contains source code listings for one embodiment of the GUI and DLL.

Preprocessing

A function of the preprocessing phase is to generate an assay database for use during the real-time analysis of an exposed beadset (clinical sample). Thus, preprocessing is performed prior to combining separately labeled bead subsets to form assay beadsets. Assay definition, discriminant function definition, and interpretation tables are created at the time an assay beadset is created. **Figure 3** shows, in flow chart form, the steps taken during the preprocessing phase.

A bead subset is characterized by (1) the analyte it is designed to identify, (2) one or more classification parameters $C_1 \ldots C_n$, and (3) one or more measurement parameters F_{m1} – F_{mx}. During the preprocessing phase the classification parameters are used to generate a set of functions, referred to as discriminant functions, that can classify a bead as belonging to one of the assay's subsets or a rejection class. Measurement parameters are used during the real-time

analysis phase to determine if a specified analyte is present in the clinical sample being analyzed.

The precise number of individual beads contained in any given subset is relatively unimportant, the only significant criterion being that a sufficient number are used so that a good statistical characterization of the subset's parameters can be achieved during the real-time analysis phase. In a current embodiment, each bead subset contains an equal number of beads. One of ordinary skill in the field will recognize that the precise number of beads within any given bead subset can vary depending upon many factors including, but not limited to, the number of analytes an assay beadset is designed to detect, the uniformity of the labeled beads (with respect to each of the measured parameters $C_1 \ldots C_n$, $F_{m1} \ldots F_{mx}$), and the penalty of misclassifying (e.g., making a type 1 or type 2 classification error) a bead during analysis.

During preprocessing, each bead in an unexposed subset is measured by a flow cytometer **100** and the resulting data values accumulated for later use **300**. For example, if the flow cytometer measures n classification parameters and x measurement parameters, i.e., generates $(n + x)$ values for each bead, data for each of the subset's $(n + x)$ parameters are updated based on each bead's measurements. This data collection step is repeated independently for each subset in the assay's beadset **305**. The collection of such data for each of an assay's subsets constitutes an assay's baseline data.

After an assay's baseline data has been collected, a set of discriminant functions are determined **310**. During real-time analysis, the discriminant functions are used to classify a bead into one of the assay's bead subsets or a rejection class based solely on the measured classification parameters, $C_1 \ldots C_n$. This step, in principle and practice, is a problem of multi-dimensional classification or cluster analysis. Many prior art techniques and commercial software programs exist to perform this task.

Beads are generally manufactured in large quantities referred to as batches. Each bead in a batch is of nearly identical size and has substantially the same dye absorption capacity. In

light of this manufacturing process, bead subsets can be created using precise dilutions of chosen dyes and, because of their nearly identical size, all classification parameters will exhibit essentially equal variances. By correcting for scaling of the photo-multipliers within a flow cytometer, a linear classification rule can be generated. Further, since there are equal quantities of beads in each subset, the prior probabilities will be equal. This allows use of Fisher's linear discriminant technique to calculate the discriminant functions which define classification boundaries. See, Fisher, "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics, 7, 179-188 (1936). For instance, linear hierarchical discriminant functions may be chosen which are equidistant, in a Euclidean sense, between the centers or centroids of any two of an assay's bead subsets. Notwithstanding the present example, other types of discriminant functions, such as quadratic functions and those discriminating on more than two classification parameters at once, are also possible.

In addition to the discriminant functions, a set of threshold values are chosen which are used during the real-time analysis phase to detect the presence of a target analyte. For example, assume measurement parameter F_{ml} is used to detect analyte-A. During preprocessing, the baseline or unexposed value for F_{ml} is measured and accumulated for that subset's beads. Analyte-A's threshold could then, for example, be set to F_{ml}'s baseline mean value plus one standard deviation of F_{ml}'s baseline value. One of ordinary skill will recognize that the precise function or value selected for a threshold depends upon the parameter being measured (e.g., its distribution) and the cost of making a classification error (e.g., a type 1 or a type 2 error). It is routine that such values be based on an empirical review of the baseline data. The important criterion is that the threshold reliably distinguish between the presence and absence of the target analyte in an exposed assay beadset.

After baseline data for each of an assay's bead subsets are collected and discriminant functions and analyte threshold values are established, an assay database is generated **315**.

Assay Database

As shown in **Figure 4**, an assay database **400** consists of an assay definition table **405**, a discriminant function table **410**, a results table **415**, and an interpretation table **420**. See **Figure 4**.

The assay definition table **405** defines an assay which, as described above, comprises two or more bead subsets each of which is designed to detect a specified analyte. Each row in the assay definition table describes a bead subset and contains the following entries: (1) assay name, (2) subset name, (3) subset token, (4) baseline values for each of the subset's measurement parameters $F_{m1} - F_{mx}$, and (5) test-type token. The subset name entry is a text string identifying the subset by, for example, the type of analyte it is labeled to detect. The subset token is a unique subset identifier. The measurement parameter baseline entries are used during the interpretation phase to associate a numerical result (collected during the real-time analysis of a clinical sample) with a textual output string. Finally, the test-type token identifies which one of a possible plurality of interpretation tests to perform on the collected (real-time) data during the interpretation phase.

The discriminant function table **410** is used to systematically set forth an assay's set of discriminant functions. Each row in the discriminant function table implements a single discriminant function and includes entries for (1) the assay's name, (2) a unique row identifier, (3) one or more classification parameters upon which to evaluate, (4) high and low discriminant values for each of the listed classification parameters, and (5) evaluation tokens which are assigned as a result of evaluating the discriminant function.

The results table **415** is used to store, or accumulate, data on an assay's beadset during the real-time analysis phase of the inventive method and is discussed further in Section 6.2(d).

The interpretation table **420** provides a means to associate text messages with each enumerated assay result and is discussed further in Section 6.2(e).

Preprocessing Example

Consider an assay beadset designed to simultaneously detect four analytes: analyte-A, analyte-B, analyte-C, and analyte-D. Thus, the assay's beadset is comprised of four bead subsets, each labeled for a different analyte. Suppose further that the assay beadset is to be processed by a Becton-Dickinson Immunocytometry Systems "FACSCAN" flow cytometer. For each bead processed, the "FACSCAN" measures forward light scatter, side light scatter, red fluorescence, orange fluorescence, and green fluorescence. Let classification parameter C_1 be forward light scatter, classification parameter C_2 be side light scatter, classification parameter C_3 be red fluorescence, classification parameter C_4 be orange fluorescence, and measurement parameter F_{m1} be green fluorescence. (This notation implies that each bead in a subset is labeled with a green fluorophore bearing, for example, an antibody or dye specifically targeted to that subset's analyte.)

After preparing each of the four subsets and before they are combined to form the assay beadset, they are processed by the flow cytometer and their measured data are accumulated: values for each of the parameters C_1, C_2, C_3, C_4, and F_{m1} are recorded for each bead. Each bead subset is similarly processed. Completion of this task constitutes completion of baseline data acquisition.

Using baseline data, the assay's beads are clustered in the four-dimensional parameter space defined by C_1, C_2, C_3, and C_4. The result of this cluster analysis is that each subset is characterized by a mean (μ) and standard deviation (σ) for each of its four classification parameters. See **Figure 5**. As previously noted, the precise number of individual beads contained in any given bead subset can be calculated by those of ordinary skill in the art. This calculation is required to obtain good statistical characterization of the subset's parameters – e.g., small, or relatively fixed, coefficient of variations for each parameter.

As shown in **Figure 6**, the assay definition table **405** is comprised of general information relevant to the overall diagnostic function of the assay. For instance, in a genotyping assay, each of the assay's subset's may be assigned a token used for identification: e.g., token 46 represents

the bead subset labeled to detect a wildtype coding sequence for a specified gene; subset tokens 21, 50, and 5 represent subsets labeled to detect various mutant type coding sequences for a specified gene(s). Additionally, measurement parameter F_{m1}'s baseline (in this example the mean) and standard deviation values are listed. Finally, a test-type token is listed. In the current embodiment a test-type token of '0' means an OVER/UNDER interpretation test is to be performed and a test-type token of '1' means a SHIFT interpretation test is to be performed. See Section 6.2(f) for further discussion of these issues.

Discriminate functions are generated by viewing the assay's baseline data graphically in three dimensions and creating planes to separate the different subset clusters. These "planes" are created by applying Fischer's Linear Discriminant to the *n*-dimensional classification parameter space. A populated discriminate function table based on the baseline data of **Figure 5** is shown in **Figure 7**.

The discriminant function table provides a systematic means of evaluating a series of classification values (C_1, C_2, C_3, C_4) in order to classify a bead. In general bead classification proceeds by entering the discriminant function table at row 0, performing a test on a specified parameter (e.g., C_1, C_2, C_3, or C_4) and then, depending upon the result, either classifying the bead or proceeding to another test which involves evaluating a different row in the table. For example, suppose bead A has the following measured classification parameter values: $C_1 = V_1$, $C_2 = V_2$, $C_3 = V_3$, and $C_4 = V_4$. Classification of bead A via the discriminant function table of **Figure 7** begins as follows (the pseudo-code below would demonstrate to those skilled in the art of programming the logic involved in the classification process):

1. Enter table at row 0 with measured values for C_1, C_2, C_3, and C_4.
2. If (LOW VALUE = 500) ≤ (PARAMETER = C_1 = V_1) ≤ (HIGH VALUE = 620) then (result = TRUE), else (result = FALSE).
3. If (result = TRUE) and (TRUE ROW ID ≠ 0), then re-enter table at TRUE ROW ID, else
4. If (result = TRUE) and (TRUE ROW ID = 0), then (class = TRUE TOKEN).

5. If (result = FALSE) and (FALSE ROW ID ≠ 0), then re-enter table at (row = FALSE ROW ID), else

6. If (result = FALSE) and (FALSE ROW ID = 0), then (class = FALSE TOKEN).

7. If (TRUE TOKEN or FALSE TOKEN) = 0, then (class = reject class).

One of ordinary skill will recognize from the above discussion that a discriminant function table embodies a (classification) decision tree. **Figure 8** shows this relationship for the discriminant function table of **Figure** 7 explicitly. A discussion of the discriminant function table as it relates to the real-time processing of an exposed assay beadset is provided in Section 6.2(d). Once a beadset is preprocessed, the data may be employed in real-time analysis of many assays using that set. One of ordinary skill in the art will also recognize that instead of a decision tree, a bitmap or look up table could be used to classify the bead sets.

Real-Time Analysis

Once a collection of bead subsets have been characterized as described above and combined to form an assay beadset, the beadset may be exposed to a test sample. That is, they may be used to analyze a clinical sample. After exposure the beadset is ready for real-time analysis. The real-time analysis phase is initiated by installing the exposed beads into a conventional flow cytometer for processing.

As described above, for each bead processed a flow cytometer **100** generates electrical signals indicative of a plurality of measured parameters, $C_1 \ldots C_n$, $F_{m1} \ldots F_{mx}$. These values are transmitted to computer **105** via data bus **110** and interface board **115**. Values for a bead's classification parameters $C_1 \ldots C_n$ are used to evaluate the assay's discriminant functions, as encoded in a discriminant function table **410**, the result of which is an initial classification of the bead into one of the assay's bead subsets or a reject class.

After this initial classification, a bead's measured classification parameter values $C_1 \ldots C_n$ can be checked against their ($C_1 \ldots C_n$) baseline values to determine if it is "reasonable" to classify the bead as belonging to the initially identified class. In a current embodiment, this

reasonableness test is implemented by computing the distance between the measured classification parameter values and the mean values obtained during preprocessing. If the measured values for $C_1 \ldots C_n$ for a particular bead are sufficiently distant from the identified subsets baseline values, the bead is assigned to a reject class. Use of this technique allows for the rejection of beads that were initially misclassified and improves the overall reliability of the analysis.

To ensure proper classification, a preferred embodiment's pooled beadset will include a bead subset which has no bound reactants (e.g., a placebo bead subset) in a known ratio to the beadset's other subsets.

It is noted that when a beadset is comprised of beads manufactured in a single batch, the above described reasonableness test can be incorporated into the linear discriminant functions by creating reject space between all subsets. However, when a beadset is comprised of beads from more than one batch a Euclidean (or similar) distance measure is needed to validate the classification result.

Once a bead is assigned its final classification, the assay's results table **415** is updated to reflect the newly classified bead's measurement parameter values $F_{m1} \ldots F_{mx}$. This data acquisition, classification, and update process is repeated for each bead in the assay beadset in real-time. **Figure 9** shows, in block diagram form, the general steps performed during the real-time analysis phase of a method in accordance with the invention.

In one embodiment the following data are accumulated in the results table for each class (subset) of bead in the assay: (1) total count of the number of beads detected in the specified class, (2) a running sum for each measurement parameter $F_{m1} - F_{mx}$, (3) for each measurement parameter the total count of the number of beads in the class whose measurement value is less than the parameter's baseline value, and (4) for each measurement parameter the total count of the number of beads in the class whose measurement value is more than the parameter's baseline value.

Real-Time Analysis Example

In the illustrative embodiment introduced in Section 6.2(c), the assay beadset is designed to simultaneously detect four analytes using four classification parameters (C_1 represents forward light scatter, C_2 represents side light scatter, C_3 represents red fluorescence, and C_4 represents orange fluorescence) and one measurement parameter (F_{m1} representing green fluorescence). After exposing the beadset to a suitable biological sample, it is placed into a flow cytometer **100** which processes each bead (e.g., measures parameters C_1, C_2, C_3, C_4, and F_{m1}) and transmits to computer **105** signals indicative of these measurements via data bus **110** and interface board **115**.

For each bead processed by the flow cytometer, values for C_1, C_2, C_3, and C_4 are evaluated in accordance with the discriminant function table shown in **Figure 7** to initially classify the bead as belonging to a particular subset, for example, in a genetic analysis intended to detect mutations in the Kras oncogene, the classification could proceed as follows: (1) class 46, Kras CODON 46 WILDTYPE, (2) class 21, Kras CODON 21 MUTANT, (3) class 50, Kras CODON 50 MUTANT, (4) class 5, Kras CODON 5 MUTANT, or (5) a reject class. (See **Figure 8** for a decision tree representation of the discriminate function table of **Figure 7**.) If the bead is initially classified as belonging to any class except the reject class, a reasonableness test is performed on the bead's classification parameter values, C_1 - C_n. For example, if the bead received an initial classification of class 50 and its C_1 value is more than two standard deviations away from its mean, the bead is given a final classification of reject. Otherwise the bead's final classification is the same as its initial classification – 50.

If the bead's final classification is other than reject, its F_{m1} value is used to update the assay's results table in the following manner (see **Figure 10**):

1. Identifying, based on the bead's classification token (i.e., subset token 46, 21, 50, or 5), the row in the results table which is to be updated.
2. Incrementing the identified row's COUNT value. The COUNT value reflects the total number of beads of the specified class that have been identified during the analysis.

3. Adding the bead's F_{mI} value to the value contained in the row's SUM column. The SUM value reflects a running sum of the identified classes measurement values.
4. If the bead's F_{mI} value is greater than F_{mI}'s base value (determined during the preprocessing phase, see **Figure 6**), then incrementing the row's OVER COUNT value. The OVER COUNT value reflects the total number of beads of the specified class that have been processed whose F_{mI} values are above that of baseline.
5. If the bead's F_{mI} value is less than F_{mI}'s base value (as determined during the preprocessing phase, see **Figure 6**), then incrementing the row's UNDER COUNT value. The UNDER COUNT value reflects the total number of beads of the specified class that have been processed whose F_{mI} values are below that of baseline.

In a preferred embodiment, data (i.e., count, and measured F_{mI} values) for each bead classified as a reject can also be collected.

Interpretation

Following the real-time classification and accumulation of results as described above, the user may select to see a text based presentation or interpretation of the assay's numerical results. During the interpretation phase the assay's real-time numerical results are associated with textual explanations. These textual explanations can be displayed to the user.

It is the function of the interpretation table **420** to associate textual descriptions of an assay's possible outcomes with an actual assay's numerical results. Each row in the interpretation table provides the necessary information to make a single interpretation and typically includes entries for (1) the assay's name, (2) a subset token identifying the class or subset on which the interpretation is based, (3) an outcome identifier for the identified subset, (4) a test-type token, (5) high and low discriminant values for each measurement parameter utilized in the identified test, and (6) a text string describing the row's result.

The test-type token identifies which one of a possible plurality of interpretation tests to perform on the collected (real-time) data during the interpretation phase. In a current

embodiment the test-type token is either '0' or '1'. A value of '0' indicates an OVER/UNDER interpretation test is to be performed. A value of '1' indicates a SHIFT interpretation test is to be performed. These tests are defined in the following manner:

$$\text{OVER/UNDER Test Value} = \frac{OVER\ COUNT}{UNDER\ COUNT}, \text{ and}$$

$$\text{SHIFT Test Value} = \frac{SUM/COUNT}{Baseline\ F_m\ Value},$$

where the variables OVER COUNT, UNDER COUNT, SUM, COUNT, and baseline F_m are described above in Section 6.2(d).

The OVER/UNDER test is generally used for qualitative measurements where the level of reactivity of beads is an indication of the condition or concentration of a biomolecule present in the sample. The shift test is used where the result sought is a determination of the a minimally detectable level of a particular biomolecule. One of ordinary skill will recognize that many other tests could be performed. Examples include ranking, stratification, ratio of means to a standard, or to each other, etc.

In general an interpretation table **420** may associate any number of entries or interpretations (e.g., rows within the table) with a single assay class or bead subset. For instance, bead subset Y could have a single measurement parameter (F_{m1}) associated with it and this measurement parameter could indicate, depending upon its value, that one or more interpretations are appropriate.

Note, the contents of the interpretation table **420** are generated during the preprocessing phase. This implies that the target assay be understood and that the various assay results be considered prior to construction of multiplexed assays.

Interpretation Example

Consider again the assay beadset, introduced above, designed to simultaneously detect four analytes. **Figure 11** shows a sample interpretation table for this assay. Interpretation of the assay's real-time numerical results is initiated by, for example, the user selecting "interpret results" via the inventive method's graphical user interface.

As described above, each bead subset (class) within an assay has an entry or row in the results table, **Figure 10**. The general procedure for interpreting an assay's real-time numerical results is shown in flow-chart form in **Figure 12**. In general, each row of the results table is matched against every row in the interpretation table with the same subset token. If the result of performing the specified test is between the identified row's low and high values, then the associated textual message is displayed to the user. When all rows in the interpretation table for a single results table row have been checked, the next results table row is evaluated. This process is repeated until the every row in the interpretation table has been compared to the appropriate results table entry.

As a specific example, consider the interpretation of subset 50's (KRAS CODON 50 MUTANT, see **Figure 6**) results table entry. The subset's token, 50, is used to identify three rows in the interpretation table (having outcome IDs of 1, 2, and 3) that contain information regarding evaluation of the mutant analyte. For the first identified row, the test-type token indicates a SHIFT type interpretation test is to be performed. Performing this test, as defined above, yields:

$$\text{SHIFT Test Value} = \frac{SUM/COUNT}{Baseline\ F_m\ Value} = \frac{1{,}700{,}000/1{,}000}{170} = 10$$

Next, the computed SHIFT test value is compared against each interval in the identified rows of the interpretation table. For the row having OUTCOME ID equal to 1, since (LOW VALUE = 10) ≤ SHIFT Test Value = 10 ≤ (HIGH VALUE = 667) is true, that row's INTERPRETATION entry – "identical complementary strand" – is displayed to the user. This

process is repeated for subset 50's remaining two rows in the interpretation table. Further, this process is repeated for each row in the results table.

The result of the interpretation phase is a series of textual messages that describe the results of the assay. Conclusion of the interpretation phase marks the end of the assay.

Operational Considerations

Assay definition, discriminant function definition, and interpretation tables are created at the time an assay beadset is created. Baseline classification data is collected only once for a given assay. That is, once an assay is defined and its baseline data is obtained, any number of beadsets can be manufactured to perform the analysis. To allow this "sharing" of baseline data the assay beadset may contain a center or calibration bead subset.

As would be known to those of ordinary skill in the field, a calibration beadset can be used to adjust any given flow cytometer to a standard. Calibration beadsets are typically processed separately from an assay. Further, calibration is generally performed daily. The purpose of calibration is to adjust the sensitivity of a flow cytometer's photomultipliers to accommodate day to day and machine to machine differences.

Unlike prior art calibration techniques which are performed manually, the processing of a calibration beadset and the adjustment of flow cytometer operational parameters (e.g., photomultiplier voltages) is performed under software control automatically. See microfiche appendix A for embodiment details.

Antibody Detection

Assays for antibody are widely used in medicine and clinical analysis for an wide variety of purposes, from detection of infections to determination of autoantibody. The following example illustrates use of the inventive method in an antibody assay and assumes the use of a flow cytometer capable of providing at least five measurements for each bead processed: forward light scatter as classification parameter C_1, side light scatter as classification parameter

C_2, red fluorescence as classification parameter C_3, orange fluorescence as classification parameter C_4, and green fluorescence as measurement parameter F_{m1}.

In one method a number of bead subsets, e.g., subsets 1 through 10 (identified as sS1-sS10), are prepared, for example, by using a cell sorter to sort a heterogeneous population to collect a homogeneous subset or alternatively, by preparing the beads using tightly controlled specifications to ensure production of a homogeneous subset. Each subset is distinguishable by its characteristic pattern of classification parameters C_1, C_2, C_3, and C_4. The beads in each subset are then labeled with a different antigen such as AgA, AgB, etc. so as to create a collection of labeled subsets as follows: sS1-AgA, sS2-AgB, sS3-AgC, sS4-AgD, sS5-AgE, sS6-AgF, sS7-AgG, sS8-AgH, sS9-AgI, and sS10-AgJ.

Antigens AgA through AgJ may be attached to the beads by any of a number of conventional procedures such as by chemical or physical absorption as described by Colvin et al., "The Covalent Binding of Enzymes and Immunoglobulins to Hydrophilic Microspheres" in Microspheres: Medical and Biological Applications, 1-13, CRC, Boca Raton, FL, 1988; Cantarero et al., "The Adsorptive Characteristics of Proteins for Polystyrene and Their Significance in Solid-Phase Immunoassays," Anal. Biochem., 105, 375-382 (1980); and Illum et al., "Attachment of Monoclonal Antibodies to Microspheres," Methods in Enzymol., 112, 67-84 (1985).

After attachment of antigen to the beads' surface, aliquots from each subset are mixed to create a pooled or assay beadset, containing known amounts of beads within each subset. Preferably, the pooled set is prepared with equal volumes of beads from each subset, so that the set contains about the same number of beads from each subset.

The assay beadset may then be incubated with a fluid sample of interest, such as serum or plasma, to test for the presence of antibodies in the fluid that are reactive with antigens on the beads. Such incubation will generally be performed under conditions of temperature, pH, ionic concentrations, and the like that facilitate specific reaction of antibodies in the fluid sample with

antigen on the bead surface. After a period for binding of antibody, the beads in the mixture are centrifuged, washed and incubated (again under controlled conditions) for another period of time with a "secondary" antibody such as, for example, fluorescein labeled goat anti human immunoglobulin. The secondary antibody will bind to and fluorescently label antibodies bound to antigen on the beads. Again after washing (or without washing), the beads are processed by the flow cytometer and the four classification parameters forward light scatter , side light scatter, red fluorescence, and orange fluorescence are measured and used to identify the subset to which each bead in the assay beadset belongs. A simultaneous measurement of green fluorescence (measurement parameter) for each bead allows one to determine whether the bead has antibody bound to it. Because the subset to which a bead belongs is correlated with the presence of a particular antigen, e.g., sS1-AgA, one may readily determine the specificity of the antibody bound to a bead as a function of the subset to which it belongs.

Experimental Example

Three different antigen-antibody pairs were used in a multiplex experiment demonstrating the ability to detect the presence or absence of several antibodies in a single sample. Antigens were coupled to latex microspheres via carbodiimide coupling, and the corresponding antibodies were fluorescently labeled with fluorescein isothiocyanate (green fluorescence – F_m). Each antigen was coupled to a unique microsphere. Baseline data for the fluorescent antibodies and antigen-microsphere complexes used in this experiment are shown in **Figure 13a**. Baseline data for the three bead subsets of **Figure 13a** are given in **Figure 13b**.

The absence of fluorescence (C_2 and C_3) immediately discriminates the clear beads (subset 50) from beads in the other two subsets. Subsets 45 and 50 were further discriminated by side light scatter (C_1) and red fluorescence (C_3). Linear discriminant functions based on these observations and created as described in Section 6.2(c); are shown in **Figure 13c**. Accepting only clear beads with side light scatter (C_1) within ± 0.25 standard deviations of the mean, doublets (two beads stuck together) were eliminated from the analyses. The remaining beads were classified by red fluorescence (C_3) at a midpoint of 59.6. A decision tree based on the discriminant function table (**Figure 13c**) is shown in **Figure 14**.

In this experiment, each of four samples (e.g., blood serum from four patients) contained all three antigen-microsphere complexes and either 1 or 2 different fluorescent antibodies in PBS buffer. After addition of the antibodies, the reactions were incubated at room temperature for 45 minutes, and then analyzed on the "FACSCAN" using side light scatter (C_1), orange fluorescence (C_2), and red fluorescence (C_3) as classification parameters. Green fluorescence was used as the measurement parameter (F_m); an increase in green fluorescence by 30-fold indicates a specific interaction between an antigen and its corresponding fluorescinated antibody. In other words, if a subset's mean measured F_m value is greater than 30-fold times that subset's baseline F_m value, then the target analyte is determined to be present. These "interpretive" observations are embodied in the interpretation table shown in **Figure 13d**.

Once the assay database was built, it was tested by running 5,000 beads from each bead subset individually through the system. After rejecting 23.8% of the beads as doublets, the remaining crimson beads (subset 18) were classified with 99.88% accuracy. Dark red beads (subset 45) were classified with 99.96% accuracy with 22.9% rejected as doublets. Clear beads (subset 50) were classified with 100% accuracy with 9.4% of the beads rejected as doublets.

The three bead subsets were pooled to form an assay beadset and divided into 4 sample tubes and processed by the system shown in **Figure 1**. The contents of each sample and the mean measured fluorescence (F_m) for each bead subset are listed in **Figure 13e**. The inventive method correctly identified the antibody or antibodies present in each sample.

An Experimental Refinement

In an alternative embodiment using a C_4 (e.g., orange fluorescence) labeled reactant as a classification parameter, a variety (for example five) of protein antigens are employed. Bead subsets are first generated based on differences in one or more of C_1, C_2, and C_3. Next, a selected antigen labeled with Cy3NHS (an orange fluorophore) is bound to the beads in each subset. To minimize the measured orange fluorescence coefficient of variation for each bead subset, the beads are sorted with a high speed cell sorter so that only a narrow range of antigen

(orange fluorophore) is found on each bead within a subset. Care should be taken to select or prepare the beadset so that different C_4 values are measured/obtained for each of the (e.g., five) different antigens used. In other words, the measured intensity of C_4 for AgA should differ from the measured intensity of C_4 from AgB, etc. To ensure that uniformity is achieved, saturation binding with fluoresceinated monoclonal antibody is tested – each bead ought to have restricted ranges of both orange and green fluorescence. While the construction of beadsets by this method is more laborious, the increase in measurement precision may be useful and will allow the sampling of fewer beads to arrive at a suitable determination of antibody concentration.

The assays previously mentioned measure any antibody with specificity for antigen upon an appropriately labeled bead. The antigen can be quite simple or rather complex and thus, the inventive methods can measure a highly restricted antibody or a broad array of antibodies. For example, a hexapeptide just large enough to bind to a monoclonal antibody can be employed as antigen or a large protein with many epitopes can be used. One of ordinary skill will recognize that the level of antibody eventually found associated with the bead (F_{ml}) is a function of the number of epitopes per bead, the concentration of epitopes, the amount of antibody and the affinity of the antibody and the valence of the antibody-antigen interaction.

Displacement Assays

Assays for many substances in a clinical laboratory are based on the interference with specific ligand-ligate or antigen-antibody interactions. In these assays, one member of the ligand-ligate pair is labeled with the F_m fluorophore and one member is immobilized on the beads. Soluble, unlabeled material (analyte) ,which may be ligand or ligate, is added to the reaction mixture to competitively inhibit interaction of the labeled component with the immobilized component. It is usually not important which member of the pair is labeled and which is immobilized; however, in certain assays, functional advantages may dictate the orientation of the assay.

In an exemplary assay of this type, each bead subset is modified with an antigen. The antigen-coated beads are then reacted with an F_m labeled antibody specific for the antigen on the

bead surface. Subsequent addition of a test fluid containing soluble analyte (inhibitor) will displace the F_m labeled antibody from the beads in direct proportion to the concentration of the soluble analyte. A standard curve of known analyte concentrations is used to provide accurate quantification of analyte in the test sample.

One of ordinary skill will recognize that the time necessary to achieve equilibrium may be quite lengthy due to the kinetics and association constant of the interaction. To lessen the time required for the assay, the fluid containing the beadset may be subjected to dissociating conditions such as a change in pH, ionic strength or temperature, after mixture of the beadset with the sample to be tested. Alternatively, the F_m labeled component may be added to the beadset after addition of the test sample. In either case, it is not necessary for equilibrium to be achieved to determine analyte concentration if the kinetics and linearity of the assays have been established.

Additional Experimental Examples

The following series of experimental examples illustrates how the above referenced techniques can be used in practice in effective diagnostic assays. In one embodiment for example, a competitive inhibition analysis is used to quantitate levels of selected analytes, here IgG, IgA, and IgM. A second experimental refinement demonstrates the utility of multiplexed assays in epitope mapping of a monoclonal antibody. In one embodiment, that approach involved the use of antibody detection technology using a fluoresceinated monoclonal antibody in combinatorial epitope screening (e.g. of peptide libraries) to map a particular epitope to which a monoclonal antibody of interest bound, together with a displacement (competitive inhibition) aspect to demonstrate the specificity of the assay. Also described is a ToRCH assay for screening of human serum for antibodies to a number of infectious agents known to pose special hazards to pregnant women. Allergy screening is exemplified by detection of serum IgE against a panel of grass antigens. Yet an additional experimental example reflects the ability of the multiplexed assay in pregnancy testing, e.g. in testing for hormones or other analytes commonly elevated during pregnancy. Each of these examples is set forth below.

Simultaneous competitive inhibition assay of human immunoglobuling G, A and M levels in serum

This example illustrates the determination of multiple analyte levels in a liquid sample simultaneously using competitive inhibition analysis. The use of a competitive inhibition assay to accurately determine analyte levels in liquid solutions is a commonly used format for many analyte assays. The uniqueness of this assay is the simultaneous determination of three distinct serum proteins at the same time in the same tube from one serum sample.

Immunoglobulins G, A and M are three distinct serum proteins whose levels are determined by a number of genetic and environmental factors in human serum. As changes to these levels may indicate the presence of disease, clinicians often request assay determinations of IgG, A and M using conventional techniques. The most common technique is nephelometry that depends upon the absorption of light by precipitates formed between these immunoglobulins and antibodies made in animals to the human immunoglobulins. As these immunoglobulins are present in human serum at fairly high levels, this type of assay is sufficient. Nephelometry however suffers from a number of limitations including the need for large quantities of reagents, long reaction times for precipitation to equilibrate and an inability to perform more than one reaction per tube or sample.

Three competitive inhibition assays are described, one for human IgG, one for human IgM and one for human IgA using three Differentially Fluorescent Microspheres (DFM). Each assay consists of a DFM coated with the immunoglobulin of choice and a polyclonal, goat anti-human Ig labeled with a green fluorescent molecule (Bodipy). In the absence of inhibitor, the Bodipy -antibody causes the immunoglobulin (Ig) coated microsphere to emit green fluorescence (F_m). In the presence of inhibitor (soluble Ig), the green signal is reduced. Each assay is balanced to reflect a sensitivity range near the physiological level of the Ig in question at a 1:500 dilution of human serum. Once balanced, the three assays were combined into a multiple analyte format and assayed simultaneously using flow cytometry.

Antibody labeling: Goat anti-human IgG, goat anti-human IgA, and goat anti-human IgM antibodies (Cappel Division, Organon Teknika, Durham, NC) were labeled with Bodipy FL-CASE (Molecular Probes, Inc., Eugene, OR) using methods described by the manufacturer of the Bodipy succinymidyl ester. The resulting Bodipy labeled antibodies were stored in PBS containing 1 mg/mL BSA as stabilizer.

Antigen conjugation to microspheres: Four DFM (5.5 μM carboxylate, Bangs Laboratories, Inc. (Carmel, IN), dyed by Emerald Diagnostics, Inc. (Eugene, OR)) were conjugated separately to human IgG, human IgA, human IgM (Cappel Division, Organon Teknika, Durham, NC) and BSA with a two-step EDC coupling method (Pierce Chemicals, Rockford, IL) using sulfo-NHS to stabilize the amino-reactive intermediate. 100 μL of each bead type (4.2×10^7 microspheres) was activated for 20 minutes in a total volume of 500 μL containing 500 μg of EDC and Sulfo-NHS in 50 mM sodium phosphate buffer, pH 7.0. The microspheres were washed twice with 500 μL PBS, pH 7.4 using centrifugation at 13,400 x g for 30 seconds to harvest the microspheres. Activated, washed beads were suspended in 250 μL of a 0.05 mg/mL solution of protein in PBS, pH 7.4. After 1 hour, the microspheres were blocked by addition of 250 μL of 1.0 mg/mL BSA, 0.02% Tween, 0.2 M glycine, in PBS, pH 7.4 and incubated for an additional 30 minutes. Protein coated microspheres were washed twice with 500 μL 0.02% Tween 20, 1 mg/mL BSA in PBS, pH 7.4 (PBSTB). and stored in PBSTB at approximately 3,000,000 microspheres/mL. Microsphere concentrations were determined using a hemacytometer.

Determination of appropriate ranges of quantitation for each Ig assay: The normal range of human Ig levels in serum as reported in Clinical Chemistry: Principles and Technics, 2nd Edition, Edited by R.J. Henry, D.C. Cannon and J.W. Winkleman are 569-2210 mg/dL for IgG, 51-425 mg/dL for IgA and 18-279 mg/dL for IgM. Each inhibition assay was designed to be sensitive to inhibition across these ranges.

Single analyte assay: 10 μL of dilutions of a serum calibrator with known Ig levels (Kamiya Biomedical, Thousand Oaks, CA) was first mixed with 10 μL of Ig loaded microspheres containing 7,500 beads. Next, 10 μL of the Bodipy-labeled Goat Anti-Ig was added and the

mixture incubated at ambient temperature for 30 minutes. The mixture was diluted to 300 µL in PBSTB and assayed by flow cytometry. For IgG, the Bodipy-labeled goat anti-hIgG was used at 30 µg/mL. For IgA, the Bodipy-labeled goat anti-hIgA was used at 8 µg/mL. For IgM, the Bodipy-labeled goat anti-hIgM was used at 2.5 µg/mL.

Cross reactivity assay: Equivalent amounts of each of the four protein loaded microspheres were mixed to produce a bead mixture. 10 µL of the bead mixture (7,500 microspheres) was mixed with 10 µL of diluted serum calibrators of known Ig level. The assay was initiated by addition of 10 µL of one of the Bodipy-labeled antibodies "spiked" with a small quantity of soluble Ig antigen to alleviate the "hook effect". The mixtures were incubated for 30 minutes, diluted to 300 µL in PBSTB and assayed by flow cytometry. As before for the single analyte assay, the Bodipy-labeled goat anti-hIgG was used at 30 µg/mL. For IgA, the Bodipy-labeled goat anti-hIgA was used at 8 µg/mL. For IgM, the Bodipy-labeled goat anti-hIgM was used at 2.5 µg/mL. The quantities of antigen "spikes" were 1.6 µg/mL for IgG, 0.6 µg/mL for IgA and 0.4 µg/mL for IgM.

Multiple analyte assay: Equivalent amounts of each of the four protein loaded microspheres were mixed to produce a bead mixture. 10 µL of the bead mixture (7,500 microspheres) was mixed with 10 µL of diluted serum calibrators of known Ig level as well as three other calibrator sera of known Ig level to serve for this purpose as unknowns. The assay was initiated by addition of 10 µL of a mixture of the three Bodipy-labeled antibodies "spiked" with a small quantity of the three soluble Ig antigen to alleviate the "hook effect". The mixtures were incubated for 30 minutes, diluted to 300 µL in PBSTB and assayed by flow cytometry. As before, the Bodipy-labeled goat anti-hIgG was used at 30 µg/mL. For IgA, the Bodipy-labeled goat anti-hIgA was used at 8 µg/mL. For IgM, the Bodipy-labeled goat anti-hIgM was used at 2.5 µg/mL. The quantities of antigen "spikes" were 1.6 µg/mL for IgG, 0.6 µg/mL for IgA and 0.4 µg/mL for IgM.

Results

IgG single analyte assay: Results of the single analyte inhibition analysis for IgG level is shown in **Table 1** and **Figure 15A**. This assay was designed to be most sensitive to inhibition in the anticipated range of IgG in human serum at a 1:500 dilution. In **Figure 15A**, the area of the inhibition curve between the dotted lines, left and right, cover the range of sensitivity. In this case, the inhibitor was known amounts of human IgG from a serum calibrator diluted into human serum containing no IgG, IgA or IgM. Dilutions of the calibrator were then diluted 1:500 in PBSTB and included as inhibitor in the assay. The Bodipy-labeled anti-hIgG was used at 30 μg/mL in PBSTB. 7,500 microspheres were used in this experiment and 250 were counted by flow cytometry. Note that as the amount of soluble IgG increased, the degree of inhibition as monitored by the MIF of F_m increased proportionally until saturation of the system was achieved. On the other end of the inhibition curve note that the lower levels of soluble inhibitor caused an elevation in the MIF of F_m as compared with the negative control (human serum with no Ig). This "hook effect" is common in immunoassay and can be adjusted up or down the inhibition curve by adjusting both the amount of antibody and antigen in the soluble portion of the assay. The "hook effect" was most prominent in the IgG assay due to the higher concentrations of both antigen and antibody per microsphere. This was necessary as IgG is found in serum at higher concentrations than IgA or IgM.

IgA single analyte assay: Results of single analyte inhibition analysis for IgA level is shown in **Table 1** and **Figure 15B**. This assay was designed to be most sensitive to inhibition in the anticipated range of IgA in human serum at a 1:500 dilution. In **Figure 15B**, the area of the inhibition curve between the dotted lines, left and right, cover the range of sensitivity. In this case, the inhibitor was known amounts of human IgA from a serum calibrator diluted into human serum containing no IgG, IgA or IgM. Dilutions of the calibrator were then diluted 1:500 in PBSTB and included as inhibitor in the assay. The Bodipy-labeled anti-hIgA was used at 8 μg/mL in PBSTB. 7,500 microspheres were used in this experiment and 250 were counted by flow cytometry. Note that as the amount of soluble IgA increased, the degree of inhibition as monitored by the MIF of F_m increased proportionally until saturation of the system was achieved. On the other end of the inhibition curve note that the lower levels of soluble inhibitor

cause a slight elevation in the MIF of F_m as compared with the negative control (human serum with no Ig). The "hook effect" was much less pronounced for both IgA and IgM due to their lower concentrations in serum.

IgM single analyte assay: Results of single analyte inhibition analysis for IgM level is shown in **Table 1** and **Figure 15C**. This assay was designed to be most sensitive to inhibition in the anticipated range of IgM in human serum at a 1:500 dilution. In **Figure 15C**, the area of the inhibition curve between the dotted lines, left and right, cover the range of sensitivity. In this case, the inhibitor was known amounts of human IgM from a serum calibrator diluted into human serum containing no IgG, IgA or IgM. Dilutions of the calibrator were then diluted 1:500 in PBSTB to be included as inhibitor in the assay.. The Bodipy-labeled anti-hIgM was used at 2.5 μg/mL in PBSTB. 7,500 microspheres were used in this experiment and 250 were counted by flow cytometry. Note that as the amount of soluble IgM increased, the degree of inhibition as monitored by the MIF of F_m increased proportionally until saturation of the system was achieved. On the other end of the inhibition curve note that the lower levels of soluble inhibitor cause a slight elevation in the MIF of F_m as compared with the negative control (PBS with no added IgM). The "hook effect" is much less pronounced for both IgA and IgM due to their lower concentrations in serum.

Cross reactivity analysis: To determine the cross-reactivity of the various assay components, a multiple analyte assay was performed using only one of the three Bodipy-labeled antibodies. Equivalent numbers of the IgG, IgA, IgM and BSA beads were mixed to make a GAM mixed bead set. To 10 μL of the bead set (7,500 microspheres) was added 10 μL of dilutions of the calibrator containing IgG, IgA and IgM. The multiple analyte assay was then performed using only one of the Bodipy-labeled anti-IgG, IgA or IgM preparations rather than a mixture. **Table 2** and **Figures 16A, 16B,** and **16C** show the results of these assays. Results indicated that Anti-IgG-Bodipy only reacted with DFM-IgG Bodipy and not the IgA or IgM beads. No cross-reactivity with IgA or IgM was noted and the assay was validated for further multiple analyte analysis. Also added to this analysis was the antigen "spike". By adding a small amount of soluble antigen to the probe antibody solution the "hook effect' can be minimized. Note in the

IgG cross-reactivity experiment that the MIF of F_m for negative control is higher than the lowest concentration of inhibitor. By spiking the experiment with 1.6 μg/mL IgG the hook effect has no effect at the lower end of inhibitor range leading to a more accurate assay over the entire dynamic range.

GAM simultaneous analysis: Equivalent numbers of the IgG, IgA, IgM and BSA beads were mixed to make a GAM mixed bead set. To 10 μL of the bead set (7,500 microspheres) was added 10 μL of dilutions of the calibrator containing IgG, IgA and IgM. Also included were several additional calibrators that served as unknowns for the demonstrative purpose of this assay. The multiple analyte assay was then initiated by adding 10 μL of a mixture of the Bodipy-labeled anti-IgG, IgA and IgM plus the soluble Ig "spikes". After a 30 minute, room temperature incubation the reaction mixture was diluted to 300 μL and 1000 microspheres counted by flow cytometry. **Tables 3-5** and **Figures 17-19** show the results of these assays. For each of the inhibition curves produced, a polynomial trendline was used as a non-linear regression analysis. The fit of this trendline to the data was demonstrated by the R^2 correlation factor (1.0 is a perfect fit). The factors of the polynomial formula were used to predict the quantity of inhibitor in each dilution of calibrator and "unknown" serum. The differences between the predicted inhibitor quantities and actual amounts were also included in **Tables 3-5**. Results indicate that this multiple analyte inhibition assay can determine the level of these 3 serum proteins with an error of less than 10 %. Coefficients of variation (CV) between the triplicate data points indicated that the assay was highly precise (no CV greater than 6%). Limits of quantitation for each assay were 400 - 3000 mg/dL for IgG, 60- 455 mg/dL for IgA, and 36 - 272 mg/dL for IgM. **Figure 20** shows the results of the three assays graphically represented on the same graph as all three assays were performed at the same time in the same tube.

A multiple analyte, competitive inhibition assay for human serum IgG, IgA, and IgM levels has been developed. This assay, that allows the simultaneous assay of these three protein levels in serum diluted 1:500, demonstrated excellent sensitivity, precision and accuracy.

TABLE 1: Single analyte inhibition assays

	IgG	MIF of		**IgA**	MIF of		**IgM**	MIF of
Tube#	**mg/dL**	*Fm*		**mg/dL**	*Fm*		**mg/dL**	*Fm*
1	0	1445		0	1654		0	1765
2	0.026	1500		0.0040	1645		0.0024	1794
3	0.11	1460		0.016	1729		0.010	1929
4	0.42	1512		0.064	1734		0.038	1921
5	1.7	1426		0.26	1733		0.15	1815
6	6.8	1619		1.02	1747		0.61	1829
7	27.1	1684		4.1	1746		2.4	1833
8	108	1943		16.4	1788		9.8	1807
9	163	1898		24.6	1813		14.7	1792
10	244	1885		36.9	1806		22.0	1723
11	366	1624		55.3	1703		33.0	1704
12	549	1456		83.0	1391		49.6	1446
13	824	998		125	971		74.3	1267
14	1235	722		187	558		112	879
15	1853	473		280	336		167	591
16	2779	350		420	240		251	360
17	4169	313		630	140		376	269
18	6253	316		945	103		564	242
19	9380	196		1418	75		847	136
20	14070	165		2127	54		1270	102

TABLE 2: Cross-reactivity analysis in multiple analyte assay

1) GAM Beads reacted with **anti-IgG -Bodipy @ 30 µg/mL** + Ag spikes.

Tube	hu IgG mg/dL	**Bead 1- MIF HuIgG**	hu IgA mg/dL	**Bead 2- MIF HuIgA**	hu IgM mg/dL	**Bead 3- MIF HuIgM**	**Bead 4- MIF BSA**
1	0	**1868**	0	4	0	6	5
2	400	**1702**	60.5	4	36.1	7	5
3	561	**1463**	84.7	4	50.6	5	5
4	785	**1218**	119	3	70.8	4	5
5	1099	**880**	166	3	99.2	3	5
6	1538	**674**	233	3	139	3	5
7	2154	**549**	326	2	194	3	5
8	3015	**450**	456	2	272	2	5

2) GAM Beads reacted with **anti-IgA -Bodipy @ 8 µg/mL** + Ag spikes.

11	0	3	0	**1800**	0	2	3
12	400	2	60.5	**1455**	36.1	2	3
13	561	8	84.7	**1225**	50.6	1	3
14	785	3	119	**930**	70.8	1	3
15	1099	2	166	**605**	99.2	1	3
16	1538	2	233	**392**	139	1	3
17	2154	2	326	**278**	194	1	3
18	3015	2	456	**163**	272	1	3

3) GAM Beads reacted with **anti-IgM -Bodipy @ 2.5 μg/mL** + Ag spikes.

21	0	3	0	6	0	**1536**	2
22	400	3	60.5	9	36.1	**1284**	2
23	561	3	84.7	2	50.6	**1135**	2
24	785	3	119	2	70.8	**1011**	2
25	1099	2	166	2	99.2	**776**	2
26	1538	2	233	1	139	**620**	2
27	2154	2	326	2	194	**463**	2
28	3015	2	456	1	272	**330**	2

TABLE 3: Multiple analyte IgG inhibition data

Tube #	hIgG mg/dL	MIF of *Fm*	Average MIF	MIF CV	Calculated mg/dL	% Difference
1		1917				
2	**0**	1943	**1926**	0.6%	**na**	na
3		1918				
4		1811				
5	**400.4**	1737	**1772**	1.7%	**399.3**	0.3%
6		1767				
7		1408				
8	**560.6**	1529	**1471**	3.4%	**566.9**	-1.1%
9		1476				
10		1250				
11	**784.8**	1163	**1236**	4.4%	**775.3**	1.2%
12		1295				
13		852				
14	**1099**	867	**862**	0.8%	**1102.5**	-0.3%
15		868				
16		661				
17	**1538**	726	**691**	3.9%	**1556.1**	-1.2%
18		687				
19		575				
20	**2154**	575	**580**	1.1%	**2126.3**	1.3%
21		589				
22		461				
23	**3015**	466	**468**	1.5%	**3025.1**	-0.3%
24		478				

"UNKNOWNS"

25		1691				
26	**446**	1657	**1657**	1.7%	**411.3**	7.8%
27		1624				
28		749				
29	**1243**	737	**763**	3.8%	**1316.8**	-5.9%
30		804				
31		464				
32	**3045**	486	**476**	1.9%	**2947.1**	3.2%
33		479				

TABLE 4: Multiple analyte IgA inhibition data

Tube #	hIgA mg/dL	MIF of *Fm*	Average MIF	MIF CV	Calculated mg/dL	% Difference
1		1954				
2	**0**	1941	**1952**	0.4%	**na**	na
3		1960				
4		1661				
5	**60.5**	1664	**1665**	0.2%	**60.5**	0.0%
6		1669				
7		1222				
8	**84.7**	1391	**1307**	5.3%	**84.7**	0.0%
9		1308				
10		1055				
11	**118.6**	974	**1051**	5.9%	**118.6**	0.0%
12		1125				
13		615				
14	**166.1**	595	**606**	1.4%	**166.1**	0.0%
15		607				
16		376				
17	**232.5**	426	**400**	5.1%	**232.6**	0.0%
18		399				
19		283				
20	**325.5**	280	**287**	2.9%	**325.4**	0.0%
21		299				
22		193				
23	**455.7**	198	**195**	1.2%	**455.7**	0.0%
24		193				

"UNKNOWNS"

25		1569				
26	**65**	1483	**1504**	3.1%	**68.2**	-4.9%
27		1460				
28		455				
29	**187**	457	**477**	6.1%	**197.2**	-5.5%
30		518				
31		187				
32	**454**	199	**201**	5.9%	**445.4**	1.9%
33		216				

TABLE 5: Multiple analyte IgM inhibition data

Tube #	hIgM mg/dL	MIF of *Fm*	Average MIF	MIF CV	Calculated mg/dL	% Difference
1		1566				
2	**0**	1615	**1605**	1.8%	**na**	na
3		1635				
4		1345				
5	**36.1**	1312	**1328**	1.0%	**35.7**	1.2%
6		1328				
7		1133				
8	**50.6**	1182	**1155**	1.8%	**52.9**	-4.6%
9		1151				
10		1038				
11	**70.8**	994	**1035**	3.2%	**68.1**	3.9%
12		1074				
13		728				
14	**99.2**	733	**735**	0.9%	**100.7**	-1.5%
15		744				
16		514				
17	**138.8**	585	**546**	5.4%	**138.3**	0.4%
18		539				
19		424				
20	**194.4**	414	**419**	1.0%	**194.0**	0.2%
21		418				
22		298				
23	**272.1**	339	**315**	5.6%	**272.4**	-0.1%
24		307				

"UNKNOWNS"

25		1266				
26	**40**	1248	**1241**	1.9%	**42.8**	-7.0%
27		1209				
28		608				
29	**113**	621	**635**	4.7%	**116.2**	-2.8%
30		677				
31		289				
32	**268**	315	**306**	3.9%	**281.0**	-4.9%
33		313				

Epitope Mapping of a Monoclonal Antibody using Flow Cytometry.

This example demonstrates the screening of combinatorial chemistry products for a biologically active molecule. The generation of random chemical products for empirical discovery of biologically significant molecules is a method that holds great promise for progress in numerous disciplines of science including biology, pharmacology and medicine. One general problem with the technique is the screening of large numbers of unique molecules for a specific activity. Screening methods are required that provide high throughput levels of screening with adequate specificity and sensitivity for detection of the biological event in question.

An experiment was designed to demonstrate the screening of peptides for the epitope of a monoclonal antibody. A monoclonal antibody (MAB 384) was chosen that was produced using the spleen cells of a mouse hyper-immunized with a defined peptide (amino acid 67-74) from the amino acid sequence of human myelin basic protein (MBP). Using the amino acid sequence of this region of MBP, nine overlapping octapeptides were synthesized that covered the predicted epitope. To the carboxyl terminal end of each peptide, glycine-lysine-biotin residues were added. Nine Differentially Fluorescent Microspheres (DFM) were each coated with avidin and one unique peptide of the set was linked through the avidin-biotin interaction to one unique member of the bead set. This resulted in a set of microspheres that contained nine

members each carrying a unique peptide either flanking or representing the monoclonal antibody's epitope. The bead carrying the epitope peptide was detected using the MAB 384 antibody labeled with a green fluorescent tag in a multiple analyte analysis. The detection was shown to be specific for the peptide in question by competitive inhibition and was not affected by high levels of free biotin.

Antibody labeling: MAB 384 (Chemicon International, Inc., Temecula, CA) was labeled with Bodipy FL-X (Molecular Probes, Inc., Eugene, OR) using methods described by the manufacturer of the Bodipy succinymidyl ester. Absorbance at 280 nm and 504 nm revealed that the resulting Bodipy-labeled antibody had a Bodipy to protein ratio of 3.31 and was stored in PBS containing 1 mg/mL BSA as stabilizer.

Avidin conjugation to microspheres: Nine distinctly dyed DFM (5.5 μM, Bangs Laboratories, Inc. (Carmel, IN), dyed by Emerald Diagnostics, Inc. (Eugene, OR)) were conjugated separately to Neutravidin (deglycosylated avidin) with a two-step EDC coupling method (Pierce Chemicals, Rockford, IL) using sulfo-NHS to stabilize the amino-reactive intermediate. 20 μL (8.4 million microspheres) of each bead type was activated for 20 minutes in a total volume of 100 μL containing 500 μg of EDC and Sulfo-NHS in 50 mM sodium phosphate buffer, pH 7.0. The microspheres were washed twice with 100 μL PBS, pH 7.4 using centrifugation at 13,400 x g for 30 seconds to harvest the microspheres. Activated, washed beads were suspended in 50 μL of a 0.25 mg/mL solution of Neutravidin in PBS, pH 7.4. After 2 hours, the microspheres were blocked by addition of 50 μL of 0.2 M glycine, 0.02% Tween 20 in PBS, pH 7.4 and incubated for an additional 30 minutes. Protein coated microspheres were washed twice with 100 μL 0.02% Tween 20, 1 mg/mL BSA in PBS, pH 7.4 (PBSTB) and stored in PBSTB at approximately 3,000,000 microspheres/mL as determined by hemocytometer count.

Peptide attachment to microspheres: Each of the nine DFM conjugated to Neutravidin were treated separately with one of the nine biotinylated peptides. 10 μL of biotinylated peptides at 100 - 200 ng/mL was mixed with 10 μL of microspheres and reacted for 5 minutes followed by

2 x 100 µL washes in PBSTB. The peptide loaded microspheres were suspended in 20 µL of PBSTB.

Single analyte assay: 10 µL of each of the peptide loaded microspheres was reacted with 10 µL of the Bodipy-labeled MAB 384 at 15.5 µg/mL in PBSTB for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry. Negative controls included the microspheres without peptide and with the Bodipy MAB 384.

Multiple analyte assay: 10 µL of each of the 9 peptide loaded microspheres was mixed to produce a bead set. 10 µL of the set was reacted with 10 µL of the Bodipy-labeled MAB 384 at 15.5 µg/mL in PBSTB for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry. Negative controls included the microsphere set without peptide and treated with the Bodipy MAB 384.

Competitive inhibition with soluble peptide: 10 µL of each of the 9 peptide loaded microspheres was mixed to produce a bead set. 10 µL of the Bodipy-labeled MAB 384 at 15.5 µg/mL in PBSTB was reacted with 10 µL of soluble peptide containing the epitope sequence HYGSLPQK (SEQ ID NO. 1) at 10 µg/mL and incubated for 1 hr. The microsphere set was then treated with peptide absorbed Bodipy-labeled MAB 384 at 15.5 µg/mL for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry.

Examination of the effects of free biotin: 10 µL of each of the 9 peptide loaded microspheres was mixed to produce a bead set. 10 µL of the mixture was reacted with 10 µL of 10 µg/mL free biotin and incubated for 1 hr. The microsphere set was then treated with Bodipy-labeled MAB 384 at 15.5 µg/mL for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry.

Results

Description of peptides to be screened: The amino acid sequence upstream and downstream from the epitope of monoclonal antibody MAB 384 (amino acid 67-74, YGSLPQ, SEQ ID NO.

2) was determined using the published amino acid sequence (Roth, H.J., et al., J. Neurosci. Res.. *17*, 321-328, 1990). The table below shows the amino acid sequence of the nine overlapping peptides produced for the screening assay. Note that to the carboxy-terminal end of all peptides was added a glycine (G)-lysine (K)-biotin.

1	GLCNMYKDGK-biotin
2	MYKDSHHPGK-biotin
3	SHHPARTAGK-biotin
4	ARTAH**YGS**GK-biotin
5	H**YGSLPQ**KGK-biotin
6	**LPQ**KSHGRGK-biotin
7	SHGRTQDEGK-biotin
8	TQDENPVVGK-biotin
9	NPVVHFFKGK-biotin

Single vs. multiple analyte analysis: Each of the nine DFM coated with Neutravidin was reacted for 5 minutes with one of the nine biotinylated peptides diluted to 250 ng/mL in PBS. For single analyte analysis, each separate microsphere was reacted with Bodipy-labeled MAB 384 at 15.5 µg/mL for 60 minutes and the mixture assayed using flow cytometry. The Mean Intensity of Fluorescence (MIF) of the green fluorescence channel (F_m) is shown for each peptide-bead as the darker set of bars in **Figure 21**. The darkest bars represents single analyte analysis of each bead in the absence of peptide as a negative control.

For multiple analyte analysis, the nine bead-peptides were mixed and then reacted with Bodipy-MAB 384 at 15.5 µg/mL. After 60 minutes, the mixture was assayed using flow cytometry and results (MIF of F_m) are also shown in **Figure 21**. Both assays minus added peptide are shown as a negative control. Results indicated that peptide #5 contained the epitope for MAB 384. Peptides #4 and #6 although containing 3 of the epitope's amino acids showed little reactivity. The multiple and single analyte assays provided identical results. Numerical data is shown in **Table 6**.

Competitive inhibition using soluble epitope peptide: To further demonstrate the specificity of the assay, soluble peptide containing the epitope (#5) was used to inhibit the reaction shown in **Figure 21**. A 10 μL aliquot of the Bodipy-labeled MAB 384 was mixed with an equal volume of the epitope containing peptide (HYGSLPQK) at 10 μg/mL. After 1 hour the mixture was reacted with 10 μL of the bead mixture for 1 hour and assayed by flow cytometry. Results shown in **Figure 22** reveal that the reaction was significantly inhibited to a MIF of F_m of 53. Numerical data for the inhibition assay is shown in **Table 7**.

Effects of free biotin: The high avidity of the biotin-avidin interaction makes it unlikely that the various peptides could be released or exchanged from microsphere to microsphere, To demonstrate that such a release or exchange does not occur under strenuous conditions the following experiment was performed. A 10 μL aliquot of free biotin at 10 μg/mL (40 μM) was incubated with 10 μL of the bead-peptide mixture for 1 hour and then the microspheres reacted with the MAB 384 Bodipy at 15.5 μg/mL for 1 hour and assayed by flow cytometry. Results shown in **Figure 23** indicate that the free biotin at 10 μg/mL did not displace significant amounts of the biotinylated epitope peptide. Numerical data for the inhibition assay is shown in **Table 8**.

This epitope mapping example demonstrates the useful application of the instant invention to the area of combinatorial screening. The peptide carrying the epitope for the mouse monoclonal antibody screened in this example was clearly identified in a set of nine peptides. The identification was further shown to be specific by competitive inhibition with soluble epitope peptide. In addition, the stability of the avidin-biotin interaction for use with flow cytometry was demonstrated in an excess of free biotin.

TABLE 6

MIF of *Fm*

Bead	Peptide plus GL-Biotin	Assayed Single	Single no peptide	Assayed Multiple	Multiple no peptide
70/50	GLCNMYKD	72	66	28	28
60/70	MYKDSHHP	57	48	36	36
40/70	SHHPARTA	47	43	36	34
40/50	ARTAH**YGS**	57	47	35	27
70/70	H**YGSLPQ**K	**1381**	66	**1348**	25
40/40	**LPQ**KSHGR	43	44	67	25
40/60	SHGRTQDE	42	54	35	26
70/60	TQDENPVV	73	70	32	23
70/40	NPVVHFFK	60	60	29	21

TABLE 7

Peptide plus GL-Biotin	Assayed w/free Biotin
GLCNMYKD	4
MYKDSHHP	7
SHHPARTA	12
ARTAH**YGS**	13
HYGSLPQK	**53**
LPQKSHGR	15
SHGRTQDE	11
TQDENPVV	9
NPVVHFFK	17

TABLE 8

Bead	Peptide plus GL-Biotin	MIF Multiple	MIF w/ Biotin
70/50	GLCNMYKD	13	17
60/70	MYKDSHHP	17	19
40/70	SHHPARTA	20	22
40/50	ARTAH**YGS**	20	26
70/70	H**YGSLPQK**	**915**	**1023**
40/40	**LPQ**KSHGR	32	20
40/60	SHGRTQDE	19	23
70/60	TQDENPVV	31	34
70/40	NPVVHFFK	31	36

Multiple Analyte Simultaneous ToRCH Assay for Seroconversion.

This example demonstrates the utility of this invention in the screening of human serum for antibodies to infectious disease agents. Screening of serum for antibodies to certain infectious disease agents is often the only method available to determine if a patient has been, or is infected with the agent in question. For example, a common method of diagnosing HIV infection is by detection of HIV specific antibodies in the serum. This phenomenon known as seroconversion is commonly employed for diagnosis of several important pathogenic infections. One of the most commonly employed assay panels of this type is the ToRCH panel. ToRCH assays detect both serum IgG and serum IgM responses to *Toxoplasma gondii*, Rubella virus, Cytomegalovirus, and Herpes Simplex Virus Types 1 and 2. The importance of this assay especially to the pregnant woman has been well documented as any one of these infectious agents is capable of crossing the placental barrier and entering the immunologically naive fetus. These infectious agents can cause severe damage to the fetus and must be avoided. Currently , all ToRCH panel assays for antibodies specific to each of these pathogens is performed separately in a unique assay tube or microtiter well. This invention provides for a multiple analyte format that allows assay for either IgG or IgM antibodies specific for each of the five pathogens at the same time in the same tube with the same sample.

A ToRCH assay using flow cytometry has been developed by coupling purified antigens of *T. gondii*, Rubella, CMV and HSV Type 1 and Type 2 to five Differentially Fluorescent Microspheres (DFM). The specificity of the assay has been demonstrated by treating this bead set with human serum calibrators certified to be either positive or negative for all five agents. After this treatment, the bead set was treated with either Goat anti-human IgG-Bodipy or Goat anti-human IgM-Bodipy used to develop the assay. In addition, a third calibrator with known levels of reactivity to each agent was assayed and the results reported.

Antibody labeling: Goat anti-human IgG and goat anti-human IgM (Cappel Division, Organon Teknika, Durham, NC) were labeled with Bodipy FL-CASE (Molecular Probes, Inc., Eugene, OR) using methods described by the manufacturer of the Bodipy succinymidyl ester. Bodipy-labeled antibodies were stored in PBS containing 1 mg/mL BSA as stabilizer.

Antigen conjugation to microspheres: Five DFM (5.5 μM carboxylate, Bangs Laboratories, Inc., Carmel, IN, dyed by Emerald Diagnostics, Inc., Eugene, OR) were conjugated separately to the five ToRCH antigens (Viral Antigens, Inc.) with a two-step EDC coupling method (Pierce Chemicals, Rockford, IL) using sulfo-NHS to stabilize the amino-reactive intermediate. All antigens were dialyzed into PBS to remove any reactive amino groups such as sodium azide or glycine. The *T. gondii* preparation (Chemicon, Inc., Temecula, CA) was sonicated for 2 minutes in PBS, 10 mM EDTA to lyse the tachyzoites. 20 μL (8.4 million microspheres) of each bead type was activated for 20 minutes in a total volume of 100 μL containing 500 μg of EDC and Sulfo-NHS in 50 mM sodium phosphate buffer, pH 7.0. Microspheres were washed twice with 200 μL PBS, pH 7.4 using centrifugation at 13,400 x g for 30 seconds to harvest the microspheres. Activated and washed beads were suspended in 100 μL of antigen at 0.05 to 0.15 mg/mL in PBS, pH 7.4. After 2 hours, the microspheres were blocked by addition of 100 μL of 0.2 M glycine, 0.02% Tween 20 in PBS, pH 7.4 and incubated for an additional 30 minutes. Antigen coated microspheres were washed twice with 200 μL 0.02% Tween 20, 1 mg/mL BSA in PBS, pH 7.4 (PBSTB). and stored in PBSTB at approximately 3,000,000 microspheres/mL as determined by hemacytometer count.

Rubella assay: Rubella antigen loaded microspheres were used to examine several parameters of the assay in a single analyte format prior to the performance of multiple analyte assays. 10 µL (30,000 microspheres) of Rubella antigen coated beads were reacted with 10 µL of a 1:10 dilution of four different Rubella calibrator sera (Consolidated Technologies, Inc., Oak Brook, IL) and the mixture incubated for 1 hour. These sera were defined using a standard assay for the anti-Rubella IgG activity by the manufacturer of the calibrators. The units were defined as International Units/ mL or IU/mL. Beads were washed in PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 40 µL of a 10 µg/mL solution of Bodipy-labeled anti-human IgG. This mixture was incubated for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry. Negative controls included the microspheres with no serum treated with the Bodipy-labeled antibodies. In addition one calibrator serum containing 70 IU/mL of anti-Rubella IgG activity was titrated in a single analyte assay.

Multiple analyte assay for IgG and IgM activities: Equivalent amounts of each of the 5 antigen loaded microspheres was mixed to produce a ToRCH bead mixture. 10 µL (30,000 microspheres) of the mixture was reacted with 10 µL of a 1:400 dilution of ToRCH control or calibrator sera and incubated for 1 hour. The positive and negative ToRCH control sera did not have defined units of activity. The ToRCH calibrator, however, did have defined levels of anti-ToRCH IgG activities as defined by INX and DiaMedix diagnostic instruments. These values were provided by the manufacturer for the lot of calibrator purchased. Beads were washed in PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 20 µL of a 40 µg/mL solution of Bodipy-labeled anti-human IgG or IgM. This mixture was incubated for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry. Negative controls included the microspheres with no serum treatment and the microspheres treated with the ToRCH negative control serum. Both negative controls were developed with the Bodipy-labeled antibodies.

Results

Rubella assay: Rubella coated DFM were reacted with 4 human serum calibrators containing known levels of IgG antibodies specific for Rubella virions defined by International units

(IU/mL). The beads were washed and developed with goat anti-human IgG-Bodipy. Results are shown in **Table 9** and **Figure 24**. Increasing units of anti-Rubella activity were reflected in the Mean Intensity of Fluorescence (MIF) of F_m (green channel). Luminex Units (LU) were defined as the MIF of F_m for each data point minus the MIF of F_m for the negative control (no serum) multiplied by 0.1, and are included in **Table 9**.

Rubella calibrator titration: The human serum calibrator containing 70 IU/mL of anti-Rubella IgG was serially diluted in PBSTB and assayed with the Rubella coated microspheres and Bodipy-labeled anti-human IgG. Results shown in **Table 10** and **Figure 25** show that, as expected, the IgG antibodies specific for Rubella were titrated with dilution.

Multiple analyte ToRCH analysis for serum IgG and IgM: Each of the five distinct DFM coated with ToRCH antigens plus one DFM coated with human serum albumin (Miles, Inc., West Haven, CT) were mixed in equal volumes and 10 µL (30,000 microspheres) of the mixture reacted for 1 hour with triplicate, 20 µL aliquots of a 1:400 dilution of the ToRCH controls as well as the Low ToRCH calibrator. The calibrator from Blackhawk Systems, Inc. contained known levels of each pathogen specific antibody as measured on other diagnostic machines. After washing, one set of triplicates was developed with Bodipy-labeled anti-human IgG and another set with Bodipy-labeled anti-human IgM. Numerical results are shown in **Tables 11** and **12**. Results are presented graphically in **Figures 26A** and **26B**. Included in the figures are standard deviation bars for the triplicate measurements. For both IgG and IgM measurements, the ToRCH negative control serum (A96601, tubes #1-3) produced MIF of F_m similar to the negative control with no serum (tubes #10-12). The ToRCH positive control serum (A96602, tubes #4-6) demonstrated significant IgG activity to all five pathogens. Conversely, the positive control serum had only slight IgM based reactivity to the five pathogens. The known levels of anti-ToRCH IgG reactivities for the ToRCH Calibrator (A96500, tubes #7-9) were compared to the Luminex units of each IgG activity as determined by the multiple analyte analysis. Luminex units were defined by subtracting the negative control serum average MIF of F_m from the average MIF of F_m for each antigen and multiplying by 0.1. The levels of the ToRCH calibrator

were defined by the manufacturer as a factor of activity for each antigen above the limit of detection for that antigen on a specific diagnostic machine. These results are listed in **Table 13**.

A demonstrative ToRCH assay has been developed to simultaneously assay for serum IgG or IgM specific for the five ToRCH pathogens in a single tube. Results of the assay indicate that it is specific for each pathogen and is as sensitive as currently available instrument based assays. The multiple analyte format provides a uniquely powerful technology for rapid and less expensive serum testing for seroconversion to ToRCH pathogens as well as other infectious agents diagnosed in this manner.

TABLE 9: Anti-Rubella calibration curve

Calibrator IU/mL	MIF of *Fm*	LU/mL
360	1419	133
225	1004	91
70	458	37
40	376	28
0	92	0

TABLE 10: Anti-Rubella calibrator titration

70 IU/mL Calibrator Reciprocal of Dilution	MIF of *Fm*
1	4510
4	2554
8	1597
16	954
32	652
64	392
128	209
256	121
512	99
0	59

TABLE 11: IgG ToRCH assay

Tube #	Calibrator (1:400)	MIF of *Fm* in Triplicate					
		Toxo.	Rubella	CMV	HSV I	HSV II	HSA
1	A96601	21	9	12	16	17	22
2	A96601	22	8	10	17	15	26
3	A96601	25	9	11	14	17	22
4	A96602	647	1786	956	1223	664	78
5	A96602	590	1677	967	1511	719	81
6	A96602	620	1670	922	1348	611	72
7	A96500	103	38	50	128	64	27
8	A96500	95	43	48	127	58	43
9	A96500	87	41	49	127	56	29
10	No Serum	21	7	15	18	13	22
11	No Serum	23	8	11	15	19	19
12	No Serum	21	5	12	12	16	23
	Calibrator	Average MIF of *Fm*					
	(1:400)	Toxo.	Rubella	CMV	HSV I	HSV II	HSA
	A96601	23	9	11	16	16	23
	A96602	619	1711	948	1361	665	77
	A96500	95	41	49	127	59	33
	No Serum	22	7	13	15	16	21

TABLE 12: IgM ToRCH assay

Tube #	Calibrator (1:400)	MIF of *Fm* in Triplicate					
		Toxo.	Rubella	CMV	HSV I	HSV II	HSA
1	A96601	40	10	17	17	21	16
2	A96601	36	9	15	15	20	17
3	A96601	39	9	19	18	23	20
4	A96602	68	109	53	80	52	27
5	A96602	69	112	56	84	52	23
6	A96602	77	133	81	91	64	60
7	A96500	66	15	27	34	26	20
8	A96500	67	15	23	37	29	22
9	A96500	66	15	28	31	31	29
10	No Serum	40	9	18	17	21	21
11	No Serum	36	8	20	17	19	16
12	No Serum	38	8	14	17	19	18
	Calibrator (1:400)	Average MIF of *Fm*					
		Toxo.	Rubella	CMV	HSV I	HSV II	HSA
	A96601	38	9	17	17	21	18
	A96602	71	118	63	85	56	37
	A96500	66	15	26	34	29	24
	No Serum	38	8	17	17	20	18

TABLE 13: Comparison of known levels of anti-ToRCH IgG for the ToRCH calibrator from Blackhawk BioSystems with Luminex Units

	T.gondii	Rubella	CMV	HSV 1	HSV 2
Diagnostic Machine used	INX	INX	INX	Diamedix	DiaMedix
Factor above Limit of Detection	1.7 x	2.7 x	1.7 x	2.5 x	1.1 x
Units of Activity	11.3 IU/mL	26.9 IU/mL	24.5 IU/mL	50 EU/mL	22 EU/mL
Luminex Units/mL	7.2 LU/mL	3.2 LU/mL	3.8 LU/mL	11.1 LU/mL	4.3 LU/mL

Simultaneous Assay of Dog Sera for Allergic IgE and Allergen-Specific IgG

This example demonstrates the screening of serum for IgE antibodies specific for allergens. Screening of serum for IgE antibodies specific to allergens is a viable option for allergy testing as compared with skin sensitivity testing. The instant invention provides for a format that can assay for either IgG or IgE responses to numerous allergens at the same time in the same tube with the same sample and is therefore a uniquely powerful method of screening.

An allergy assay has been developed including 16 grass allergens in a multiple analyte, simultaneous format. A panel of 16 grass allergens were attached to 16 Differentially Fluorescent Microspheres (DFM) with one grass allergen being coated onto one unique member of the bead set. The allergen bead set was treated with diluted dog serum for 1 hour and treated with a solution of either Goat anti-Dog IgE or goat anti-dog IgG-FITC for an additional hour. For the IgE assay, beads were washed clear of this antibody and the bead set treated with an affinity purified rabbit anti-goat IgG-FITC antibody as probe.

Results demonstrate a uniquely powerful method of serum screening for allergies that provides a true multiple analyte format, as well as sensitivity and specificity.

Allergen conjugation to microspheres: Sixteen DFM (5.5 μM carboxylate) were conjugated separately to 16 soluble grass allergens (provided by Dr. Bill Mandy, BioMedical Services, Austin, TX) with a two-step EDC coupling method (Pierce Chemicals, Rockford, IL) using sulfo-NHS to stabilize the amino-reactive intermediate. All grass allergens were diluted 1:100 into PBS, pH 7.4. 20 μL (8.4 million microspheres) of each bead type was activated for 20 minutes in a total volume of 100 μL containing 500 μg of EDC and Sulfo-NHS in 50 mM sodium phosphate buffer, pH 7.0. Microspheres were washed twice with 100 μL PBS, pH 7.4 using centrifugation at 13,400 x g for 30 seconds to harvest the microspheres. Activated, washed beads were suspended in 50 μL of diluted allergen. After 2 hours, the microspheres were blocked by addition of 50 μL of 0.2 M glycine, 0.02% Tween 20 in PBS, pH 7.4 and incubated for an additional 30 minutes. Protein coated microspheres were washed twice with

100 µL 0.02% Tween 20, 1 mg/mL BSA in PBS, pH 7.4 (PBSTB). and stored in PBSTB at approximately 3,000,000 microspheres/mL as determined by hemacytometer count.

Multiplexed K-9 grass allergen IgE assay: Equivalent amounts of each of the 16 grass allergen loaded microspheres was mixed to produce a bead mixture. 20 µL (60,000 microspheres) of the mixture was reacted with 60 µL of a 1:3 dilution of dog serum in PBSTB and the mixture incubated for 1 hour. Beads were washed in 200 µL PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 40 µL of a 50 µg/mL solution of anti-dog IgE (provided by Dr. Bill Mandy, BioMedical Services, Austin, TX). After incubation for 1 hour, beads were washed in 200 µL PBSTB by centrifugation at 13,400 x g for 30 seconds. Beads were then treated with 40 µL of rabbit anti-goat IgG-FITC (Sigma, St. Louis, MO) at 20 µg/mL. After one hour the bead mixture was diluted to 300 µL in PBSTB and assayed using flow cytometry. Negative controls included the microspheres with dog serum, without the goat anti-dog IgE and with the rabbit anti-goat IgG-FITC. A negative control of the bead set with no dog serum was also included. Allergen specific dog IgE was determined by subtraction of the mean intensity of fluorescence (MIF) of the green channel (F_m) for the negative controls for each grass allergen from the MIF of F_m for the tubes including the goat anti-dog IgE.

Simultaneous K-9 grass allergen IgG assay: Equivalent amounts of each of the 16 grass allergen loaded microspheres was mixed to produce a bead mixture. 20 µL (8.4 million microspheres) of the mixture was reacted with 20 µL of a 1:10 dilution of dog serum in PBSTB and the mixture incubated for 1 hour. Beads were washed in 200 µL PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 25 µL of a 50 µg/mL solution of goat anti-dog IgG-FITC. After one hour the bead mixture was diluted to 300 µL in PBSTB and assayed using flow cytometry. Negative controls included the microspheres with no dog serum and with the goat anti-dog IgG-FITC. Allergen specific dog IgG was determined by subtraction of the mean intensity of fluorescence (MIF) of the green channel (F_m) for the negative control for each grass allergen from the MIF of F_m for the tubes including dog serum.

Results

Multiple analyte dog anti-grass allergen IgG assay: Grass allergen coated DFM were reacted with 6 dog sera provided by BioMedical Services, Austin, TX that had been characterized by ELISA for anti-grass allergen IgE. The IgG response to these grass allergens was not measured by BioMedical Services. The beads were washed and developed with goat anti-dog IgG-FITC. Results are shown in **Figure 27**. The MIF of F_m in the absence of dog serum was subtracted from the MIF of F_m for each bead with each dog serum. Two dogs, A96324 and A96326 demonstrated relatively high IgG reactivity to most of the grass allergens. Two dogs, A96325 and A96317 demonstrated relatively medium IgG reactivity to most of the grass allergens. Two dogs, A96319 and A96323 demonstrated relatively low IgG reactivity to most of the grass allergens.

Multiple analyte dog anti-grass allergen IgE assay: Grass allergen coated DFM were reacted with 6 dog sera provided by BioMedical Services, Austin, TX that had been characterized by ELISA for anti-grass allergen IgE. The beads were washed and treated with goat anti-dog IgE for 1 hour. The assay was developed with rabbit anti-goat IgG-FITC. Results are shown in **Figure 28**. The MIF of F_m in the absence of dog serum was subtracted from the MIF of F_m for each bead with each dog serum. Two dogs, A96325 and A96326 demonstrated relatively low reactivity to most of the grass allergens with the exception of Wheat grass and several others for A96326. These results agree with the ELISA results provided by BioMedical Services. A96325 was negative for 11 grass allergens (only ones tested) and A96326 was negative for the same 11 grass allergens except for a "Borderline" result in ELISA against a mixture of Wheat and Quack grass (due to the non-multiplexed format of ELISA assays, allergens are often mixed to increase the throughput levels). The other four dog sera demonstrated medium to high IgE responses to several of the grass allergens. Although agreement between ELISA and flow cytometry assay results was not absolute, the two assays followed the same trends. Dogs with IgE reactivity to grass allergens were detected by both assays.

Comparison of multiple analyte IgG and IgE results: The IgG and IgE anti-grass allergen response to each of the 16 allergens was compared by graphing. **Figures 29-34** demonstrate that

there was no correlation between IgG and IgE response to grass allergens in the six dogs. Some dogs were low responders for both IgE and IgG, some were reactive with both immunoglobulin subclasses, and some demonstrated IgE reactivity in a low background of IgG specific for the grass allergens. Examination of the IgG reactivity in a serum could identify those sera in which the IgE reactivity could be masked by the high IgG reactivity.

A demonstrative assay for serum IgG or IgE activity to 16 grass allergens has been developed that allows simultaneous assay of all 16 allergens at the same time in the same tube using the same sample. Results with 6 dog sera suggested that IgE anti-grass allergen activity as determined by ELISA was in general agreement with results provided using flow cytometry. In addition, the ease of determination for IgG anti-grass allergen activity in the six dogs was demonstrated.

A Simultaneous Immunometric Assay For Human Chorionic Gonadotropin and Alpha-Fetoprotein

This example illustrates the determination of multiple analyte levels in a liquid sample simultaneously by immunometric or capture-sandwich assay. The use of capture-sandwich assays to accurately determine analyte levels in liquid solutions is a commonly used format for many analyte assays. The technique is especially useful for those analytes present in low quantities as the first step serves to capture and thus concentrate the analyte. The uniqueness of this assay is the multiple analyte format allowing the simultaneous determination of two distinct serum proteins at the same time in the same tube from the same serum sample.

Human chorionic gonadotropin (hCG), a gonadotropic hormone secreted by the placenta, is the primary hormonal marker utilized for pregnancy testing. hCG is elevated both in urine and serum during pregnancy. Alpha fetoprotein (AFP) is the fetal cell equivalent to human serum albumin. AFP is elevated in pregnancy and in certain types of malignancies. Many clinical fertility or pregnancy test panels include immunometric assays for these two serum proteins. Immunometric or capture-sandwich assays for hCG and AFP were developed separately and then combined in a multiple analyte format.

The hCG assay was developed by examining several antibody pairs for their ability to capture and quantitate hCG levels in solution. First, a monoclonal antibody was coupled through carbodiimide linkage to a carboxylate substituted Differentially Fluorescent Microsphere (DFM). Next, a polyclonal, affinity purified antibody was Bodipy-labeled and used to probe DFM captured hormone. Once this assay was adjusted to include physiological sensitive ranges, the process was repeated for AFP. Cross-reactivity of the two assays was examined to demonstrate that the two assays would not interfere. The assays were then performed simultaneously. Commercially available serum calibrators were used to demonstrate that both assays were sensitive in clinically relevant ranges and an unknown was include to demonstrate how the two assays work simultaneously.

Antibody labeling: The two affinity purified polyclonal anti-hCG (AB633) and anti-AFP (M20077) antibodies (Chemicon, Inc., Temecula, CA and Medix Division, Genzyme, San Carlos, CA) were labeled with Bodipy FL-CASE (Molecular Probes, Inc., Eugene, OR) using methods described by the manufacturer of the Bodipy succinymidyl ester. The resulting Bodipy-labeled antibodies were stored in PBS containing 1 mg/mL BSA as stabilizer.

Antibody conjugation to microspheres: Monoclonal anti-hCG (MAB602) and anti-AFP (S10473) capture antibodies were conjugated to microspheres with a two-step EDC coupling method (Pierce Chemicals, Rockford, IL) using sulfo-NHS to stabilize the amino-reactive intermediate. 20 µL (8.4 million microspheres) of each DFM was activated for 20 minutes in a total volume of 100 µL containing 500 µg of EDC and Sulfo-NHS in 50 mM sodium phosphate buffer, pH 7.0. Microspheres were washed twice with 200 µL PBS, pH 7.4 using centrifugation at 13,400 x g for 30 seconds to harvest the microspheres. Washed, activated beads were suspended in 50 µL of a 0.05 mg/mL solution of antibody in PBS, pH 7.4. After 2 hours, microspheres were blocked by addition of 50 µL of 0.5 mg/mL BSA, 0.02% Tween 20 in PBS, pH 7.4 and incubated for an additional 30 minutes. Protein coated microspheres were washed twice with 200 µL 0.02% Tween 20, 1 mg/mL BSA in PBS, pH 7.4 (PBSTB) and stored in PBSTB at approximately 3,000,000 microspheres/mL. Microsphere concentrations were determined using a hemacytometer.

Antibody pairs analysis of hormone capture assay: Capture assay antibody pairs were screened by coupling potential capture antibodies to microspheres and assaying them using all potential combinations of capture antibody-bead/ Bodipy-labeled probe antibody. Assays were performed using 10 μL of capture antibody microspheres (approximately 30,000) plus 20 μL of antigen solution at 10 μg/mL in PBSTB for a 1 hour incubation. Beads were washed in PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 20 μL of a 25 μg/mL solution of Bodipy-labeled probe antibody. This mixture was incubated for 1 hour, diluted to 300 μL in PBSTB and assayed using flow cytometry.

Antigen titration assay: Once an antibody pair was chosen for use, the pair was analyzed for sensitivity and limit of detection by titration of antigen. Assays were performed using 10 μL of capture antibody microspheres plus 20 μL of antigen dilutions in PBSTB for a 1 hour incubation. Beads were washed in 200 μL PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 20 μL of a 25 μg/mL solution of Bodipy-labeled probe antibody. This mixture was incubated for 1 hour, diluted to 300 μL in PBSTB and assayed using flow cytometry.

Cross-reactivity analysis: To examine the possibility of cross-reactivity, 10 μL of MAB 602 anti-hCG capture beads (5,000 microspheres) were treated with 20 μL dilutions of hCG or AFP. After 1 hour the beads were washed in 200 μL PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 20 μL of either Bodipy-labeled anti-hCG or Bodipy-labeled anti-AFP at 25 μg/mL. Conversely, 10 μL of S-10473 anti-AFP capture beads (5,000 microspheres) were treated with 20 μL dilutions of hCG or AFP. After 1 hour, beads were washed in 200 μL PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 20 μL of either Bodipy-labeled anti-hCG or Bodipy-labeled anti-AFP at 25 μg/mL. Mixtures were incubated for 1 hour, diluted to 300 μL in PBSTB and assayed using flow cytometry.

Washed vs. no-wash assay format: An AFP/hCG capture antibody bead mixture was made by mixing equal amounts of the two bead types. In duplicate, 10 μL of this bead mixture (10,000 microspheres) was mixed with 20 μL dilutions of AFP/ hCG and incubated for 1 hour. One set

of beads were washed in PBSTB by centrifugation at 13,400 x g for 30 seconds and suspended in 20 µL of a mixture of Bodipy-labeled anti-hCG and anti-AFP both at 25 µg/mL. This mixture was incubated for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry. The second set of beads were treated directly with 20 µL of a mixture of Bodipy-labeled anti-hCG and anti-AFP both at 25 µg/mL. This mixture representing a homogenous (no-wash) assay was also incubated for 1 hour, diluted to 300 µL in PBSTB and assayed using flow cytometry.

Multiple analyte assay: Once the AFP and hCG antibody pairs were shown not to cross-react and were adjusted to provide clinically relevant ranges of sensitivity in a homogenous assay, the assays were performed simultaneously using commercially available serum calibrators as the source of AFP and hCG antigens. Equivalent amounts of each of the two capture antibody loaded microspheres was mixed to produce an AFP/hCG capture mixture. In triplicate, 10 µL of this bead mixture (5,000 of each microsphere) was mixed with 20 µL of three serum calibrators (high, medium and low) containing known levels of AFP and hCG and incubated for 1 hour. Mixtures were treated directly with 20 µL of a blend of Bodipy-labeled anti-hCG and anti-AFP both at 25 µg/mL. Mixtures were incubated for 1 hour, diluted to 300 µL in PBSTB and assayed by flow cytometry.

Results

Antibody pair analysis for hCG capture assay: For hCG antibody pair analysis, five capture antibody/microspheres were prepared and the identical five antibodies were Bodipy-labeled to serve as probes. Three of the antibodies were specific for the alpha sub-unit of hCG and two for the beta sub-unit. The three anti-alpha sub-unit antibody/microspheres were assayed for utility with the two Bodipy-labeled anti-beta hCG antibodies. Conversely, the two anti-beta sub-unit antibody/microspheres were assayed for utility with the three Bodipy-labeled anti-alpha hCG antibodies. Results of this screen are shown in **Table 14** and **Figure 35**. The 12 combinations of antibodies are shown with (odd numbers) and without (even numbers) hCG at 20 µg/mL. It is apparent that the first two antibody pairs, #1 and #3 demonstrated the highest mean intensity of fluorescence (MIF) of the F_m (green channel). Further examination of these two pairs led to

the decision to chose the #3 pair of MAB 602 for capture antibody and AB633-Bodipy as probe antibody for the hCG capture/sandwich assay.

Antigen titration: The MAB 602/AB633 anti-hCG capture system was assayed by hCG titration to determine if the level of sensitivity required for clinical assay was achievable. A limit of detection of at least 1 ng hCG/mL was the target as this was the level of hCG in the low serum calibrator to be used later in this project. The results of this antigen titration is shown in **Table 15** and **Figure 36**. The limit of detection was between 20 and 200 pg/mL. This revealed that the MAB602/AB633 anti-hCG antibody pair was sufficiently sensitive for hCG analysis. Included in this analysis was MIF of F_m measurements from counting of 100 or 1000 microspheres. Results were similar. A similar analysis of antibody pairs and antigen titration for AFP identified an AFP pair that could be further developed.

Cross-reactivity assay: The MAB 602/AB633 anti-hCG capture system and S-10473/M20077 anti-AFP capture system were examined for cross reactivity by assaying each capture bead with each antigen and Bodipy-labeled antibody. Results are shown in **Table 16** and **Figures 37A** and **37B**. No significant cross-reactivity between the hCG and AFP capture systems was detected.

No-wash vs. washed assay format: The hCG and AFP assays were performed simultaneously and examined for the limit of quantitation or dynamic range in both a washed format and no wash or homogenous format. Result of these antigen titrations are shown in **Table 17** and **Figures 38A** and **38B**. Results indicated that the homogenous format provided sufficient dynamic range for the purposes of clinical relevance.

Multiple analyte hCG/AFP assay: The two assays were performed simultaneously using serum calibrators of known hCG and AFP levels to generate a standard curve. For each standard curve one serum of unknown hCG and AFP level was included to demonstrate how the assay would determine the level of hCG and AFP in the serum.

The Randox Tri-level calibrators consisted of three serum samples with high, medium and low levels of hCG and AFP documented in mU or U/mL for hCG and AFP respectively. These calibrators are used in at least 12 different diagnostic instruments including those of Abbott (Abbott Park, IL), bioMerieux (St. Louis, MO), Ciba Corning (Medfield, MA), Diagnostics Products (Los Angeles, CA), Kodak (Rochester, NY), Syva (San Jose, CA), Tosoh (Atlanta, GA) and Wallac (Gaithersburg, MD). Literature with the Randox Tri-Level control listed the units of each known analyte as defined by each diagnostic machine. We calculated the average of the hCG mU/mL and AFP U/mL for the three calibrators. In the case of the hCG, the low and medium calibrators contained 22.8 and 26.4 mU/mL which were extremely close considering the distance to the high calibrator (436 mU/mL). For this reason, we included a 1:2 dilution of the high range calibrator into hCG/AFP certified negative serum to produce a fourth level serum calibrator termed Level 3D. Calibrator 3D was only used in construction of the hCG standard curve so each of the assays was effectively defined by three point calibration.

Table 18 shows the results of this multiple analyte assay. The assay was performed in triplicate and the average MIF of F_m computed for graphing. Coefficients of variation (CV) for the triplicates were consistently less than 10% are shown. Also included in the table are the number of microspheres correctly identified by the flow cytometry out of the 400 counted per tube. Of the 400 beads counted the expected ratio of MAB 602 containing 60/40 beads to S-10473 containing 40/60 beads was 1:1. Therefore of the 200 beads expected, this was the number of beads correctly identified and used to compute the MIF of F_m for that data point.

Figures 39A and **39B** graphically represent the data of **Table 18**. For both hCG and AFP a plot of the MIF vs. the log of antigen concentration produced a line that was best fit using a third level polynomial equation. The fit for the hCG curve provided an R^2 of 1.0 and for AFP an R^2 of 0.9999 was achieved. Using the polynomial equation, the concentration of the unknowns was computed. Results of these analyses are seen in **Table 18**. The unknown serum contained 218.55 ± 6.56 mU/mL of hCG and 39.59 ± 1.19 U/mL of AFP.

A demonstrative immunometric assay for hCG and AFP in serum has been developed. Assays were first developed as single analyte or single bead assays, and optimized with regards to sensitivity, limit of quantitation and cross-reactivity. The assays were then combined to quantitatively determine multiple analyte levels in a liquid solution in the same tube from the same sample at the same time. Results, using commercially available calibrator sera, has proven that this invention is effective for this type of quantitative assay.

TABLE 14

Sample	Description	hCG Conc. (μg/mL)	MIF of *Fm*
1	A1-Beads + B1 Ab-BD with hCG	20.0	8790
2	A1-Beads + B1 Ab-BD with no hCG	0.0	108
3	A2-Beads + B1 Ab-BD with hCG	20.0	9441
4	A2-Beads + B1 Ab-BD with no hCG	0.0	163
5	A3-Beads + B1 Ab-BD with hCG	20.0	3150
6	A3-Beads + B1 Ab-BD with no hCG	0.0	2984
7	A1-Beads + B2 Ab-BD with hCG	20.0	2287
8	A1-Beads + B2 Ab-BD with no hCG	0.0	37
9	A2-Beads + B2 Ab-BD with hCG	20.0	1232
10	A2-Beads + B2 Ab-BD with no hCG	0.0	42
11	A3-Beads + B2 Ab-BD with hCG	20.0	566
12	A3-Beads + B2 Ab-BD with no hCG	0.0	560
13	B1-Beads + A1 Ab-BD with hCG	20.0	70
14	B1-Beads + A1 Ab-BD with no hCG	0.0	23
15	B2-Beads + A1 Ab-BD with hCG	20.0	346
16	B2-Beads + A1 Ab-BD with no hCG	0.0	20
17	B1-Beads + A2 Ab-BD with hCG	20.0	107
18	B1-Beads + A2 Ab-BD with no hCG	0.0	33
19	B2-Beads + A2 Ab-BD with hCG	20.0	886
20	B2-Beads + A2 Ab-BD with no hCG	0.0	56
21	B1-Beads + A3 Ab-BD with hCG	20.0	105
22	B1-Beads + A3 Ab-BD with no hCG	0.0	196
23	B2-Beads + A3 Ab-BD with hCG	20.0	143
24	B2-Beads + A3 Ab-BD with no hCG	0.0	609

TABLE 15

Sample	hCG Conc. (ng/mL)	MIF of *Fm* (1000 Beads)	MIF of *Fm* (1000 Beads)
1	20000	9337	9222
2	2000	9286	9392
3	200	9233	9400
4	20	8497	8664
5	2	1286	1382
6	0.2	258	254
7	0.02	120	147
8	0.002	122	121
9	0.0002	122	149
10	0	128	111

TABLE 16A

MAB602 BEADS -		**Anti-hCG**			
Samp.	**Antigen** ng/mL	hCG anti-hCG	AFP anti-hCG	hCG anti-AFP	AFP anti-AFP
1	1000.0	792	53	47	52
2	100.0	761	47	48	48
3	10.0	530	47	47	48
4	1.0	104	47	48	48
5	0.1	55	52	49	48
6	0.0	48	71	72	48

TABLE 16B

M20077 BEADS -		**Anti-AFP**			
Samp.	**Antigen** ng/mL	hCG anti-hCG	AFP anti-hCG	hCG anti-AFP	AFP anti-AFP
1	1000.0	99	57	78	348
2	100.0	54	75	44	356
3	10.0	44	44	45	103
4	1.0	51	50	44	98
5	0.1	42	75	49	44
6	0.0	43	61	45	45

- 82 -

TABLE 17

	AFP				hCG		
Sample No.	AFP ng/mL	No Wash	Washed		hCG ng/mL	No Wash	Washed
1	1000	379	1481		2000	491	3194
2	500	643	1376		1000	770	3158
3	250	956	1205		500	1198	3342
4	125	1063	1052		250	1521	2755
5	62	980	814		125	2068	2949
6	31	639	612		62	2417	3200
7	16	359	347		31	2514	3183
8	8	190	205		16	2440	2528
9	4	94	108		8	1761	1955
10	2	51	59		4	1122	1300
11	1	33	35		2	650	547
12	0.5	24	25		1	330	359
13	0.25	17	24		0.5	166	175
14	0	15	13		0	15	18

TABLE 18

		hCG capture system					AFP capture system				
Tube No.	Descript.	hCG mU/mL	AFP U/mL	MIF of FL1	MIF AVG	MIF CV%	Beads IDed	MIF of FL1	MIF AVG	MIF CV%	Beads IDed
1	Level 1			23			114	74			98
2	Level 1	22.8	10.7	27	25.67	7%	130	83	78.00	5%	80
3	Level 1			27			140	77			72
4	Level 2			32			93	351			58
5	Level 2	26.4	53.8	34	31.67	6%	85	365	362.67	2%	61
6	Level 2			29			94	372			65
7	Level 3D			268			92	535			56
8	Level 3D	218	111.5	276	271.67	1%	101	562	552.00	2%	69
9	Level 3D			271			96	559			61
10	Level 3			631			106	1109			46
11	Level 3	436	223	601	624.00	3%	99	994	1061.00	5%	38
12	Level 3			640			97	1080			40
13	Negative			8			99	11			104
14	Negative	10	2	7	7.33	6%	111	13	12.33	8%	106
15	Negative			7			119	13			95
16	Unknown			270			140	268			67
17	Unknown	218.55	39.59	264	272.33	3%	141	274	276.00	3%	81
18	Unknown			283			113	286			84

Multiplexed Beadset Standard Curve Using an Inhibition Assay

This example provides a demonstration of the measurement of ligand-ligate reactions using a multiplexed beadset standard curve. Commonly for ligand-ligate reactions quantitation, known amounts of the ligand or ligate are introduced to the reaction leading to the production of a standard curve. Values for unknown samples are compared to the standard curve and quantified. The true multiple assay capability of this invention allows for an additional type of standard to be utilized. A multiplexed beadset standard curve for measuring analyte concentration is created by using several Differentially Fluorescent Microspheres (DFM) coated with either 1) different amounts of ligand (antigen), or 2) different amounts of ligate (antibody), or 3) different ligates possessing different avidities for the ligand (different monoclonal antibodies). We have demonstrated an example of the first type of multiple analyte standard curve by developing a competitive inhibition assay for human IgG.

Four DFM were coated with human IgG at four different concentrations. When probed with goat anti-human IgG-Bodipy the Mean Intensity of Fluorescence (MIF) of F_m (green channel) for each bead subset was different. The MIF of F_m correlated with the amount of hIgG used to coat the beads in each subset. If soluble hIgG was mixed with the reaction in a competitive manner the MIF of F_m was reduced for each bead as less of the probe antibody was bound to the beads. In a normal standard curve, the signal (MIF of F_m) is plotted against the concentration of the inhibitor. For the multiplexed beadset standard curve, the slope of the MIF of F_m across the beads within a subset is plotted against the concentration of inhibitor. Comparison of the two types of standard curves revealed them to be of equivalent value for prediction of an unknown amount of inhibitor.

Human IgG conjugation to microspheres: Four DFM (5.5 μM carboxylate, Bangs Laboratories, Inc., Carmel, IN, dyed by Emerald Diagnostics, Inc., Eugene, OR) were conjugated separately to 4 different concentrations of hIgG (Cappel Division, Organon Teknika, Durham, NC) with a two-step EDC coupling method (Pierce Chemicals, Rockford, IL) using sulfo-NHS to stabilize the amino-reactive intermediate. 20 μL (8.4 million microspheres) of each bead type was activated for 20 minutes in a total volume of 100 μL containing 500 μg of EDC and Sulfo-NHS

in 50 mM sodium phosphate buffer, pH 7.0. The microspheres were washed twice with 200 µL PBS, pH 7.4 using centrifugation at 13,400 x g for 30 seconds to harvest the microspheres. Activated, washed beads were suspended in 50 µL of hIgG at 50, 10, 5, and 1 µg/mL in PBS, pH 7.4. After 2 hours, the microspheres were blocked by addition of 50 µL of 0.2 M glycine, 0.02% Tween 20 in PBS, pH 7.4 and incubated for an additional 30 minutes. Protein coated microspheres were washed twice with 200 µL 0.02% Tween 20, 1 mg/mL BSA in PBS, pH 7.4 (PBSTB). and stored in PBSTB at approximately 3,000,000 microspheres/mL as determined by hemacytometer count.

Antibody labeling: Goat anti-human IgG (Cappel Division, Organon Teknika, Durham, NC) was labeled with Bodipy FL-CASE (Molecular Probes, Inc., Eugene, OR) using methods described by the manufacturer of the Bodipy succinymidyl ester. The resulting Bodipy-labeled antibody was stored in PBS containing 1 mg/mL BSA as stabilizer.

Multiplexed beadset standard curve: Equivalent amounts of each of the 4 differentially loaded IgG microspheres was mixed to produce a bead mixture. 10 µL of the goat anti-hIgG-Bodipy at 25 µg/mL in PBSTB was mixed with 10 µL of a dilution of hIgG in PBSTB. Immediately 10 µL (30,000 microspheres) of the bead mixture was added and the mixture incubated at room temperature for 30 minutes. Beads were diluted to 300 µL in PBSTB and assayed using flow cytometry. A negative control included the microspheres with the goat anti-hIgG-Bodipy with no inhibitor (hIgG). Each bead subset was assigned the value of a consecutive integer (i.e. the bead subset coupled with the lowest concentration of IgG was given a value of 1, the next highest concentration was given a value of 2, etcetera) and those numbers plotted against the MIF of each bead subset at each concentration of inhibitor tested. The slopes (designated here as inter-bead subset slopes) were computed using linear regression analysis. The inter-subset slopes were then plotted against the concentration of inhibitor using a logarithmic scale for the concentration of inhibitor. Results were plotted as the slope of the MIF of F_m across the bead set against the log of hIgG concentration. Curve fitting was performed using a power function trendline and the R^2 correlation was reported. For a perfect fit, R^2=1.0.

Common standard curve: Using the data from the assay described above, a common standard curve was constructed to compare results with the multiple analyte standard curve. Data from the DFM coated at 50 µg/mL hIgG was utilized to produce a five-point standard curve by plotting the MIF of F_m against the log of hIgG concentration. Curve fitting was performed using a power function trendline and the R^2 correlation was reported.

Results

Multiplexed beadset standard curve for a competitive inhibition assay: Four differentially loaded IgG microspheres were utilized in a multiple beadset competitive inhibition assay for hIgG at five different concentrations of soluble inhibitor (hIgG). Results of the assay are shown in **Table 19**. The inhibition pattern on each bead is plotted in **Figure 40**. The inter-bead subset slopes are plotted against the log concentration of inhibitor in **Figure 41**. A Power Trendline in Excel was used to produce the R^2 of 0.9933.

Common standard curve using one bead of the multiple analyte assay: Data from the 50 µg/mL hIgG bead was utilized to produce a five-point standard curve by plotting the MIF of F_m against the log of hIgG concentration. Results are shown in **Figure 42**. Curve fitting was performed using a Power function trendline and $R^2 = 0.9942$.

A novel type of standard curve for ligand-ligate measurement was demonstrated. Results suggested that the multiplexed beadset standard curve was of similar utility as the common multi-point standard curve in quantitation of unknown samples. Advantages of the multiplexed beadset standard curve include the inclusion of the standard curve microspheres in each point of a multiplexed beadset assay, and the extension of an assay's dynamic range. This may be achieved by increasing the concentration range of ligand or ligate on the microspheres or by increasing the range of avidities for ligand on a series of microspheres.

TABLE 19

Samp	Inhibitor Conc (μg/mL)	Bead 1 1.0 μg/mL IgG	Bead 2 5 μg/mL IgG	Bead 3 10 μg/mL IgG	Bead 4 50 μg/mL IgG	SLOPE
1	100	14	77	108	288	85.3
2	50	21	100	162	428	128.3
3	25	40	166	267	844	251.3
4	12.5	110	463	747	1467	435.5
5	6.25	257	1226	1629	2316	658
6	0	134	793	1432	2217	688.8

Nucleic Acid Measurement

The power and sensitivity of PCR has prompted its application to a wide variety of analytical problems in which detection of DNA or RNA sequences is required. One major difficulty with the PCR technique is the cumbersome nature of the methods of measuring the reaction's products - amplified DNA.

A major advance in this area is here. That advance employs a flow cytometric bead-based hybridization assay which permits the extremely rapid and accurate detection of genetic sequences of interest. In a preferred embodiment of that invention, a bead to which a nucleic acid segment of interest has been coupled is provided. A PCR product of interest (or any other DNA or cDNA segment) is detected by virtue of its ability to competitively inhibit hybridization between the nucleic acid segment on the bead and a complementary fluorescent nucleic acid probe. The method is so sensitive and precise as to allow the detection of single point mutations in the PCR product or nucleic acid of interest. Although that method in itself provides a pivotal advance in the art of analyzing PCR reaction products, the further discovery of methods of multiplexing such an analysis, compounds the method's power and versatility to allow simultaneously analysis of a number of nucleic acid products or a number of sequences within a single product in a single sample.

The multiplexed DNA analysis method described here can be applied to detect any PCR product or other DNA of interest for specific polymorphisms or mutations or for levels of expression, e.g. mRNA. With the multiplexed techniques provided by the instant invention, individuals can be screened for the presence of histocompatibility alleles associated with susceptibility to diseases, mutations associated with genetic diseases, autoimmune diseases, or mutations of oncogenes associated with neoplasia or risk of neoplasia. The analysis of DNA sequences occurs generally as follows:

1. A beadset containing subsets of beads coupled to nucleic acid sequences of interest is prepared by coupling a unique synthetic or purified DNA sequence to the beads within each subset.
2. Fluorescent probes complementary to the DNA coupled to each bead subset are prepared. Methods known in the art, e.g., as described in U.S. Patent No. 5,403,711, issued April 4, 1995 and incorporated herein by reference, or other methods may be used to fluorescently label the DNA. Since each probe will bind optimally only to its complementary DNA-containing subset, under the conditions of the assay, the fluorescent probes may be added to the subsets before or after the subsets are pooled, and before or after addition of the DNA test sample(s) of interest.
3. Tissue, fluid or other material to be analyzed is obtained, and DNA is purified and/or amplified with PCR as necessary to generate the DNA products to be tested.
4. The DNA samples of interest are then mixed with the pooled beadset under suitable conditions to allow competitive hybridization between the fluorescent probes and the DNA of interest.
5. The beadset is then analyzed by flow cytometry to determine the reactivity of each bead subset with the DNA sample(s). If the test sample contains a DNA sequence complementary to the DNA of a given bead subset then that subset will exhibit a decreased F_m value relative to the F_m value of beads to which a control DNA has been added. A computer executed method in accordance with the current invention can determine the subset from which each bead is derived, and therefore, the identity of the DNA sequence on the bead and any change in F_m.

Detection of Foreign DNA

The methods of the present invention find wide utility in the detection of foreign DNA's in, for example, diagnostic assays. Although the DNA segment to be analyzed can be any DNA sequence, in accordance with this embodiment the selected segment will be a DNA segment of a pathogenic organism such as, but not limited to, bacterial, viral, fungal, mycoplasmal, rickettsial, chlamydial, or protozoal pathogens. The procedure has particular value in detecting infection by pathogens that are latent in the host, found in small amounts, do not induce inflammatory or immune responses, or are difficult or cumbersome to cultivate in the laboratory. The multiplexed DNA detection method of the present invention is likely to find particular utility as a diagnostic assay for analysis of a sample from a patient having clinical symptoms known to be caused by a variety of organisms using a beadset designed to detect DNAs from the variety of organisms known to cause such symptoms to determine which of such organisms is responsible for the symptoms. DNA would be extracted from tissue, fluid or other sources and analyzed as described above.

Analysis of Genetic Polymorphisms

The invention may also be used to measure a variety of genetic polymorphisms in a target DNA of interest. For example, there are several genes in the MHC and many are polymorphic. There are at least two applications in which determination of the alleles at each position of the MHC is of critical importance. The first is the determination of haplotype for transplantation, and the second is determination of haplotype as indicator of susceptibility to disease. See Gross et al., "The Major Histocompatibility Complex-Specific Prolongation of Murine Skin and Cardiac Allograft Survival After In Vivo Depletion of $V\beta^+$ T Cells," J. Exp. Med., 177, 35-44 (1993). The MHC complex contains two kinds of polymorphic molecules, Class I genes, HLA A, B and D which have 41, 61 and 18 known alleles and Class 10 genes, HLA-DRI,3,4,5 HLA-DQAI and BI HLA-DP, DPA1, DPB1, also with many alleles. Each human can have up to 6 co-dominant Class I genes and 12 co-dominant Class 10 genes.

In the case of transplantation, the closer the match between the donor and recipient the greater the chance of transplant acceptance. A multiplexed assay in accordance with the

invention may be employed to perform tissue typing quickly and accurately to identify suitable matches for transplantation.

In the situation of disease association, it has been found that individuals bearing certain alleles are more prone to some diseases than the remainder of the population. The frequency of alleles of the MHC genes is not equal, and sets of alleles are frequently found (linkage disequilibrium) so that the identification of the exact set of alleles associated with many diseases is feasible. As one example, insulin-dependent diabetes mellitus (IDDM) is associated with certain HLA-DQ alleles. The number of alleles of DQ in the population is modest and genetic typing by PCR amplification and hybridization with allele specific probes has been shown to be practical. See Saiki et al., "Genetic Analysis of Amplified DNA with Immobilized Sequence-Specific Oligonucleotide Probes," Proc. Natl. Acad. Sci. U.S.A., 86, 6230- 6234 (1989).

For an assay of MHC in accordance with the invention, DNA is obtained from blood or other extractable source, and amplified with primers specific for the MHC genes under analysis, for example, HLA-DQA. For a full genotyping of the MHC, several samples of DNA would be amplified with different sets of primers to accommodate the large number of loci and the high degree of polymorphism. The PCR products are then screened for specific alleles using beadsets and fluorescent probes as described above.

Mutation Analysis of Selected Genes: Screening Procedures

There are several methodologies for determining and comparing DNA sequences in order to detect mutations which are associated with disease or neoplasia. When adapted to a bead-based, multiplexed format in accordance with the current invention, hybridization analysis allows for the rapid screening of multiple genetic loci for multiple wild type and mutant sequences.

In a preferred embodiment of the invention, a given genetic locus, or multiple loci, can be simultaneously screened for the presence of wild type or mutant sequences. In the same analysis, multiple known mutations can be distinguished from each other and from the wild type

sequence and uncharacterized mutations. In addition, the homozygosity or heterozygosity of known sequences can be determined.

A general approach for detecting a DNA mutation in accordance with this aspect of the invention is as follows. In a first step, a suitable probe for detecting a mutation of interest is selected. In an illustrative embodiment, selected oligonucleotides, representing wild-type and mutant sequences, from a region of a gene known to contain a mutation are prepared. Such oligonucleotides are coupled to microspheres by techniques known in the art, (e.g., carbodiimide coupling, or other means) to produce individual aliquots of beads having known oligonucleotides coupled thereto. The oligonucleotides must be a sufficient length to allow specific hybridization in the assay, e.g., generally between about 10 and 50 nucleotides, more preferably between about 20 and 30 nucleotides in length. In a preferred embodiment, a saturating amount of the oligonucleotide is bound to the bead. Fluorescent oligonucleotides, complementary to all or part of the sequences attached to each bead, are also prepared.

Next, PCR primers are selected to amplify that region of the test DNA corresponding to the selected probe, which are then used to amplify the particular region of DNA in the sample that contains the sequence corresponding to the oligonucleotide coupled to the beads. Either double stranded or single stranded PCR techniques may be used. If double stranded product is produced, the amplified PCR product is made single stranded by heating to a sufficient temperature to and for a sufficient time to denature the DNA (e.g., for about 1 to about 5 minutes at about 90-95°C in 2.3X SSC hybridization buffer). The mixture is cooled, and the beads are added and incubated with the PCR product under conditions suitable to allow hybridization to occur between the oligonucleotide on the beads and the PCR product (e.g., at room temperature for about 10 minutes). The fluorescent DNA probe may then be added and the entire mixture incubated under hybridization conditions suitable to allow competitive hybridization to occur (e.g., 5 minutes at 65°C, then cooling to room temperature over a period of several hours in 2.3X SSC buffer). As those of skill in the art will recognize, the concentrations of the PCR product and fluorescent probe to be used may vary and may be adjusted to optimize the reaction.

In general, the concentrations of PCR product and fluorescent probe to be used are adjusted so as to optimize the detectable loss of fluorescence resulting from competitive inhibition without sacrificing the ability of the assay to discriminate between perfect complementarity and one or more nucleotide mismatches. In an exemplary assay, the concentration of PCR product complementary to the oligonucleotide bound to the beads may be on the order of 1 to 10 times the molar concentration of fluorescent probe used. The fluorescent probe should preferably be added in an amount sufficient to achieve slightly less than saturation of the complementary oligonucleotide on the beads in order to obtain maximum sensitivity for competitive inhibition.

In a multiplexed assay employing the above principles, beadsets are separately prepared, pooled, and the bead-based hybridization analysis performed. In order to screen a given locus for mutations, beadset subsets are prepared such that subset 1 is coupled to a DNA segment identical to the wild type sequence, subset 2 is coupled to a DNA segment identical to a known mutation 1 (which may represent a single or multiple point mutations, deletions or insertions), subset 3 is coupled to a DNA segment identical to a second known mutation 2, and so on. The subsets are then mixed to create a pooled beadset.

When a nucleic acid sample is analyzed with such a beadset, only the bead subsets containing sequences identical to the test sample will show a large decrease in fluorescence (F_m). Bead subsets containing unrelated or greatly disparate sequences will show little or no decrease in fluorescence (F_m) and bead subsets containing very closely related sequences, such as point mutants, will show an intermediate decrease in fluorescence (F_m). Thus, a large decrease in the F_m of only subset 1 would indicate homozygous wild-type; a large decrease in the F_m of both subset 1 and subset 2 would indicate heterozygous wild-type / mutant 1 and so on. If the test sample is less inhibitory than the perfectly complementary sequence for any of the known sequences represented by the subsets then a new uncharacterized mutation is indicated. The test sample could then be sequenced to characterize the new mutation, and this sequence information used to construct a new subset for the beadset to detect the newly discovered mutation.

The present invention has wide-spread advantages for detection of any of a number of nucleic acid sequences of interest in the genomic DNA of an individual or organism and has the advantages of being both rapid and extremely accurate in effecting the detection of such mutations. The invention will find wide applicability in diagnosis of a number of genetically associated disorders as well as in other applications where identification of genetic mutations may be important. Exemplary diseases include without limitation, diseases such as cystic fibrosis, generalized myotonia and myotonia congenita, hyperkalemic periodic paralysis, hereditary ovalocytosis, hereditary spherocytosis and glucose malabsorption; which are associated with mutations in the genes encoding ion transporters; multiple endocrine neoplasia, which is associated with mutations in the MEN2a, b, and MEN1 genes; familial medullary thyroid carcinoma, and Hirschsprung's disease, which are associated with mutations in the ret proto-oncogene; familial hypercholesterolemia, which is associated with mutations in the LDL receptor gene; neurofibromatosis and tuberous sclerosis, which are associated with mutations in the NF1 gene, and NF type 2 gene; breast and ovarian cancer, which are associated with mutations in the BRCA1, BRCA2, BRCA3 genes; familial adenomatous polyposis, which is associated with mutations in the APC gene; severe combined immunodeficiency, which is associated with mutations in the adenosine deaminase gene; xeroderma pigmentosum, which is associated with mutations in the XPAC gene; Cockayne's syndrome, which is associated with mutations in the ERCC6 excision repair gene; fragile X, which is associated with mutations in the fmr1 gene; Duchenne's muscular dystrophy, which is associated with mutations in the Duchenne's muscular dystrophy gene; myotonic dystrophy, which is associated with mutations in the myotonic dystrophy protein kinase gene; bulbar muscular dystrophy, which is associated with mutations in the androgen receptor genes; Huntington's disease, which is associated with mutations in the Huntington's gene; Peutz-jegher's syndrome; Lesch-Nyhan syndrome, which is associated with mutations in the HPRT gene; Tay-Sachs disease, which is associated with mutations in the HEXA gene; congenital adrenal hyperplasia, which is associated with mutations in the steroid 21-hydroxylase gene; primary hypertension, which is associated with mutations in the angiotensin gene; hereditary non-polyposis, which is associated with mutations in the hNMLH1 gene; colorectal carcinoma, which is associated with mutations in the 2 mismatch repair genes; colorectal cancer, which is associated with mutations in the APC gene; forms of Alzheimer's disease which have been associated with the

apolipoprotein E gene, retinoblastoma, which is associated with mutations in the Rb gene; Li-Fraumeni syndrome, which is associated with mutations in the p53 gene; various malignancies and diseases that are associated with translocations: e.g., in the bcr/abl, bcl-2 gene; chromosomes 11 to 14 and chromosomes 15 to 17 transpositions. The references at the end of the specification which are expressly incorporated herein by reference describe genetic mutations associated with certain diseases which may be tested for in accordance with the invention as well as sequences provided in GENBANK, the contents of which are also expressly incorporated herein by reference.

Double Stranded Experiment

For the purposes of illustration, the two complementary strands of a double-stranded DNA segment are referred to as strand "A" and strand "B". Either strand may be designated "A" or "B". The wild-type "B" strand oligo (ras codon 12) having the oligonucleotide sequence 5'-GCCTACGCCACCAGCTCCAACTAC-3' (SEQ ID NO. 3) was coupled to 3.0 micrometers (μm) latex microspheres (manufactured by Interfacial Dynamics, Portland, OR) by carbodiimide coupling. Double stranded competitor was prepared by combining equal amounts of both the "A" and "B" strands of either the wild-type or mutant version of the oligo, mutant "B" strand having the sequence 5'-GCCTACGCCACAAGCTCCAACTAC-3' (SEQ ID NO. 4) (ras codon 12) in 5X SSC buffer. Annealing was accomplished by heating the mixture to 65°C for five minutes, then cooling slowly to room temperature. Competitive hybridization was accomplished by combining approximately 40 picomoles of the bead-attached oligo (wild-type "B" strand) with the indicated amounts of double stranded competitor in 2.3X SSC buffer at approximately 25°C. Finally, 100 picomoles of the fluorescinated oligo (wild-type "A" strand) was added to the reaction mixture. This mixture was incubated for two hours at room temperature, and then diluted with 300 μl of saline pH 7.3, and analyzed on the "FACSCAN" (manufactured by Becton-Dickinson Immunocytometry Systems, San Jose, CA). The results are shown in **Table 20** below and in **Figures 43a** through **43c**.

TABLE 20: Double-Stranded Experimental Results Using Wild-Type "B" Oligonucleotide

Double Stranded Competitor (picomole)	Percent Inhibition (%)		Fold Competition
	Wild-Type	Mutant	Wild-Type/Mutant
10	20	9	2.2
100	35	12	2.9
1000	56	17	3.3

These results clearly show that the DNA containing the single point mutation ("Mutant") was a detectably less effective inhibitor of hybridization between the DNA on the beads and the fluorescent oligonucleotide probe at all concentrations of competitor tested.

Single Stranded Experiment

The wild-type "B" strand oligo (ras codon 12) was coupled to 3.0 μm latex microspheres (manufactured by Interfacial Dynamics) by carbodiimide coupling. Competitive hybridization was accomplished by combining approximately 40 picomoles of the bead-attached oligo with 100 picomoles of the fluorescinated oligo (wild-type "A" strand) in 2.3X SSC buffer. Finally, the indicated amounts of single stranded competitor (either mutant or wild-type) were added to two separate aliquots of the reaction mixture. These aliquots were incubated for two hours at room temperature, and then diluted with 300 μl of saline pH 7.3. and analyzed on the FACSCAN flow cytometer. The result of these experiments are set forth in **Table 21** below and in **Figures 44a** and **44b**.

TABLE 21: Single-Stranded Experimental Results

Single Stranded Competitor (picomole)	Percent Inhibition (%)		Fold Competition
	Wild-Type	Mutant	Wild-Type/Mutant
100 "A" Strand	14	6	2.4
1000 "A" Strand	25	11	2.3

These results clearly show that the DNA containing the single point mutation ("Mutant") was a detectably less effective inhibitor of hybridization between the DNA on the beads and the florescent oligonucleotide probe at all concentrations of competitor tested.

Resequencing analysis of PCR products using multiplexed analysis.

This example demonstrates the ability of flow cytometry to perform resequencing analysis of PCR products. As a model system, PCR products were derived from the DQA1 gene, in the region of the gene which determines the major alleles of DQA1. The DQA1 gene represents the DNA coding sequence for the alpha chain of the DQ molecule. DQ is classified as a class II histocompatibility locus and is expressed in allelic form in all humans. Most individuals are heterozygous for DQA , i.e., they express two different DQA alleles. The determination of DQA alleles is used in identity testing for paternity and forensic purposes.

Seventeen alleles of DQA1 have been defined by DNA sequencing; however, eight major alleles account for the large majority of the population. These alleles are determined by fourteen unique DNA sequences contained within four regions of the DQA1 gene; all four regions are contained within a 227 base pair PCR product derived from human genomic DNA.

Flow cytometry was used to determine the presence or absence of all fourteen DNA sequences in a PCR product simultaneously in a single reaction tube, thereby allowing determination of the DQA alleles expressed in a given sample. The system is based on competitive hybridization between the PCR product and complementary oligonucleotide pairs representing each of the fourteen unique DNA sequences. One strand of each oligonucleotide

pair is coupled to a unique subset of microspheres and the complementary strand is labeled with a green emitting fluorophore. After coupling, the fourteen unique microsphere subsets were pooled to produce the mixed bead set. After addition of the fourteen fluorescent oligonucleotides and the PCR product to the beadset, the mixture is hybridized and then analyzed by flow cytometry. The ability of the PCR product to inhibit the hybridization of the complementary fluorescent oligonucleotides to their respective microsphere subsets is used to determine the DNA sequences, and thus, the allele(s) present in the PCR product.

Microspheres: Carboxylate-modified latex (CML) microspheres of 5.5 micron mean diameter were obtained from Bangs Laboratories, Inc. (Carmel, IN). The microspheres were differentially dyed with varying concentrations of two fluorescent dyes with orange and red emission spectra to produce fourteen unique microsphere subsets.

Oligonucleotides: Fourteen oligonucleotide pairs (complementary strands designated "A" and "B") corresponding to allelic sequences within the DQA1 gene (**Table 22**) were synthesized by Oligos, Etc. (Wilsonville, OR). using standard automated techniques. Each eighteen-base oligonucleotide was substituted at the 5' end with an amino-terminal linker during synthesis.

Oligonucleotide coupling to microspheres: The "B" strand of each oligonucleotide pair was coupled to a unique subset of CML microspheres using carbodiimide chemistry. Briefly, 0.1 mL of a 1 mM solution of oligonucleotide in 0.1 M MES (2-[*N*-morpholino]ethanesulfonic acid), pH 4.5 was added to 1.0 mL of microspheres (1% solids) in 0.1 M MES, pH 4.5. To this mixture, 0.05 mL of a 10 mg/mL solution of EDC (1-ethyl-3-[3-dimethylaminopropyl]-carbodiimide hydrochloride) was added and mixed vigorously. The mixture was incubated for 30 minutes at room temperature, followed by another addition of EDC, mixing, and incubation as above. Following the second incubation period, the microspheres were pelleted by centrifugation and resuspended in 0.4 mL of 0.1 M MES, pH 4.5 and stored at 4°C.

Oligonucleotide labeling: The "A" strand of each oligonucleotide pair was fluorescently labeled with Bodipy FL-X (6-((4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-propionyl)

amino)hexanoic acid, succinimidyl ester) (Molecular Probes, Inc. , Eugene, OR). Briefly, a 400 μL solution containing 20 μM oligonucleotide in 0.1 M sodium bicarbonate and 5% DMSO, pH 8.2 was reacted with 30 μL Bodipy FL-X (10 mg/mL in DMSO) for 16-18 hours at room temperature. The mixture was desalted on a PD10 column equilibrated in TE (10 mM TrisHCl, 1 mM ethylenediamine tetraacetic acid (EDTA), pH 8.0) to remove excess unreacted dye and stored at 4°C.

DNA extraction: Tissue sample (template) DNA was purified using the QIAmp Blood Kit (Qiagen, Chatsworth, CA) for DNA purification. Briefly, 1 x 10^7 tissue culture cells or 200 μL whole blood is lysed with Qiagen protease and Buffer AL. The lysate is incubated at 70°C for 10 minutes followed by addition of 210 μL ethanol. The mixture is applied to a QIAmp spin column and centrifuged at 8,000 x g for 1 minute. The filtrate is discarded, 500 μL Buffer AW is added to the column and the centrifugation is repeated; this step is repeated. The filtrates are discarded and the DNA is eluted into a new tube by addition of 200 μL Buffer AE, incubation at room temperature for 1 minute, followed by centrifugation as above.

Polymerase chain reaction (PCR: PCR primers designated DQA AMP-A (5'-ATGGTGTAAA CTTGTACCAGT-3', SEQ ID NO. 5) and DQA AMP-B (5'-TTGGTAGCAG CGGTAGAGTTG-3', SEQ ID NO. 6) (World Health Organization, 1994) were synthesized by Oligos, Etc. (Wilsonville, OR) using standard automated techniques. PCR was performed with reagents (PCR buffer, dNTPs, $MgCl_2$, and TAQ DNA polymerase) from Life Technologies, Inc.(Gaithersburg, MD). The reaction mixture (50 μL) contained 1 μM of each primer, 200 nM dNTPs, 3 mM $MgCl_2$, 4 - 10 μg/mL DNA template, and 2.5 units TAQ DNA polymerase in PCR buffer. The PCR reaction was performed on an Idaho Technologies thermal cycler (Idaho Falls, ID) using and initial step at 94°C for 45 sec, and 32 cycles of 94°C for 30 sec, 48°C for 60 sec, and 72°C for 60 sec followed by a final hold at 72°C for 7 minutes. Production of the product was verified by agarose electrophoresis and was quantified by size exclusion chromatography on a Superdex 75 (10/30) column (Pharmacia, Piscataway, NJ). The PCR product was used without purification.

Competitive hybridization analysis: The hybridization reaction was performed in a total volume of 40 μL, containing approximately 8,000 of each bead subset for a total of approximately 110,000 microspheres, 50 nM of each fluorescent oligonucleotide, and 10 - 200 nM PCR product, as competitor, in hybridization buffer (3 M trimethyl ammonium chloride, 0.15% sodium dodecyl sulfate, 3 mM EDTA, and 75 mM TrisHCl, pH 8.0). Briefly, the beadset mixture, in hybridization buffer, was equilibrated at 55°C. The mixture of fluorescent oligonucleotides and PCR product was denatured in a boiling water bath for 10 minutes followed by quick-chilling on ice for 2 minutes. The microspheres were added, mixed well, and the entire reaction was allowed to hybridize for 30 minutes at 55°C. Following hybridization, the mixture was diluted to 250 μL using hybridization buffer and analyzed by flow cytometry.

Results

Microspheres for multiple analytes: **Figure 45** illustrates the classification, using orange and red fluorescence, of the fourteen microsphere subsets used in the DQA1 analysis. Each distinct microsphere subset bears one of the fourteen unique oligonucleotide capture probes on its surface. The level of green fluorescence associated with each subset, after hybridization with the fluorescent oligonucleotide probes, is also determined simultaneously, and measures the reactivity of the fluorescent oligonucleotides (and therefore, the reactivity of the PCR product) with each unique oligonucleotide sequence.

Titration of fluorescent oligonucleotide: To optimize the system for detection of PCR products, fluorescent oligonucleotide was titered in the presence or absence of PCR competitor. **Figure 46** illustrates the hybridization of increasing concentrations of fluorescent oligonucleotide "5503A" to microspheres coupled to oligonucleotide "5503B" in the presence or absence of a 200 nM concentration of double-stranded 0301 PCR product which contains the 5503 sequence. In the absence of competitor, the level of "5503A" which hybridizes to the microspheres, detected as FL1, increases in a linear manner and reaches saturation at approximately 10 nM. In the presence of competitor, the binding curve is shifted to the right indicating inhibition of "5503A" hybridization.

Concentration dependence of inhibition and detection of point mutations: **Figure 47** illustrates the inhibition of fluorescent oligonucleotide hybridization by varying concentrations of complementary and point mutant competitors in the presence of a fixed concentration of fluorescent oligonucleotide. The solid lines show the inhibition of hybridization to bead "3401B" induced by competitors 3401 (u) or 3402 (n). The dashed lines show inhibition of hybridization to bead "3402B" induced by competitors 3401(s) or 3402 (l). Even at the lowest competitor concentration (10 nM), there is approximately a two-fold difference between the reactivity of the identical sequence versus the point mutant.

Specificity of the multiple analyte assay: The specificity of the reaction of each DNA competitor sequence with the multiplexed microsphere subsets is illustrated in **Table 23** and **Figure 48**, using double-stranded oligonucleotide competitors. The pattern of reactivity is consistent with the homology of the different oligonucleotides with identical sequences showing maximal reactivity, closely related sequences showing less reactivity, and unrelated sequences showing little or no reactivity.

Allele-specific reactivity patterns: In order to establish the reactivity patterns of the DQA1 alleles in a model system, simulated alleles were prepared by mixing the oligonucleotides representing the DNA sequences that would be present within a single PCR product for a given allele. **Figure 49** illustrates the typing of four simulated alleles of DQA1. By comparison to the allele reactivity chart shown in **Table 24**, it can be seen that each of the simulated alleles types correctly.

Typing of homozygous genomic DNA: To verify the ability of flow cytometry to correctly type PCR products prepared from genomic DNA, samples of DNA of known, homozygous DQA1 type were obtained from the UCLA Tissue Typing Laboratory, Los Angeles, CA. After PCR amplification, these samples were typed using flow cytometry; the results are shown in **Figure 50**. By comparison to the allele reactivity chart **(Table 24)**, it can be seen that the system correctly types these samples.

Typing of heterozygous genomic DNA: To determine the ability of multiplexed flow analysis to accurately type heterozygous DQA1 haplotypes, twenty-five samples of known heterozygous DQA1 type were obtained from the UCLA Tissue Typing Laboratory, Los Angeles, CA. The samples of homozygous DNA used above were added to the panel and all of the samples were coded and typed in a blinded study. The data from this study are presented in **Table 25**. The last column of **Table 25** entitled "Type" indicates whether the haplotype indicated by UCLA and the Luminex analysis agreed. In 34 of 35 samples, the haplotypes reported by both laboratories agreed; sample number 19 was not typed by the UCLA laboratory, but typed clearly as an 0501/0201 heterozygote in the Luminex analysis. Thus, the multiplexed analysis is capable of typing the DQA1 haplotypes with at least 97% accuracy.

These studies have demonstrated that flow cytometry can rapidly and accurately perform resequencing analysis of PCR products. The model system used here required the analysis of fourteen DNA sequences to determine eight different DQA1 alleles. Flow cytometry was able to perform this analysis in a true simultaneous format, using a single sample of a single PCR product in a single reaction tube. The entire analysis, including setup, hybridization, flow analysis, and data collection and analysis can be accomplished within an hour after PCR amplification of the DNA sample. Thus, it is possible to perform tissue typing or other genetic analysis in less than three hours after obtaining a sample of blood, tissue, etcetera, including the time required for extraction of DNA and PCR amplification.

TABLE 22: DQA1 DNA Sequences

Name	Sequence "A" Strand	Sequence "B" Strand	Allele Specificities
DQA2501	TGGCCAGTACACCCATGA (SEQ ID NO. 7)	TCATGGGTGTACTGGCCA (SEQ ID NO. 8)	**0101, 0401, 0501**
DQA2502	TGGCCAGTTCACCCATGA (SEQ ID NO. 9)	TCATGGGTGAACTGGCCA (SEQ ID NO. 10)	**0103, 0201, 0601**
DQA2503	TGGGCAGTACAGCCATGA (SEQ ID NO. 11)	TCATGGCTGTACTGCCCA (SEQ ID NO. 12)	**0301**
DQA3401	GAGATGAGGAGTTCTACG (SEQ ID NO. 13)	CGTAGAACTCCTCATCTC (SEQ ID NO. 14)	**0101, 0104**
DQA3402	GAGATGAGCAGTTCTACG (SEQ ID NO. 15)	CGTAGAACTGCTCATCTC (SEQ ID NO. 16)	**0102, 0103, 0501**
DQA3403	GAGACGAGCAGTTCTACG (SEQ ID NO. 17)	CGTAGAACTGCTCGTCTC (SEQ ID NO. 18)	**0401, 0601**
DQA3404	GAGACGAGGAGTTCTATG (SEQ ID NO. 19)	CATAGAACTCCTCGTCTC (SEQ ID NO. 20)	**0201, 0301**
DQA4101N	ACCTGGAGAGGAAGGAGA (SEQ ID NO. 21)	TCTCCTTCCTCTCCAGGT (SEQ ID NO. 22)	**0101, 0102, 0201, 0301**
DQA4102	ACCTGGAGAAGAAGGAGA (SEQ ID NO. 23)	TCTCCTTCTTCTCCAGGT (SEQ ID NO. 24)	**0103**
DQA4103	ACCTGGGGAGGAAGGAGA (SEQ ID NO. 25)	TCTCCTTCCTCCCCAGGT (SEQ ID NO. 26)	**0401, 0501, 0601**
DQA5501N	TCAGCAAATTTGGAGGTT (SEQ ID NO. 27)	AACCTCCAAATTTGCTGA (SEQ ID NO. 28)	**0101, 0102, 0103**
DQA5502N	TCCACAGACTTAGATTTG (SEQ ID NO. 29)	CAAATCTAAGTCTGTGGA (SEQ ID NO. 30)	**0201**
DQA5503	TCCGCAGATTTAGAAGAT (SEQ ID NO. 31)	ATCTTCTAAATCTGCGGA (SEQ ID NO. 32)	**0301**
DQA5504	TCAGACAATTTAGATTTG (SEQ ID NO. 33)	CAAATCTAAATTGTCTGA (SEQ ID NO. 34)	**0401, 0501, 0601**

Table 23: Specificity of Oligonucleotide Inhibition

	Bead #													
	2501	2502	2503	3401	3402	3403	3404	4101	4102	4103	5501	5502	5503	5504
Oligo														
none	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
2501	**64%**	10%	5%	-1%	-6%	4%	1%	4%	1%	3%	5%	-9%	2%	3%
2502	19%	**77%**	-7%	-2%	0%	1%	-3%	-3%	-1%	-4%	-4%	-4%	-5%	4%
2503	7%	1%	**85%**	1%	2%	4%	-3%	-2%	4%	-1%	1%	-1%	5%	6%
3401	-1%	-1%	-3%	**76%**	1%	6%	1%	3%	2%	2%	5%	-4%	2%	5%
3402	-1%	-8%	-12%	14%	**83%**	11%	-5%	-5%	-3%	-10%	0%	-7%	-4%	1%
3403	-2%	-3%	-1%	0%	7%	**73%**	-1%	2%	1%	-1%	5%	-9%	-4%	2%
3404	5%	5%	4%	6%	2%	8%	**62%**	10%	9%	6%	10%	0%	7%	9%
4101	6%	4%	6%	10%	5%	10%	6%	**79%**	18%	22%	31%	5%	12%	12%
4102	0%	-1%	-4%	0%	-3%	5%	0%	8%	**79%**	3%	7%	-1%	5%	4%
4103	-2%	11%	3%	5%	6%	7%	5%	4%	7%	**71%**	0%	8%	9%	6%
5501	-3%	3%	0%	2%	5%	5%	1%	-1%	6%	1%	**79%**	9%	6%	4%
5502	3%	5%	5%	5%	7%	3%	-1%	1%	2%	-4%	-7%	**86%**	4%	1%
5503	-5%	0%	-6%	1%	2%	0%	-2%	-8%	0%	-7%	-9%	13%	**80%**	-5%
5504	3%	8%	9%	5%	6%	5%	2%	2%	7%	6%	4%	13%	4%	**93%**

TABLE 24. Allele Reactivity Chart

Allele	Pattern	Sequence
0101	(1,1,1,1)	2501 3401 4101 5501
0102	(1,2,1,1)	2501 3402 4101 5501
0103	(2,2,2,1)	2502 3402 4102 5501
0201	(2,4,1,2)	2502 3404 4101 5502
0301	(3,4,1,3)	2503 3404 4101 5503
0401	(1,3,3,4)	2501 3403 4103 5504
0501	(1,2,3,4)	2501 3402 4103 5504
0601	(2,3,3,4)	2502 3403 4103 5504

Table 25. Blinded typing of Genomic DNA Samples for DQA1

	BEAD SUBSET														
	2501	2502	2503	3401	3402	3403	3404	4101	4102	4103	5501	5502	5503	5504	TYPE
1	4%	25%	35%	7%	69%	8%	58%	36%	61%	-1%	78%	2%	75%	-8%	Y
2	13%	-5%	-7%	63%	-4%	-2%	-3%	37%	2%	-9%	77%	0%	1%	-11%	Y
3	19%	-2%	40%	66%	-3%	0%	65%	51%	-2%	-2%	78%	-2%	76%	-14%	Y
4	22%	13%	18%	33%	68%	9%	1%	-3%	6%	33%	-14%	0%	-11%	36%	Y
5	38%	1%	-7%	68%	78%	6%	-1%	40%	-1%	36%	76%	-3%	-5%	48%	Y
6	-13%	-10%	11%	-10%	-10%	-7%	45%	21%	-4%	-12%	-24%	-9%	75%	-19%	Y
7	-7%	8%	-6%	-11%	-8%	19%	27%	7%	-5%	16%	-27%	31%	-6%	30%	Y
8	20%	-5%	-9%	1%	65%	53%	-2%	37%	-2%	30%	76%	-2%	-7%	43%	Y
9	32%	5%	-1%	7%	76%	5%	2%	57%	0%	-1%	84%	-3%	-11%	-25%	Y
10	32%	10%	60%	4%	72%	1%	71%	60%	-4%	7%	83%	-32%	76%	-40%	Y
11	27%	7%	5%	10%	70%	5%	2%	-14%	-5%	45%	-21%	-6%	-18%	44%	Y
12	29%	3%	57%	8%	71%	6%	67%	60%	0%	3%	84%	-4%	77%	-29%	Y
13	25%	8%	2%	66%	-4%	-8%	-8%	47%	-1%	-4%	83%	-12%	-22%	-36%	Y
14	16%	8%	33%	-3%	0%	16%	29%	24%	2%	26%	-25%	-7%	47%	17%	Y
15	7%	19%	8%	-6%	26%	0%	31%	26%	-5%	28%	-31%	36%	-21%	23%	Y
16	32%	9%	2%	18%	76%	8%	5%	-1%	-2%	48%	-4%	-6%	-17%	52%	Y
17	8%	0%	35%	6%	53%	4%	58%	45%	1%	48%	-13%	-1%	77%	41%	Y
18	1%	-2%	55%	3%	1%	1%	75%	54%	4%	6%	-14%	-3%	84%	-12%	Y
19	12%	18%	2%	5%	47%	8%	54%	37%	5%	46%	-11%	46%	1%	39%	?
20	63%	15%	6%	42%	87%	17%	20%	64%	10%	65%	87%	6%	5%	60%	Y
21	44%	62%	8%	23%	76%	13%	73%	70%	11%	20%	88%	64%	4%	1%	Y
22	-6%	26%	5%	4%	-3%	2%	55%	48%	6%	-1%	-7%	58%	1%	-11%	Y
23	56%	67%	13%	38%	87%	22%	24%	12%	75%	63%	89%	15%	15%	63%	Y
24	15%	25%	5%	60%	58%	7%	10%	42%	61%	-4%	90%	-2%	-3%	-17%	Y
25	47%	15%	12%	75%	15%	50%	17%	56%	4%	55%	83%	2%	-5%	51%	Y
26	30%	8%	56%	15%	71%	5%	67%	65%	2%	9%	85%	-1%	82%	-18%	Y
27	13%	10%	10%	7%	27%	21%	5%	1%	3%	47%	-16%	5%	-5%	51%	Y
28	23%	2%	0%	17%	23%	60%	14%	-2%	-1%	58%	-18%	-2%	-3%	59%	Y
29	24%	6%	10%	48%	46%	5%	10%	56%	5%	8%	86%	2%	-2%	-16%	Y
30	38%	12%	11%	73%	14%	48%	7%	55%	6%	55%	84%	1%	-7%	50%	Y
31	-1%	-1%	19%	0%	20%	-1%	26%	29%	-3%	31%	-28%	-4%	70%	31%	Y
32	57%	16%	6%	83%	81%	11%	6%	59%	7%	60%	86%	9%	-1%	53%	Y
33	-13%	17%	24%	2%	-1%	29%	47%	37%	-6%	50%	-11%	0%	80%	52%	Y
34	33%	7%	5%	19%	75%	6%	13%	54%	1%	2%	85%	24%	-3%	-3%	Y
35	-11%	19%	31%	10%	-14%	0%	70%	46%	-4%	-2%	-57%	44%	79%	-8%	Y

Measuring Enzymes with Bead-Based Assays

The invention may also be used in several formats for measurement of enzymes, enzyme inhibitors and other analytes. For example, bead subsets can be generated that are modified with selected fluorescent substrates which can be enzymatically cleaved from the bead, resulting in a loss of fluorescence (F_m). Enzymes that can be detected and measured using the invention include but are not restricted to, protease, glycosidase, nucleotidase, and oxidoreductase. Any enzyme that results in selected bond cleavage can be measured. A cartoon of the action of enzyme on a bead-bound enzyme is shown in **Figure 51a**. An enzyme that acts upon a bead-bound substrate so that the bead-bound substrate becomes fluorescent or loses fluorescence comprises an assay for the level of enzyme affecting such a change. **Figures 51b** and **51c** depict these situations. Alteration of the substrate could be an oxidation or reduction, alteration of a chemical bond such a hydrolysis or other alteration of the bead-bound substrate so that the fluorescence of the substrate is altered in intensity or spectrum.

Enzymes that act upon pro-enzymes (convertases) can be measured using a bead-bound substrate providing the reaction mixture contains the pro-enzyme and beads bearing a substrate that can be acted upon by the active form of the enzyme. (Providing the specificity of each activated enzyme is distinct, a multiplexed assay is achievable in which several pro-enzymes can be measured at the same time.) The sample is introduced into a mixture of pro-enzymes under reaction conditions. After a fixed time interval during which the convertase acts upon the pro-enzyme, the beadsets specific for each enzyme generated from each pro-enzyme are added and the newly generated activities measured in a subsequent time period which is terminated when the beadsets are analyzed by flow cytometry. Such a process for a single pro-enzyme to enzyme conversion is illustrated by the cartoon in **Figure 51d**.

The action of the enzyme can be measured in an indirect but facile manner using a bead bound substrate as depicted in **Figure 51e**. The action of the enzyme on the bead-bound substrate results in the formation or revelation of a ligate for a fluorescent ligand present in the reaction mixture. The bead bearing the modified substrate then becomes fluorescent by virtue of binding of the fluorescent ligand to the newly formed ligate. In practice, the enzyme(s) would

be added to the beadset under reactive conditions. After a defined time period during which the enzyme acts upon the bead bound substrate, the enzyme action would be stopped and the fluorescent ligands added and after a period for association of ligand with the beadsets, the mixture analyzed by flow cytometry. The fluorescent ligands could be of a single reactivity or multiple ligands employed, the critical specificity is that of the enzyme for the substrate.

The bead-bound substrate may be used to detect the activation of enzyme when the enzyme requires a cofactor for activity. Under this circumstance, the level of the cofactor becomes the limiting component of the reaction mixture and determination of the level of cofactor can be measured. Such a configuration is illustrated in **Figure 51f**. The reaction mixture contains the bead-bound substrate as well as the apo-enzyme. After introduction of the analyte (enzyme cofactor), the reaction mixture is held under reactive conditions for a fixed period of time followed by analysis of the beads by flow cytometry, the level of cofactor limits the level of enzyme activity. Providing the enzymes present require different cofactors and have action on different substrate-bearing beadsets, several cofactors could be measured in a single assay mixture.

In short, bead-borne substrates can be used as reagent as are soluble substrates for enzymes. However, because each type of bead bearing a unique substrate can be distinguished, a mixture of bead subsets can be used to measure several enzyme activities simultaneously in the same reaction mixture.

Fluids that can be analyzed using these techniques include plasma, serum, tears, mucus, saliva, urine, pleural fluid, spinal fluid and gastric fluid, sweat, semen, vaginal secretions, fluid from ulcers and other surface eruptions, blisters, and abscesses, and extracts of tissues including biopsies of normal, malignant, and suspect tissues. An assay according to this aspect of the invention proceeds as follows:

1. Beads containing reactive surface groups (one of the following: amino, aldehyde, acid chloride, amidine, phenolic hydroxyl, phenyl amine, carboxyl) are obtained that can be

discriminated on the basis of, for example, forward angle light scatter, C_1, right angle light scatter, C_2, and one of several wavelengths of fluorescence $C_3 \ldots C_n$ which are designated as orange and red fluorescence, for example, and comprise a number of subsets.

2. Subsets thus obtained are derivatized with a peptide (substrate) having a terminal fluorescent green group, for example fluorescein (F_m).
3. Unreacted surface groups and hydrophobic surface of the bead are blocked and the subsets are processed by a particle analyzer and sorter (FACSCAN) and a uniform population of particles are separated which have a low coefficient of variance for F_m. (e.g., 3%).
4. A fluid to be tested is diluted with an appropriate buffer and added to the beadset mixture to allow enzymes present in the sample to react with (cleave) their corresponding substrate on the surfaces of the beads.
5. After a defined period of time, the reaction is stopped and the entire mixture processed by a flow cytometer and results are determined.

The presence of an enzyme in the clinical sample is indicated by loss of fluorescence resulting from the cleavage of the fluorescent F_m substrate from the bead surface. Because the beads are analyzed in a very small volume (e.g., about 6 picoliters) as they are passed through the flow cytometer's laser beam, interference from free fluorescent molecules (cleaved substrate) will not significantly interfere with the assay. This obviates the requirement of washing of the beads prior to assay and simplifies the procedure significantly.

Time

Time measurement is an important feature of the analysis. The essence of the measurement of an enzyme activity is a change in substrate with time. The activity can be determined by setting a period of time during which the clinical sample is in contact with the beads using standard conditions of pH, ionic composition and temperature. Two alternative processes are available for determination of the bead-bound substrate with time, that is the time expired while the enzyme(s) is (are) acting on each beadset(s).

External Time

In this configuration, as each bead is measured by the flow cytometer, the time at which each measurement was obtained is recorded along with the bead's other measurements. Prior to the beginning of the assay, the baseline measurement is determined. Once the enzyme (clinical sample) is added to the bead mixture, the sample analysis begins. As the beads proceed through the instrument, the time data collected is used to determine the length of time that the bead has been exposed to the clinical sample. The F_m data collected over the period of the assay is used to determine the rate of change of substrate on the beads (kinetics) and thus the rate readily derived for each bead subset in the mixture exposed to the clinical sample.

Internal Time

Time can be determined and at the same time a quality control internally generated by including a "timer" bead subset that bears a substrate which is acted on by an enzyme that does not naturally occur in the clinical sample to be tested. The use of non-pathogenic microbial enzymes and substrates with human samples, for example, would suffice. The corresponding "timer" enzyme is added to the dilution buffer so that a known concentration of the "timer" enzyme is present in the buffer. The degree of action of the "timer" enzyme upon the beads in the "timer" subset can be measured as a function of the loss of fluorescence of the beads in the subset to ensure that proper reaction conditions are achieved. The level of fluorescence of the timer beads can thus be used as an internal standard and an estimation of time.

Determination of Enzyme Inhibitors or Regulators

In addition to direct assay of enzymes, an assay of this type may also be used to detect enzyme inhibitors or regulators. In accordance with this variation, samples being tested for inhibitors are added to the beadset followed by the corresponding enzymes. If inhibitors are present, the measured fluorescent (F_m) values will not be decreased to the same extent as a control containing no inhibitors. In accordance with **Figure 52**, in a similar manner, inhibitors of enzyme activators or binders of cofactors can be measured.

The present invention provides numerous advantages and overcomes many problems associated with prior art techniques of multiplexed diagnostic and genetic analysis apparatus and methods. It will be appreciated by those of ordinary skill having the benefit of this disclosure that numerous variations from the foregoing illustration will be possible without departing from the inventive concept described herein. Accordingly, it is the claims set forth below, and not merely the foregoing illustration, which are intended to define the exclusive rights claimed in this application program.

REFERENCES

1. Hum. Biol. 64: 167-174 (1992) Mutation in Cystic Fibrosis: a Review Spatial Distribution of the DF508. DeBraekeleer, M. and Daigeneault, J.;

2. Science 257: 797-800 (1992) [92358240] The skeletal muscle chloride channel in dominant and recessive human myotonia. M. C. Koch, K. Steinmeyer, C. Lorenz, K. Ricker, F. Wolf, M. Otto, B. Zoll, Lehmann-Horn, K. H. Grzeschik & T. J. Jentsch;

3. Neuron 12: 281-94 (1994) [94153549] Sodium channel mutations in paramyotonia congenita uncouple inactivation from activation. M. Chahine, A. L. George, M. Zhou, S. Ji, W. Sun, R. L. Barchi & R. Horn. Ann. Neurol. 33: 300-7 (1993) [93270429]; Sodium channel mutations in paramyotonia congenita and hyperkalemic periodic paralysis. L. J. Ptacek, L. Gouw, H. Kwiecinski, P. McManis, J. R. Mendell, R. J. Barohn, A. L. George, R. L. Barchi, M. Robertson & M. F. Leppert;

4. Ann. Neurol. 33: 300-7 (1993) [93270429] Sodium channel mutations in paramyotonia congenita and hyperkalemic periodic paralysis. L. J. Ptacek, L. Gouw, H. Kwiecinski, P. McManis, J. R. Mendell, R. J. Barohn, A. L. George, R. L. Barchi, M. Robertson & M. F. Leppert; Cell 67: 1021-7 (1991) [92069747] Identification of a mutation in the gene causing hyperkalemic periodic paralysis. L. J. Ptacek, A. L. George, R. C. Griggs, R. Tawil, R. G. Kallen, R. L. Barchi, M. Robertson & M. F. Leppert;

5. Nature 355: 836-8 (1992) [92168137] Defective anion transport activity of the abnormal band 3 in hereditary ovalocytic red blood cells. A. E. Schofield, D. M. Reardon & M. J. Tanner;

6. J. Clin. Invest. 93: 121-30 (1994) [94110314] Duplication of 10 nucleotides in the erythroid band 3 (AE1) gene in a kindred with hereditary spherocytosis and band 3 protein deficiency (band 3PRAGUE). P. Jarolim, H. L. Rubin, S. C. Liu, M. R. Cho, V. Brabec, L. H. Derick, S. J. Yi, S. T. Saad, S. Alper, C. Brugnara et al.;

7. Acta Physiol. Scand. Suppl. 607: 201-7 (1992) [93080072] The Na+/glucose cotransporter (SGLT1). E. M. Wright, E. Turk, K. Hager, L. Lescale-Matys, B. Hirayama, S. Supplisson & D. D. Loo. Nature 350: 354-6 (1991) [91179516]; Glucose/galactose malabsorption caused by a defect in the Na+/glucose cotransporter. E. Turk, B. Zabel, S. Mundlos, J. Dyer & E. M. Wright;

8. Nature 363: 458-60 (1993) [93275414] Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A.L. M. Mulligan, J. B. Kwok, C. S. Healey, M. J. Elsdon, C. Eng, E. Gardner, D.R. Love, S. E. Mole, J. K. Moore, L. Papi, et al.;

9. Nature 367: 375-6 (1994) [94159102] A mutation in the RET proto-oncogene associated with multiple endocrine neoplasia type 2B and sporadic medullary thyroid carcinoma [see comments] R. M. Hofstra, R. M. Landsvater, I. Ceccherini, R. P. Stulp, T. Stelwagen, Y. Luo, B. Pasini, J. W. Hoppener, H. K. van Amstel, G. Romeo, et al.;

10. Nature 367: 378-80 (1994) [94159104] Mutations of the RET proto-oncogene in Hirschsprung's disease [see comments] P. Edery, S. Lyonnet, L. M. Mulligan, A. Pelet, E. Dow, L. Abel, S. Holder, C. Nihoul-Fekete, B. A. Ponder & A. Munnich; Nature 367: 377-8 (1994) [94159103] Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease [see comments] G. Romeo, P. Ronchetto, Y. Luo, V. Barone, M. Seri, I. Ceccherini, B. Pasini, R. Bocciardi, M. Lerone, H. Kaariainen, et al.;

11. Hum. Mutat. 1: 445-66 (1992) [93250847] Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. H. H. Hobbs, M. S. Brown & J. L. Goldstein; Clin. Chem. 36: 900-3 (1990) [90291682] Use of polymerase chain reaction to detect heterozygous familial hypercholesterolemia. M. Keinanen, J. P. Ojala, E. Helve, K. Aalto-Setala, K. Kontula & P. T. Kovanen;

12. Hum. Genet. 93: 351-2 (1994) [94171244] Two CA/GT repeat polymorphisms in intron 27 of the human neurofibromatosis (NF1) gene. C. Lazaro, A. Gaona & X. Estivill; Am J

Hum Genet 54: 424-36 (1994) [94160989] Deletions spanning the neurofibromatosis 1 gene: identification and phenotype L. M. Kayes, W. Burke, V. M. Riccardi, R. Bennett, P. Ehrlich, A. Rubenstein & K. Stephens; Cell 75: 1305-15 (1993) [94094325] Identification and characterization of the tuberous sclerosis gene on chromosome 16. The European Chromosome 16 Tuberous Sclerosis Consortium;

13. Hum. Mol. Genet. 2: 1823-8 (1993) [94108432] Genetic analysis of the BRCA1 region in a large breast/ovarian family: refinement of the minimal region containing BRCA1. D. P. Kelsell, D. M. Black, D. T. Bishop & N. K. Spurr;

14. Hum. Mutat. 3: 12-8 (1994) [94163183] Exon eight APC mutations account for a disproportionate number of familial adenomatous polyposis families. D. J. Koorey, G. W. McCaughan, R. J. Trent & N. D. Gallagher; Hum. Mutat. 1: 467-73 (1992) [93250848] Screening for germ-line mutations in familial adenomatous polyposis patients: 61 new patients and a summary of 150 unrelated patients. H. Nagase, Y. Miyoshi, A. Horii, T. Aoki, G. M. Petersen, B. Vogelstein, E. Maher, M. Ogawa, M. Maruyama, J. Utsunomiya, et al.; Cell 66: 589-600 (1991) [91330306] Identification and characterization of the familial adenomatous polyposis coli gene. J. Groden, A. Thliveris, W. Samowitz, M. Carlson, L. Gelbert, H. Albertsen, G. Joslyn, J. Stevens, L. Spirio, M. Robertson, et al.;

15. Hum. Mol. Genet. 2: 1307-8 (1993) [94004878] A missense mutation in exon 4 of the human adenosine deaminase gene causes severe combined immunodeficiency. U. Atasoy, C. J. Norby-Slycord & M. L. Markert; Hum. Mol. Genet. 2: 1099-104 (1993) [94004847] The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated in X- linked severe combined immunodeficiency, SCIDX1 J. M. Puck, S. M. Deschenes, J. C. Porter, A. S. Dutra, C. J. Brown, H. F. Willard & P. S. Henthorn; Cell 73: 147-57 (1993) [93214986] Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans. M. Noguchi, H. Yi, H. M. Rosenblatt, A. H. Filipovich, S. Adelstein, W. S. Modi, O. W. McBride & W. J. Leonard; Am. J. Med. Genet. 42: 201-7 (1992) [92125333] Five missense mutations at the adenosine deaminase locus (ADA) detected by altered restriction fragments and their frequency in ADA--patients with severe combined immunodeficiency (ADA-SCID). R. Hirschhorn, A. Ellenbogen & S. Tzall;

16. Mutat. Res. 273: 193-202 (1992) [92186915] Three nonsense mutations responsible for group A xeroderma pigmentosum. I. Satokata, K. Tanaka, N. Miura, M. Narita, T. Mimaki, Y. Satoh, S. Kondo & Y. Okada; J. Biol. Chem. 266: 19786-9 (1991) [92011785] Identification and characterization of xpac protein, the gene product of the human XPAC (xeroderma pigmentosum group A complementing) gene. N. Miura, I. Miyamoto, H. Asahina, I. Satokata, K. Tanaka & Y. Okada;

17. Nucleic Acids Res. 21: 419-26 (1993) [93181229] Structure and expression of the excision repair gene ERCC6, involved in the human disorder Cockayne's syndrome group B. C. Troelstra, W. Hesen, D. Bootsma & J. H. Hoeijmakers;

18. Am. J. Hum. Genet. 51: 299-306 (1992) [92351926] A microdeletion of less than 250 kb, including the proximal part of the FMR-I gene and the fragile-X site, in a male with the clinical phenotype of fragile-X syndrome. D. Wohrle, D. Kotzot, M. C. Hirst, A. Manca, B. Korn, A. Schmidt, G. Barbi, H. D. Rott, A. Poustka, K. E. Davies, et al.;

19. Lancet 341: 273-5 (1993) [93148721] Direct diagnosis of carriers of point mutations in Duchenne muscular dystrophy. S. C. Yau, R. G. Roberts, M. Bobrow & C. G. Mathew. Hum. Genet. 90: 65-70 (1992) [93052247] Molecular genetic analysis of 67 patients with Duchenne/Becker muscular dystrophy. S. Niemann-Seyde, R. Slomski, F. Rininsland, U. Ellermeyer, J. Kwiatkowska & J. Reiss. Hum. Genet. 84: 228-32 (1990) [90152651] Rapid detection of deletions in the Duchenne muscular dystrophy gene by PCR amplification of deletion-prone exon sequences. M. Hentemann, J. Reiss, M. Wagner & D. N. Cooper, Nature 322: 73-7 (1986) [86257412] Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. L. M. Kunkel;

20. Genomics 18: 673-9 (1993) [94140369] Genomic organization and transcriptional units at the myotonic dystrophy locus. D. J. Shaw, M. McCurrach, S. A. Rundle, H. G. Harley, S. R. Crow, R. Sohn, J. P. Thirion, M. G. Hamshere, A. J. Buckler, P. S. Harper, et al. Arch. Neurol. 50: 1173-9 (1993) [94029649] The myotonic dystrophy gene. A. Pizzuti, D. L. Friedman & C. T. Caskey; Hum. Mol. Genet. 2: 299-304 (1993) [93271990] Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene. M. S. Mahadevan, C. Amemiya, G. Jansen, L. Sabourin, S. Baird, C. E. Neville, N. Wormskamp, B. Segers, M. Batzer, J. Lamerdin, et al.;

21. Nature 352: 77-9 (1991) [91287825] Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. A. R. La Spada, E. M. Wilson, D. B. Lubahn, A. E. Harding & K. H. Fischbeck; Neurology 42: 2300-2 (1992) [93096171] Strong correlation between the number of CAG repeats in androgen receptor genes and the clinical onset of features of spinal and bulbar muscular atrophy. S. Igarashi, Y. Tanno, O. Onodera, M. Yamazaki, S. Sato, A. Ishikawa, N. Miyatani, M. Nagashima, Y. Ishikawa, K. Sahashi, et al.; Science 256: 784-9 (1992) [92271195] Triplet repeat mutations in human disease. C. T. Caskey, A. Pizzuti, Y. H. Fu, R. G. Fenwick & D. L. Nelson;

22. Hum. Mol. Genet. 2: 1713-5 (1993) [94093563] Analysis of the huntingtin gene reveals a trinucleotide-length polymorphism in the region of the gene that contains two CCG-rich stretches and a correlation between decreased age of onset of Huntington's disease and CAG repeat number. D. C. Rubinsztein, D. E. Barton, B. C. Davison & M. A. Ferguson-Smith; Mol. Cell. Probes 7: 235-9 (1993) [93375991] A new polymerase chain reaction (PCR) assay for the trinucleotide repeat that is unstable and expanded on Huntington's disease chromosomes. J. P. Warner, L. H. Barron & D. J. Brock; Cell 72: 971-83 (1993) [93208892] A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group;

23. J. Clin. Invest. 83: 11-3 (1989) [89093407] Identification of a single nucleotide change in the hypoxanthine-guanine phosphoribosyltransferase gene (HPRTYale) responsible for Lesch-Nyhan syndrome. S. Fujimori, B. L. Davidson, W. N. Kelley & T. D. Palella; Proc. Natl. Acad. Sci. U.S.A. 86: 1919-23 (1989) [89184538] Identification of mutations leading to the Lesch-Nyhan syndrome by automated direct DNA sequencing of in vitro amplified cDNA. R. A. Gibbs, P. N. Nguyen, L. J. McBride, S. M. Koepf & C. T. Caskey; Genomics 7: 235-44 (1990) [90269813] Multiplex DNA deletion detection and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families. R. A. Gibbs, P. N. Nguyen, A. Edwards, A. B. Civitello & C. T. Caskey;

24. Nature 333: 85-6 (1988) [88202110] Identification of an altered splice site in Ashkenazi Tay-Sachs disease. E. Arpaia, A. Dumbrille-Ross, T. Maler, K. Neote, M. Tropak, C. Troxel, J. L. Stirling, J. S. Pitts, B. Bapat, A. M. Lamhonwah, et al.; J. Biol. Chem. 263:

18587-9 (1988) [89066640] The major defect in Ashkenazi Jews with Tay-Sachs disease is an insertion in the gene for the alpha-chain of beta-hexosaminidase R. Myerowitz & F. C. Costigan; Hum. Mutat. 1: 303-9 (1992) [93250824] A mutation common in non-Jewish Tay-Sachs disease: frequency and RNA studies. B. R. Akerman, J. Zielenski, B. L. Triggs-Raine, E. M. Prence, M. R. Natowicz, J. S. Lim-Steele, M. M. Kaback, E. H. Mules, G. H. Thomas, J. T. Clarke, et al.;

25. Clin. Endocrinol (Oxf) 38: 421-5 (1993) [93306853] Prenatal diagnosis of congenital adrenal hyperplasia by direct detection of mutations in the steroid 21-hydroxylase gene. G. Rumsby, J. W. Honour & C. Rodeck; Proc. Natl. Acad. Sci. U.S.A. 90: 4552-6 (1993) [93281617] Mutations in the CYP11B1 gene causing congenital adrenal hyperplasia and hypertension cluster in exons 6, 7, and 8. K. M. Curnow, L. Slutsker, J. Vitek, T. Cole, P. W. Speiser, M. I. New, P. C. White & L. Pascoe; Hum. Genet. 89: 109-10 (1992) [92250001] Prenatal diagnosis of 21-hydroxylase deficiency congenital adrenal hyperplasia using the polymerase chain reaction. D. Owerbach, M. B. Draznin, R. J. Carpenter & F. Greenberg;

26. Nucleic Acids Res. 20: 1433 (1992) [92220641] PCR detection of the insertion/deletion polymorphism of the human angiotensin converting enzyme gene (DCP1) (dipeptidyl carboxypeptidase 1). B. Rigat, C. Hubert, P. Corvol & F. Soubrier; Biochem. Biophys. Res. Commun. 184: 9-15 (1992) [92231988] Association of a polymorphism of the angiotensin I-converting enzyme gene with essential hypertension. R. Y. Zee, Y. K. Lou, L. R. Griffiths & B. J. Morris;

27. Nature 368: 258-61 (1994) [94195398] Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. C. E. Bronner, S. M. Baker, P. T. Morrison, G. Warren, L. G. Smith, M. K. Lescoe, M. Kane, C. Earabino, J. Lipford, A. Lindblom, et al.; Oncogene 9: 991-4 (1994) [94151027] DNA alterations in cells from hereditary non-polyposis colorectal cancer patients. C. Wu, Y. Akiyama, K. Imai, S. Miyake, H. Nagasaki, M. Oto, S. Okabe, T. Iwama, K. Mitamura, H. Masumitsu, et al.;

28. Science 263: 1625-9 (1994) [94174309] Mutation of a mutL homolog in hereditary colon cancer [see comments] N. Papadopoulos, N. C. Nicolaides, Y. F. Wei, S. M. Ruben, K. C.

Carter, C. A. Rosen, W. A. Haseltine, R. D. Fleischmann, C. M. Fraser, M. D. Adams, et al.; Cell 75: 1215-25 (1993) [94084796] Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. F. S. Leach, N. C. Nicolaides, N. Papadopoulos, B. Liu, J. Jen, R. Parsons, P. Peltomaki, P. Sistonen, L. A. Aaltonen, M. Nystrom-Lahti;

29. Hum. Mutat. 3: 12-8 (1994) [94163183] Exon eight APC mutations account for a disproportionate number of familial adenomatous polyposis families. D. J. Koorey, G. W. McCaughan, R. J. Trent & N. D. Gallagher, Hum. Mutat. 2: 478-84 (1993) [94154735] Simple, rapid, and accurate determination of deletion mutations by automated DNA sequencing of heteroduplex fragments of the adenomatous polyposis coli (APC) gene generated by PCR amplification. K. Tamura, Y. Yamamoto, Y. Saeki, J. Furuyama & J. Utsunomiya;

30. Biochim. Biophys. Acta 1155: 43-61 (1993) [93277907] Molecular characterization of the retinoblastoma susceptibility gene. D. W. Goodrich & W. H. Lee; Br. J. Cancer 68, 958-64 (1993) Mechanisms of oncogenesis in patients with familial retinoblastoma. Onadim,Z., Hogg,A. & J.K Cowell;

31. Cancer Res. 54: 1298-304 (1994) [94163623] Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. J. M. Birch, A. L. Hartley, K. J. Tricker, J. Prosser, A. Condie, A. M. Kelsey, M. Harris, P. H. Jones, A. Binchy, D. Crowther, et al.;

32. Leukemia 8: 186-9 (1994) [94118546] An optimized multiplex polymerase chain reaction (PCR) for detection of BCR-ABL fusion mRNAs in haematological disorders. N. C. Cross, J. V. Melo, L. Feng & J. M. Goldman; Blood 69: 971-3 (1987) [87129392] bcr-abl RNA in patients with chronic myelogenous leukemia. E. Shtivelman, R. P. Gale, O. Dreazen, A. Berrebi, R. Zaizov, I. Kubonishi, I. Miyoshi & E. Canaani bcl-2; Diagn. Mol. Pathol. 2: 241-7 (1993) [94163382] Rearrangement of the BCL-2 gene in follicular lymphoma. Detection by PCR in both fresh and fixed tissue samples. J. Liu, R. M. Johnson & S. T. Traweek; Blood 83: 1079-85 (1994) [94154269] Cytometric detection of DNA amplified with fluorescent primers: applications to analysis of clonal bcl-2 and IgH gene rearrangements in malignant lymphomas. R. L. Barker, C. A. Worth & S. C. Peiper; Br. J. Cancer 67: 922-5 (1993) [93264208] Detection of bcl-2/JH rearrangement in follicular and

diffuse lymphoma: concordant results of peripheral blood and bone marrow analysis at diagnosis. R. Yuan, P. Dowling, E. Zucca, H. Diggelmann & F. Cavalli;

33. Molecular Cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press (1989). J. Sambrook, E. Fritch, & T. Maniatis;

34. Proc. Natl. Acad. Sci. U.S.A. 74: 5463-5467 (1977), DNA Sequencing with Chain Terminating Inhibitors, F. Sanger, S. Niklen & A.R. Coulsen.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Luminex Corporation, Chandler, Van S.
Fulton, R. Jerrold
Chandler, Mark B.

(ii) TITLE OF INVENTION: Multiplexed Analysis of Clinical Specimens Aparatus and Method

(iii) NUMBER OF SEQUENCES: 34

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Melinda L. Patterson, Arnold, White & Durkee
(B) STREET: P.O. Box 4433
(C) CITY: Houston
(D) STATE: Texas
(E) COUNTRY: USA
(F) ZIP: 77210-4433

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Patterson, Melinda L.

(B) REGISTRATION NUMBER: 33,062

(C) REFERENCE/DOCKET NUMBER: ILAB:005

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (713) 787-1592

(B) TELEFAX: (713) 789-2689

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
His Tyr Gly Ser Leu Pro Gln Lys
1               5
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Tyr Gly Ser Leu Pro Gln
1 5

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCTACGCCA CCAGCTCCAA CTAC 24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCCTACGCCA CAAGCTCCAA CTAC 24

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGGTGTAAA CTTGTACCAG T 21

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTGGTAGCAG CGGTAGAGTT G 21

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGGCCAGTAC ACCCATGA 18

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TCATGGGTGT ACTGGCCA 18

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGCCAGTTC ACCCATGA 18

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TCATGGGTGA ACTGGCCA 18

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGGGCAGTAC AGCCATGA 18

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCATGGCTGT ACTGCCCA 18

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGATGAGGA GTTCTACG 18

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGTAGAACTC CTCATCTC 18

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAGATGAGCA GTTCTACG 18

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGTAGAACTG CTCATCTC 18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GAGACGAGCA GTTCTACG 18

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGTAGAACTG CTCGTCTC 18

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAGACGAGGA GTTCTATG 18

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CATAGAACTC CTCGTCTC 18

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ACCTGGAGAG GAAGGAGA 18

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TCTCCTTCCT CTCCAGGT 18

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACCTGGAGAA GAAGGAGA 18

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TCTCCTTCTT CTCCAGGT 18

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ACCTGGGGAG GAAGGAGA 18

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCTCCTTCCT CCCCAGGT 18

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCAGCAAATT TGGAGGTT 18

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AACCTCCAAA TTTGCTGA 18

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TCCACAGACT TAGATTTG 18

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CAAATCTAAG TCTGTGGA 18

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TCCGCAGATT TAGAAGAT 18

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATCTTCTAAA TCTGCGGA 18

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TCAGACAATT TAGATTTG 18

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CAAATCTAAA TTGTCTGA 18

MICROFICHE APPENDIX A

```
FORM1.FRM - 1

VERSION 2.00
Begin Form Form1
   BackColor          =   &H00FFFF80&
   BorderStyle        =   3  'Fixed Double
   Caption            =   "Digital Diagnostics, Inc."
   ClientHeight       =   6912
   ClientLeft         =   840
   ClientTop          =   1452
   ClientWidth        =   9480
   FillColor          =   &H00FFFFFF&
   FillStyle          =   4  'Upward Diagonal
   ForeColor          =   &H00808080&
   Height             =   7332
   Left               =   792
   LinkTopic          =   "Form1"
   MaxButton          =   0   'False
   MinButton          =   0   'False
   ScaleHeight        =   6912
   ScaleWidth         =   9480
   Top                =   1080
   Width              =   9576
   Begin Data Dataint
      Caption          =   "Dataint"
      Connect          =   ""
      DatabaseName     =   "C:\ACCESS\ORBIT.MDB"
      Exclusive        =   0   'False
      Height           =   252
      Left             =   8160
      Options          =   0
      ReadOnly         =   0   'False
      RecordSource     =   "interpret"
      Top              =   0
      Visible          =   0   'False
      Width            =   912
   End
   Begin Data Datatst
      Caption          =   "Datatst"
      Connect          =   ""
      DatabaseName     =   "C:\ACCESS\ORBIT.MDB"
      Exclusive        =   0   'False
      Height           =   252
      Left             =   8160
      Options          =   0
      ReadOnly         =   0   'False
      RecordSource     =   "testdef"
      Top              =   0
      Visible          =   0   'False
      Width            =   912
   End
   Begin Data Datahld
      Caption          =   "Datahld"
      Connect          =   ""
      DatabaseName     =   "C:\ACCESS\ORBIT.MDB"
      Exclusive        =   0   'False
      Height           =   252
      Left             =   8160
      Options          =   0
      ReadOnly         =   0   'False
      RecordSource     =   "class_tab"
      Top              =   0
      Visible          =   0   'False
```

```
FORM1.FRM - 2

      Width             =   912
   End
   Begin PictureBox WWPic
      BackColor         =   &H00C0C0C0&
      BorderStyle       =   0  'None
      Height            =   495
      Index             =   1
      Left              =   9000
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   372
      TabIndex          =   38
      Top               =   2400
      Width             =   375
   End
   Begin PictureBox PRPic
      BackColor         =   &H00C0C0C0&
      BorderStyle       =   0  'None
      Height            =   495
      Index             =   1
      Left              =   9000
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   372
      TabIndex          =   37
      Top               =   3000
      Width             =   375
   End
   Begin PictureBox PRPic
      BackColor         =   &H00C0C0C0&
      BorderStyle       =   0  'None
      Height            =   495
      Index             =   0
      Left              =   9000
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   372
      TabIndex          =   36
      Top               =   3000
      Visible           =   0   'False
      Width             =   375
   End
   Begin PictureBox WWPic
      BackColor         =   &H00C0C0C0&
      BorderStyle       =   0  'None
      Height            =   495
      Index             =   0
      Left              =   9000
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   372
      TabIndex          =   35
      Top               =   2400
      Visible           =   0   'False
      Width             =   375
   End
   Begin PictureBox SFPic
      BackColor         =   &H00C0C0C0&
      BorderStyle       =   0  'None
      Height            =   495
      Index             =   1
```

```
FORM1.FRM - 3

      Left              =   9000
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   372
      TabIndex          =   34
      Top               =   1800
      Width             =   375
   End
   Begin PictureBox SFPic
      BackColor         =   &H00C0C0C0&
      BorderStyle       =   0  'None
      Height            =   495
      Index             =   0
      Left              =   9000
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   372
      TabIndex          =   33
      Top               =   1800
      Visible           =   0   'False
      Width             =   375
   End
   Begin PictureBox FCPic
      BackColor         =   &H00C0C0C0&
      BorderStyle       =   0  'None
      Height            =   495
      Index             =   2
      Left              =   9000
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   372
      TabIndex          =   32
      Top               =   1200
      Width             =   375
   End
   Begin Timer Timer1
      Interval          =   2000
      Left              =   9120
      Top               =   0
   End
   Begin PictureBox Picture1
      BackColor         =   &H00FFFF80&
      BorderStyle       =   0  'None
      Height            =   495
      Left              =   960
      Picture           =   (Icon)
      ScaleHeight       =   492
      ScaleWidth        =   492
      TabIndex          =   31
      Top               =   120
      Width             =   495
   End
   Begin Data Data2
      Caption           =   "Data2"
      Connect           =   ""
      DatabaseName      =   "C:\ACCESS\orbit.MDB"
      Exclusive         =   0   'False
      Height            =   270
      Left              =   8160
      Options           =   0
      ReadOnly          =   0   'False
```

```
FORM1.FRM - 4

      RecordSource     =   "assay"
      Top              =   0
      Visible          =   0   'False
      Width            =   1455
   End
   Begin ComboBox CubeSel
      Height           =   300
      Left             =   120
      Style            =   2  'Dropdown List
      TabIndex         =   30

      Top              =   960
      Width            =   1815
   End
   Begin Data Data1
      Caption          =   "Data1"
      Connect          =   ""
      DatabaseName     =   "C:\ACCESS\orbit.MDB"
      Exclusive        =   0   'False
      Height           =   270
      Left             =   8355
      Options          =   0
      ReadOnly         =   0   'False
      RecordSource     =   "cubes"
      Top              =   0
      Visible          =   0   'False
      Width            =   1140
   End
   Begin SSCommand ResPrint
      Caption          =   "Print"
      Font3D           =   0  'None
      Height           =   495
      Left             =   7920
      Picture          =   (none)
      TabIndex         =   22
      Top              =   6120
      Width            =   1455
   End
   Begin SSCommand ResSave
      Caption          =   "Save"
      Font3D           =   0  'None
      Height           =   495
      Left             =   7920
      Picture          =   (none)
      TabIndex         =   21
      Top              =   5520
      Width            =   1455
   End
   Begin SSCommand ResClear
      Caption          =   "Clear"
      Font3D           =   0  'None
      Height           =   495
      Left             =   7920
      Picture          =   (none)
      TabIndex         =   20
      Top              =   4920
      Width            =   1455
   End
   Begin Grid Grid1
      FixedCols        =   0
      Height           =   2295
```

```
FORM1.FRM - 5
      Left            =   2640
      Rows            =   64
      ScrollBars      =   2  'Vertical
      TabIndex        =   19
      Top             =   4440
      Width           =   5175
   End
   Begin GRAPH Graph1




      AsciiSymbol     =   "0"



      GraphCaption    =   "Classification"

      GraphTitle      =   "Classification"
      GraphType       =   9  'Scatter
      GridStyle       =   3  'Horizontal and Vertical
      Height          =   2295

      Left            =   120

      NumPoints       =   20
      NumSets         =   16


      TabIndex        =   18
      Top             =   4440
      Width           =   2415

   End
   Begin SSPanel Panel3D2
      Alignment       =   0  'Left Justify - TOP
      BackColor       =   &H00C0C0C0&
      BevelInner      =   1  'Inset
      Caption         =   "Operator Insturctions"
      Font3D          =   0  'None
      Height          =   615
      Left            =   120
      TabIndex        =   15
      Top             =   3720
      Width           =   9375
      Begin TextBox OpInst
         Height       =   375
         Left         =   2040
         MultiLine    =   -1  'True
         TabIndex     =   16
         Top          =   120
         Width        =   5655
      End
```

```
FORM1.FRM - 6

      Begin SSCommand OpInstOK
         Caption          =   "OK"
         Font3D           =   0  'None
         Height           =   375
         Left             =   8040
         Picture          =   (none)
         TabIndex         =   17
         Top              =   120
         Width            =   855
      End
   End
   Begin TextBox SIBox
      Height           =   1815
      Left             =   120
      MultiLine        =   -1  'True
      TabIndex         =   14
      Top              =   1680
      Width            =   3735
   End
   Begin SSPanel Panel3D1
      Alignment        =   0  'Left Justify - TOP
      BackColor        =   &H00C0C0C0&
      Caption          =   "Cytometer Status"
      Font3D           =   0  'None
      Height           =   2535
      Left             =   7680
      TabIndex         =   12
      Top              =   960
      Width            =   1815
      Begin PictureBox FCPic
         BackColor        =   &H00C0C0C0&
         BorderStyle      =   0  'None
         Height           =   495
         Index            =   1
         Left             =   1320
         Picture          =   (Icon)
         ScaleHeight      =   492
         ScaleWidth       =   372
         TabIndex         =   25
         Top              =   240
         Visible          =   0   'False
         Width            =   375
      End
      Begin PictureBox FCPic
         BackColor        =   &H00C0C0C0&
         BorderStyle      =   0  'None
         Height           =   495
         Index            =   0
         Left             =   1320
         Picture          =   (Icon)
         ScaleHeight      =   492
         ScaleWidth       =   372
         TabIndex         =   24
         Top              =   240
         Visible          =   0   'False
         Width            =   375
      End
      Begin Label Label10
         Alignment        =   1  'Right Justify
         BackStyle        =   0  'Transparent
         Caption          =   "Pressure"
```

```
FORM1.FRM - 7

         Height          =   255
         Left            =   120
         TabIndex        =   29
         Top             =   2160
         Width           =   1095
      End
      Begin Label Label9
         Alignment       =   1  'Right Justify
         BackStyle       =   0  'Transparent
         Caption         =   "Waste Water"
         Height          =   255
         Left            =   0
         TabIndex        =   28
         Top             =   1560
         Width           =   1215
      End
      Begin Label Label8
         Alignment       =   1  'Right Justify
         BackStyle       =   0  'Transparent
         Caption         =   "Sheath Fluid"
         Height          =   255
         Left            =   120
         TabIndex        =   27
         Top             =   960
         Width           =   1095
      End
      Begin Label Label7
         Alignment       =   1  'Right Justify
         BackStyle       =   0  'Transparent
         Caption         =   "Flow Control"
         Height          =   255
         Left            =   120
         TabIndex        =   26
         Top             =   360
         Width           =   1095
      End
   End
   Begin SSFrame TstCtl
      Caption         =   "Test Control"
      Font3D          =   0  'None
      Height          =   2535
      Left            =   3960
      TabIndex        =   4
      Top             =   960
      Width           =   1935
      Begin SSCommand TCHalt
         Caption         =   "Counts"
         Enabled         =   0   'False
         Font3D          =   0  'None
         Height          =   495
         Left            =   240
         Picture         =   (none)
         TabIndex        =   7
         Top             =   1680
         Width           =   1455
      End
      Begin SSCommand TCStart
         Caption         =   "Start Test"
         Enabled         =   0   'False
         Font3D          =   0  'None
         Height          =   495
```

```
FORM1.FRM - 8
         Left              =   240
         Picture           =   (none)
         TabIndex          =   6
         Top               =   1080
         Width             =   1455
      End
      Begin SSCommand TCInit
         Caption           =   "Initialize"
         Enabled           =   0   'False
         Font3D            =   0   'None
         Height            =   495
         Left              =   240
         Picture           =   (none)
         TabIndex          =   5
         Top               =   480
         Width             =   1455
      End
   End
   Begin SSFrame MacCtl
      Caption           =   "Machine Control"
      Font3D            =   0   'None
      Height            =   2535
      Left              =   5880
      TabIndex          =   3
      Top               =   960
      Width             =   1815
      Begin SSCommand MCEnd
         Caption           =   "Exit"
         Font3D            =   0   'None
         Height            =   495
         Left              =   240
         Picture           =   (none)
         TabIndex          =   10
         Top               =   1680
         Width             =   1335
      End
      Begin SSCommand MCAdjust
         Caption           =   "Manual Adjust"
         Font3D            =   0   'None
         Height            =   495
         Left              =   240
         Picture           =   (none)
         TabIndex          =   9
         Top               =   1080
         Width             =   1335
      End
      Begin SSCommand MCCalib
         Caption           =   "Calibrate"
         Font3D            =   0   'None
         Height            =   495
         Left              =   240
         Picture           =   (none)
         TabIndex          =   8
         Top               =   480
         Width             =   1335
      End
   End
   Begin ComboBox AssaySel
      Height            =   300
      Left              =   2040
      Style             =   2   'Dropdown List
```

```
FORM1.FRM - 9

      TabIndex          =   2

      Top               =   960
      Width             =   1815
   End
   Begin Label Label6
      Alignment         =   2  'Center
      BackStyle         =   0  'Transparent
      BorderStyle       =   1  'Fixed Single
      Caption           =   "Results"
      FontBold          =   -1  'True
      FontItalic        =   -1  'True
      FontName          =   "MS Sans Serif"
      FontSize          =   12
      FontStrikethru    =   0   'False
      FontUnderline     =   0   'False
      ForeColor         =   &H00000080&
      Height            =   375
      Left              =   7920
      TabIndex          =   23
      Top               =   4440
      Width             =   1455
   End
   Begin Label Label4
      BackStyle         =   0  'Transparent
      Caption           =   "Sample Identification"
      Height            =   255
      Left              =   120
      TabIndex          =   13
      Top               =   1440
      Width             =   2535
   End
   Begin Line Line1
      X1                =   0
      X2                =   9480
      Y1                =   3600
      Y2                =   3600
   End
   Begin Label Label2
      BackStyle         =   0  'Transparent
      Caption           =   "Specific Assay"
      Height            =   255
      Left              =   2040
      TabIndex          =   0
      Top               =   720
      Width             =   1575
   End
   Begin Label Label1
      BackStyle         =   0  'Transparent
      Caption           =   "Assay Cube"
      Height            =   255
      Left              =   120
      TabIndex          =   1
      Top               =   720
      Width             =   1575
   End
   Begin Label Label3
      Alignment         =   2  'Center
      BackStyle         =   0  'Transparent
      Caption           =   "Orbit Diagnostic Operating System"
      FontBold          =   -1  'True
```

```
FORM1.FRM - 10

      FontItalic      =   0   'False
      FontName        =   "MS Sans Serif"
      FontSize        =   18
      FontStrikethru  =   0   'False
      FontUnderline   =   0   'False
      ForeColor       =   &H00808000&
      Height          =   495
      Left            =   120
      TabIndex        =   11
      Top             =   120
      Width           =   9255
   End
End
```

```
FORM2.FRM - 1

VERSION 2.00
Begin Form Form2
   BorderStyle     =   3  'Fixed Double
   Caption         =   "Manual Adjust"
   ClientHeight    =   6612
   ClientLeft      =   1212
   ClientTop       =   1128
   ClientWidth     =   5508
   Height          =   7032
   Left            =   1164
   LinkTopic       =   "Form2"
   MaxButton       =   0   'False
   MinButton       =   0   'False
   ScaleHeight     =   6612
   ScaleWidth      =   5508
   Top             =   756
   Width           =   5604
   Begin SSFrame Frame3D5
      Caption          =   "DDM Select"
      Font3D           =   3  'Inset w/light shading
      ForeColor        =   &H00000000&
      Height           =   735
      Left             =   2760
      TabIndex         =   63
      Top              =   5880
      Width            =   1455
      Begin OptionButton Option1
         BackColor        =   &H000000FF&
         Caption          =   "Option1"
         Height           =   255
         Index            =   2
         Left             =   1080
         TabIndex         =   66
         Top              =   360
         Width            =   255
      End
      Begin OptionButton Option1
         BackColor        =   &H000080FF&
         Caption          =   "Option1"
         Height           =   255
         Index            =   1
         Left             =   600
         TabIndex         =   65
         Top              =   360
         Width            =   255
      End
      Begin OptionButton Option1
         BackColor        =   &H0000FF00&
         Caption          =   "Option1"
         Height           =   255
         Index            =   0
         Left             =   120
         TabIndex         =   64
         Top              =   360
         Value            =   -1  'True
         Width            =   255
      End
      Begin Shape Shape3
         BackColor        =   &H000000FF&
         BackStyle        =   1  'Opaque
         Height           =   495
```

```
FORM2.FRM - 2

         Left            =   960
         Top             =   240
         Width           =   495
      End
      Begin Shape Shape2
         BackColor       =   &H000080FF&
         BackStyle       =   1  'Opaque
         Height          =   495
         Left            =   480
         Top             =   240
         Width           =   495
      End
      Begin Shape Shape1
         BackColor       =   &H0000FF00&
         BackStyle       =   1  'Opaque
         Height          =   495
         Left            =   0
         Top             =   240
         Width           =   495
      End
   End
   Begin SSPanel Panel3D3
      Alignment       =   0  'Left Justify - TOP
      BackColor       =   &H00C0C0C0&
      BevelInner      =   1  'Inset
      BorderWidth     =   4
      Font3D          =   0  'None
      Height          =   1095
      Left            =   4200
      TabIndex        =   56
      Top             =   2880
      Width           =   1335
      Begin PictureBox Picture1
         BackColor       =   &H00FFFF80&
         BorderStyle     =   0  'None
         Height          =   855
         Left            =   120
         Picture         =   (Icon)
         ScaleHeight     =   852
         ScaleWidth      =   1092
         TabIndex        =   57
         Top             =   120
         Width           =   1095
         Begin Data facset
            Caption         =   "Data1"
            Connect         =   ""
            DatabaseName    =   "C:\ACCESS\orbit.MDB"
            Exclusive       =   0   'False
            Height          =   270
            Left            =   0
            Options         =   0
            ReadOnly        =   0   'False
            RecordSource    =   "facs_settings"
            Top             =   600
            Visible         =   0   'False
            Width           =   1140
         End
         Begin Label Label11
            BackStyle       =   0  'Transparent
            Caption         =   "ORBIT"
            ForeColor       =   &H00808000&
```

```
FORM2.FRM - 3

            Height          =   255
            Left            =   360
            TabIndex        =   58
            Top             =   480
            Width           =   735
         End
      End
   End
   Begin SSPanel Panel3D2
      Alignment       =   0  'Left Justify - TOP
      BackColor       =   &H00C0C0C0&
      BevelInner      =   1  'Inset
      BorderWidth     =   4
      Font3D          =   0  'None
      Height          =   2655
      Left            =   4200
      TabIndex        =   54
      Top             =   3960
      Width           =   1335
      Begin SSCommand Command3D5
         Caption         =   "Done"
         Font3D          =   0  'None
         Height          =   495
         Left            =   120
         Picture         =   (none)
         TabIndex        =   62
         Top             =   1920
         Width           =   1095
      End
      Begin SSCommand Command3D4
         Caption         =   "Save"
         Font3D          =   0  'None
         Height          =   495
         Left            =   120
         Picture         =   (none)
         TabIndex        =   61
         Top             =   1320
         Width           =   1095
      End
      Begin SSCommand Command3D3
         Caption         =   "Reset"
         Font3D          =   0  'None
         Height          =   495
         Left            =   120
         Picture         =   (none)
         TabIndex        =   60
         Top             =   720
         Width           =   1095
      End
      Begin SSCommand Command3D2
         Caption         =   "Set"
         Font3D          =   0  'None
         Height          =   495
         Left            =   120
         Picture         =   (none)
         TabIndex        =   59
         Top             =   120
         Width           =   1095
      End
      Begin SSCommand Command3D1
         Caption         =   "exit"
```

```
FORM2.FRM - 4

         Font3D          =   0  'None
         Height          =   375
         Left            =   240
         Picture         =   (none)
         TabIndex        =   55
         Top             =   2880
         Width           =   735
      End
   End
   Begin SSPanel Panel3D1
      Alignment       =   0  'Left Justify - TOP
      BackColor       =   &H00C0C0C0&
      BevelInner      =   1  'Inset
      BorderWidth     =   4
      Caption         =   "Status"
      Font3D          =   0  'None
      Height          =   2895
      Left            =   4200
      TabIndex        =   53
      Top             =   0
      Width           =   1335
      Begin Label Stat
         Alignment       =   2  'Center
         BackStyle       =   0  'Transparent
         Caption         =   "Label15"
         Height          =   255
         Index           =   3
         Left            =   120
         TabIndex        =   74
         Top             =   2520
         Width           =   1095
      End
      Begin Label Label15
         Alignment       =   2  'Center
         BackStyle       =   0  'Transparent
         Caption         =   "DDM"
         Height          =   255
         Left            =   120
         TabIndex        =   73
         Top             =   2280
         Width           =   1095
      End
      Begin Label Stat
         Alignment       =   2  'Center
         BackStyle       =   0  'Transparent
         Caption         =   "Label15"
         Height          =   255
         Index           =   2
         Left            =   120
         TabIndex        =   72
         Top             =   1920
         Width           =   1095
      End
      Begin Label Stat
         Alignment       =   2  'Center
         BackStyle       =   0  'Transparent
         Caption         =   "Label15"
         Height          =   255
         Index           =   1
         Left            =   120
         TabIndex        =   71
```

```
FORM2.FRM - 5

         Top              =   1320
         Width            =   1095
      End
      Begin Label Stat
         Alignment        =   2  'Center
         BackStyle        =   0  'Transparent
         Caption          =   "Label15"
         Height           =   255
         Index            =   0
         Left             =   120
         TabIndex         =   70
         Top              =   720
         Width            =   1095
      End
      Begin Label Label14
         Alignment        =   2  'Center
         BackStyle        =   0  'Transparent
         Caption          =   "Samp. Volts"
         Height           =   255
         Left             =   120
         TabIndex         =   69
         Top              =   1680
         Width            =   1095
      End
      Begin Label Label13
         Alignment        =   2  'Center
         BackStyle        =   0  'Transparent
         Caption          =   "Laser Amps"
         Height           =   255
         Left             =   120
         TabIndex         =   68
         Top              =   1080
         Width            =   1095
      End
      Begin Label Label12
         Alignment        =   2  'Center
         BackStyle        =   0  'Transparent
         Caption          =   "Laser Volts"
         Height           =   255
         Left             =   120
         TabIndex         =   67
         Top              =   480
         Width            =   1095
      End
   End
   Begin SSFrame Frame3D4
      Caption          =   "DDM Amp"
      Font3D           =   3  'Inset w/light shading
      ForeColor        =   &H00000000&
      Height           =   975
      Left             =   0
      TabIndex         =   46
      Top              =   5640
      Width            =   2775
      Begin HScrollBar ddmscroll
         Height           =   255
         Index            =   0
         LargeChange      =   10
         Left             =   720
         Max              =   999
         TabIndex         =   51
```

```
FORM2.FRM - 6

         Top               =   240
         Width             =   1335
      End
      Begin TextBox ddmtxt
         Height            =   285
         Index             =   0
         Left              =   2160
         TabIndex          =   50
         Text              =   "Text1"
         Top               =   240
         Width             =   495
      End
      Begin HScrollBar ddmscroll
         Height            =   255
         Index             =   1
         LargeChange       =   10
         Left              =   720
         Max               =   999
         TabIndex          =   48
         Top               =   600
         Width             =   1335
      End
      Begin TextBox ddmtxt
         Height            =   285
         Index             =   1
         Left              =   2160
         TabIndex          =   47
         Text              =   "Text1"
         Top               =   600
         Width             =   495
      End
      Begin Label Label10
         BackStyle         =   0  'Transparent
         Caption           =   "FLA"
         Height            =   255
         Index             =   5
         Left              =   120
         TabIndex          =   52
         Top               =   240
         Width             =   615
      End
      Begin Label Label10
         BackStyle         =   0  'Transparent
         Caption           =   "FLW"
         Height            =   255
         Index             =   4
         Left              =   120
         TabIndex          =   49
         Top               =   600
         Width             =   615
      End
   End
   Begin SSFrame Frame3D3
      Caption           =   "Trigger Level"
      Font3D            =   3  'Inset w/light shading
      ForeColor         =   &H00000000&
      Height            =   2055
      Left              =   2760
      TabIndex          =   38
      Top               =   3840
      Width             =   1455
```

FORM2.FRM - 7

```
      Begin VScrollBar VScroll1
         Height          =   1575
         LargeChange     =   10
         Left            =   960
         Max             =   999
         TabIndex        =   41
         Top             =   360
         Width           =   255
      End
      Begin SSOption trigger
         Caption         =   "FL3"
         Font3D          =   0  'None
         Height          =   255
         Index           =   4
         Left            =   120
         TabIndex        =   45
         Top             =   1680
         Width           =   615
      End
      Begin SSOption trigger
         Caption         =   "FL2"
         Font3D          =   0  'None
         Height          =   255
         Index           =   3
         Left            =   120
         TabIndex        =   44
         Top             =   1440
         Width           =   615
      End
      Begin SSOption trigger
         Caption         =   "FL1"
         Font3D          =   0  'None
         Height          =   255
         Index           =   2
         Left            =   120
         TabIndex        =   43
         Top             =   1200
         Width           =   615
      End
      Begin SSOption trigger
         Caption         =   "SSC"
         Font3D          =   0  'None
         Height          =   255
         Index           =   1
         Left            =   120
         TabIndex        =   42
         Top             =   960
         Width           =   615
      End
      Begin TextBox trigval
         Height          =   285
         Left            =   120
         TabIndex        =   40
         Text            =   "Text2"
         Top             =   360
         Width           =   615
      End
      Begin SSOption trigger
         Caption         =   "FSC"
         Font3D          =   0  'None
         Height          =   255
```

```
FORM2.FRM - 8

         Index           =   0
         Left            =   120
         TabIndex        =   39
         Top             =   720
         Width           =   615
      End
   End
   Begin SSFrame Frame3D2
      Caption         =   "Compensation"
      Font3D          =   3  'Inset w/light shading
      ForeColor       =   &H00000000&
      Height          =   1815
      Left            =   0
      TabIndex        =   25
      Top             =   3840
      Width           =   2775
      Begin TextBox Text1
         Height          =   285
         Index           =   3
         Left            =   2160
         TabIndex        =   33
         Text            =   "Text1"
         Top             =   1440
         Width           =   495
      End
      Begin TextBox Text1
         Height          =   285
         Index           =   2
         Left            =   2160
         TabIndex        =   32
         Text            =   "Text1"
         Top             =   1080
         Width           =   495
      End
      Begin TextBox Text1
         Height          =   285
         Index           =   1
         Left            =   2160
         TabIndex        =   31
         Text            =   "Text1"
         Top             =   720
         Width           =   495
      End
      Begin TextBox Text1
         Height          =   285
         Index           =   0
         Left            =   2160
         TabIndex        =   30
         Text            =   "Text1"
         Top             =   360
         Width           =   495
      End
      Begin HScrollBar HScroll1
         Height          =   255
         Index           =   3
         LargeChange     =   10
         Left            =   720
         Max             =   999
         TabIndex        =   29
         Top             =   1440
         Width           =   1335
```

FORM2.FRM - 9

```
End
Begin HScrollBar HScroll1
   Height          =   255
   Index           =   2
   LargeChange     =   10
   Left            =   720
   Max             =   999
   TabIndex        =   28
   Top             =   1080
   Width           =   1335
End
Begin HScrollBar HScroll1
   Height          =   255
   Index           =   1
   LargeChange     =   10
   Left            =   720
   Max             =   999
   TabIndex        =   27
   Top             =   720
   Width           =   1335
End
Begin HScrollBar HScroll1
   Height          =   255
   Index           =   0
   LargeChange     =   10
   Left            =   720
   Max             =   999
   TabIndex        =   26
   Top             =   360
   Width           =   1335
End
Begin Label Label10
   BackStyle       =   0  'Transparent
   Caption         =   "FL3-2"
   Height          =   255
   Index           =   3
   Left            =   120
   TabIndex        =   37
   Top             =   1440
   Width           =   615
End
Begin Label Label10
   BackStyle       =   0  'Transparent
   Caption         =   "FL2-3"
   Height          =   255
   Index           =   2
   Left            =   120
   TabIndex        =   36
   Top             =   1080
   Width           =   615
End
Begin Label Label10
   BackStyle       =   0  'Transparent
   Caption         =   "FL2-1"
   Height          =   255
   Index           =   1
   Left            =   120
   TabIndex        =   35
   Top             =   720
   Width           =   615
End
```

```
FORM2.FRM - 10

      Begin Label Label10
         BackStyle          =   0  'Transparent
         Caption            =   "FL1-2"
         Height             =   255
         Index              =   0
         Left               =   120
         TabIndex           =   34
         Top                =   360
         Width              =   615
      End
   End
   Begin SSFrame modefr
      Caption             =   "Mode"
      Font3D              =   3  'Inset w/light shading
      ForeColor           =   &H00000000&
      Height              =   975
      Left                =   0
      TabIndex            =   14
      Top                 =   2880
      Width               =   4215
      Begin SSRibbon fl3mod
         AutoSize         =   0  'None
         BackColor        =   &H00C0C0C0&
         GroupAllowAllUp  =   0   'False
         GroupNumber      =   5
         Height           =   240
         Index            =   1
         Left             =   2760
         PictureDisabled  =   (none)
         PictureDn        =   (none)
         PictureDnChange  =   1  'Dither 'PictureUp' Bitmap
         PictureUp        =   (none)
         Top              =   600
         Value            =   -1  'True
         Width            =   495
      End
      Begin SSRibbon fl3mod
         AutoSize         =   0  'None
         BackColor        =   &H00C0C0C0&
         GroupAllowAllUp  =   0   'False
         GroupNumber      =   5
         Height           =   240
         Index            =   0
         Left             =   2760
         PictureDisabled  =   (none)
         PictureDn        =   (none)
         PictureDnChange  =   1  'Dither 'PictureUp' Bitmap
         PictureUp        =   (none)
         Top              =   360
         Width            =   495
      End
      Begin SSRibbon fl2mod
         AutoSize         =   0  'None
         BackColor        =   &H00C0C0C0&
         GroupAllowAllUp  =   0   'False
         GroupNumber      =   4
         Height           =   240
         Index            =   1
         Left             =   2280
         PictureDisabled  =   (none)
         PictureDn        =   (none)
```

```
FORM2.FRM - 11

      PictureDnChange =   1  'Dither 'PictureUp' Bitmap
      PictureUp       =   (none)
      Top             =   600
      Value           =   -1  'True
      Width           =   495
   End
   Begin SSRibbon fl2mod
      AutoSize        =   0  'None
      BackColor       =   &H00C0C0C0&
      GroupAllowAllUp =   0   'False
      GroupNumber     =   4
      Height          =   240
      Index           =   0
      Left            =   2280
      PictureDisabled =   (none)
      PictureDn       =   (none)
      PictureDnChange =   1  'Dither 'PictureUp' Bitmap
      PictureUp       =   (none)
      Top             =   360
      Width           =   495
   End
   Begin SSRibbon fl1mod
      AutoSize        =   0  'None
      BackColor       =   &H00C0C0C0&
      GroupAllowAllUp =   0   'False
      GroupNumber     =   3
      Height          =   240
      Index           =   1
      Left            =   1800
      PictureDisabled =   (none)
      PictureDn       =   (none)
      PictureDnChange =   1  'Dither 'PictureUp' Bitmap
      PictureUp       =   (none)
      Top             =   600
      Value           =   -1  'True
      Width           =   495
   End
   Begin SSRibbon fl1mod
      AutoSize        =   0  'None
      BackColor       =   &H00C0C0C0&
      GroupAllowAllUp =   0   'False
      GroupNumber     =   3
      Height          =   240
      Index           =   0
      Left            =   1800
      PictureDisabled =   (none)
      PictureDn       =   (none)
      PictureDnChange =   1  'Dither 'PictureUp' Bitmap
      PictureUp       =   (none)
      Top             =   360
      Width           =   495
   End
   Begin SSRibbon sscmod
      AutoSize        =   0  'None
      BackColor       =   &H00C0C0C0&
      GroupAllowAllUp =   0   'False
      GroupNumber     =   2
      Height          =   240
      Index           =   1
      Left            =   1320
      PictureDisabled =   (none)
```

FORM2.FRM - 12

```
      PictureDn          =   (none)
      PictureDnChange    =   1  'Dither 'PictureUp' Bitmap
      PictureUp          =   (none)
      Top                =   600
      Value              =   -1  'True
      Width              =   495
   End
   Begin SSRibbon sscmod
      AutoSize           =   0  'None
      BackColor          =   &H00C0C0C0&
      GroupAllowAllUp    =   0   'False
      GroupNumber        =   2
      Height             =   240
      Index              =   0
      Left               =   1320
      PictureDisabled    =   (none)
      PictureDn          =   (none)
      PictureDnChange    =   1  'Dither 'PictureUp' Bitmap
      PictureUp          =   (none)
      Top                =   360
      Width              =   495
   End
   Begin SSRibbon fscmod
      AutoSize           =   0  'None
      BackColor          =   &H00C0C0C0&
      GroupAllowAllUp    =   0   'False
      Height             =   240
      Index              =   1
      Left               =   840
      PictureDisabled    =   (none)
      PictureDn          =   (none)
      PictureDnChange    =   1  'Dither 'PictureUp' Bitmap
      PictureUp          =   (none)
      Top                =   600
      Value              =   -1  'True
      Width              =   495
   End
   Begin SSRibbon fscmod
      AutoSize           =   0  'None
      BackColor          =   &H00C0C0C0&
      GroupAllowAllUp    =   0   'False
      Height             =   240
      Index              =   0
      Left               =   840
      PictureDisabled    =   (none)
      PictureDn          =   (none)
      PictureDnChange    =   1  'Dither 'PictureUp' Bitmap
      PictureUp          =   (none)
      Top                =   360
      Width              =   495
   End
   Begin Label Label9
      Alignment          =   2  'Center
      BackStyle          =   0  'Transparent
      Caption            =   "F13"
      Height             =   255
      Left               =   2760
      TabIndex           =   15
      Top                =   120
      Width              =   495
   End
```

```
FORM2.FRM - 13

      Begin Label Label8
         Alignment        =    2  'Center
         BackStyle        =    0  'Transparent
         Caption          =    "FL2"
         ForeColor        =    &H00000000&
         Height           =    255
         Left             =    2280
         TabIndex         =    16
         Top              =    120
         Width            =    495
      End
      Begin Label Label7
         Alignment        =    2  'Center
         BackStyle        =    0  'Transparent
         Caption          =    "FL1"
         ForeColor        =    &H00000000&
         Height           =    255
         Left             =    1800
         TabIndex         =    17
         Top              =    120
         Width            =    495
      End
      Begin Label Label6
         Alignment        =    2  'Center
         BackStyle        =    0  'Transparent
         Caption          =    "SSC"
         Height           =    255
         Left             =    1320
         TabIndex         =    20
         Top              =    120
         Width            =    495
      End
      Begin Label Label4
         Alignment        =    2  'Center
         BackColor        =    &H00000000&
         BackStyle        =    0  'Transparent
         Caption          =    "FSC"
         ForeColor        =    &H00000000&
         Height           =    255
         Left             =    840
         TabIndex         =    21
         Top              =    120
         Width            =    495
      End
      Begin Label Label5
         BackStyle        =    0  'Transparent
         Caption          =    "Log"
         Height           =    255
         Left             =    120
         TabIndex         =    24
         Top              =    360
         Width            =    615
      End
      Begin Label Linear
         BackStyle        =    0  'Transparent
         Caption          =    "Linear"
         Height           =    255
         Left             =    120
         TabIndex         =    23
         Top              =    600
         Width            =    855
```

```
FORM2.FRM - 14

      End
   End
   Begin SSFrame Frame3D1
      Caption          =   "Amplifier"
      Font3D           =   3  'Inset w/light shading
      ForeColor        =   &H00000000&
      Height           =   1455
      Left             =   0
      TabIndex         =   8
      Top              =   1440
      Width            =   4215
      Begin SpinButton Spina
         BackColor        =   &H000000FF&
         Delay            =   100
         ForeColor        =   &H00C0C0C0&
         Height           =   495
         Index            =   4
         Left             =   3360
         LightColor       =   &H000000C0&
         ShadeColor       =   &H00000080&
         ShadowBackColor  =   &H000000FF&
         ShadowForeColor  =   &H000000FF&
         SpinBackColor    =   &H000000FF&
         SpinForeColor    =   &H00404040&
         TdThickness      =   2
         Top              =   720
         Width            =   495
       End
       Begin SpinButton Spina
         BackColor        =   &H000080FF&
         Delay            =   100
         ForeColor        =   &H00C0C0C0&
         Height           =   495
         Index            =   3
         Left             =   2640
         LightColor       =   &H000080FF&
         ShadeColor       =   &H00404080&
         ShadowBackColor  =   &H000080FF&
         ShadowForeColor  =   &H000080FF&
         SpinBackColor    =   &H000080FF&
         SpinForeColor    =   &H00404040&
         TdThickness      =   2
         Top              =   720
         Width            =   495
       End
       Begin SpinButton Spina
         BackColor        =   &H0000FF00&
         Delay            =   100
         ForeColor        =   &H00C0C0C0&
         Height           =   495
         Index            =   2
         Left             =   1920
         LightColor       =   &H0000C000&
         ShadeColor       =   &H00008000&
         ShadowBackColor  =   &H0000FF00&
         ShadowForeColor  =   &H0000FF00&
         SpinBackColor    =   &H0000FF00&
         SpinForeColor    =   &H00404040&
         TdThickness      =   2
         Top              =   720
         Width            =   495
```

FORM2.FRM - 15

```
End
Begin SpinButton Spina
   BackColor        =   &H00FFFF00&
   Delay            =   100
   Height           =   495
   Index            =   1
   Left             =   1080
   SpinBackColor    =   &H00FFFF00&
   TdThickness      =   1
   Top              =   720
   Width            =   495
End
Begin SpinButton Spina
   BackColor        =   &H0080FF80&
   ForeColor        =   &H00FFFF00&
   Height           =   495
   Index            =   0
   Left             =   360
   LightColor       =   &H00FFFF80&
   ShadeColor       =   &H00FFFF80&
   ShadowForeColor  =   &H0080FF80&
   SpinBackColor    =   &H0080FF80&
   SpinForeColor    =   &H80000008&
   TdThickness      =   1
   Top              =   720
   Width            =   495
End
Begin TextBox amp
   Height           =   285
   Index            =   4
   Left             =   3360
   TabIndex         =   13
   Text             =   "Text1"
   Top              =   360
   Width            =   495
End
Begin TextBox amp
   Height           =   285
   Index            =   3
   Left             =   2640
   TabIndex         =   12
   Text             =   "Text1"
   Top              =   360
   Width            =   495
End
Begin TextBox amp
   Height           =   285
   Index            =   2
   Left             =   1920
   TabIndex         =   11
   Text             =   "Text1"
   Top              =   360
   Width            =   495
End
Begin TextBox amp
   Height           =   285
   Index            =   1
   Left             =   1080
   TabIndex         =   10
   Text             =   "Text1"
   Top              =   360
```

FORM2.FRM - 16

```
      Width          =   495
   End
   Begin TextBox amp
      Height         =   285
      Index          =   0
      Left           =   360
      TabIndex       =   9
      Text           =   "Text1"
      Top            =   360
      Width          =   495
   End
   Begin Label Label3
      Alignment      =   2  'Center
      BackStyle      =   0  'Transparent
      Caption        =   "SSC"
      Height         =   255
      Left           =   1080
      TabIndex       =   22
      Top            =   1200
      Width          =   495
   End
   Begin Label Label1
      Alignment      =   2  'Center
      BackStyle      =   0  'Transparent
      Caption        =   "FSC"
      Height         =   255
      Index          =   4
      Left           =   360
      TabIndex       =   19
      Top            =   1200
      Width          =   495
   End
   Begin Label Label1
      Alignment      =   2  'Center
      BackStyle      =   0  'Transparent
      Caption        =   "FSC"
      Height         =   15
      Index          =   3
      Left           =   0
      TabIndex       =   18
      Top            =   240
      Width          =   495
   End
End
Begin TextBox txtNumber
   Height         =   285
   Index          =   4
   Left           =   3360
   TabIndex       =   5
   Text           =   "Text1"
   Top            =   360
   Width          =   495
End
Begin TextBox txtNumber
   Height         =   285
   Index          =   3
   Left           =   2640
   TabIndex       =   4
   Text           =   "Text1"
   Top            =   360
   Width          =   495
```

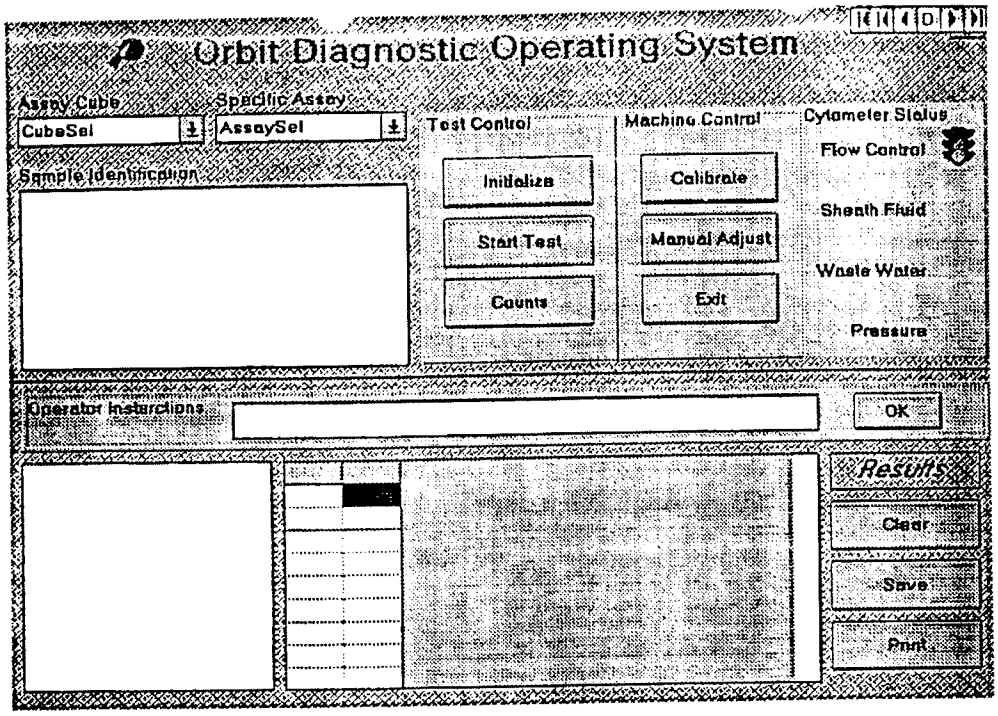
FORM2.FRM - 17

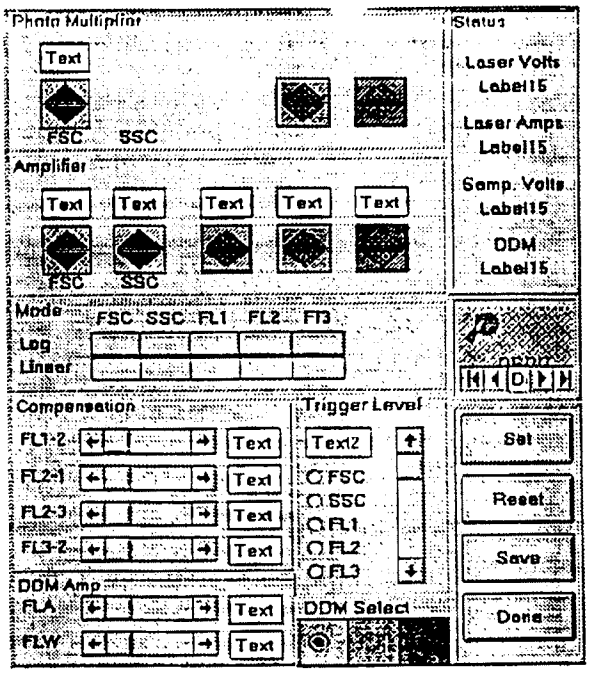
```
   End
   Begin TextBox txtNumber
      Height          =   285
      Index           =   2
      Left            =   1920
      TabIndex        =   3
      Text            =   "Text1"
      Top             =   360
      Width           =   495
   End
   Begin TextBox txtNumber
      Height          =   285
      Index           =   1
      Left            =   1080
      TabIndex        =   2
      Text            =   "Text1"
      Top             =   360
      Width           =   495
   End
   Begin SpinButton Spin1
      BackColor       =   &H0000FF00&
      Delay           =   100
      ForeColor       =   &H00C0C0C0&
      Height          =   495
      Index           =   2
      Left            =   1920
      LightColor      =   &H0000C000&
      ShadeColor      =   &H00008000&
      ShadowBackColor =   &H0000FF00&
      ShadowForeColor =   &H0000FF00&
      SpinBackColor   =   &H0000FF00&
      SpinForeColor   =   &H00404040&
      TdThickness     =   2
      Top             =   720
      Width           =   495
   End
   Begin SpinButton Spin1
      BackColor       =   &H00FFFF00&
      Delay           =   100
      Height          =   495
      Index           =   1
      Left            =   1080
      SpinBackColor   =   &H00FFFF00&
      TdThickness     =   1
      Top             =   720
      Width           =   495
   End
   Begin SSFrame PMT
      Caption         =   "Photo Multiplier"
      Font3D          =   3  'Inset w/light shading
      Height          =   1455
      Index           =   0
      Left            =   0
      TabIndex        =   0
      Top             =   0
      Width           =   4215
      Begin SpinButton Spin1
         BackColor       =   &H000000FF&
         Delay           =   100
         ForeColor       =   &H00C0C0C0&
         Height          =   495
```

FORM2.FRM - 18

```
      Index           =   4
      Left            =   3360
      LightColor      =   &H000000C0&
      ShadeColor      =   &H00000080&
      ShadowBackColor =   &H000000FF&
      ShadowForeColor =   &H000000FF&
      SpinBackColor   =   &H000000FF&
      SpinForeColor   =   &H00404040&
      TdThickness     =   2
      Top             =   720
      Width           =   495
   End
   Begin SpinButton Spin1
      BackColor       =   &H000080FF&
      Delay           =   100
      ForeColor       =   &H00C0C0C0&
      Height          =   495
      Index           =   3
      Left            =   2640
      LightColor      =   &H000080FF&
      ShadeColor      =   &H00404080&
      ShadowBackColor =   &H000080FF&
      ShadowForeColor =   &H000080FF&
      SpinBackColor   =   &H000080FF&
      SpinForeColor   =   &H00404040&
      TdThickness     =   2
      Top             =   720
      Width           =   495
   End
   Begin TextBox txtNumber
      Height          =   285
      Index           =   0
      Left            =   360
      TabIndex        =   1
      Text            =   "Text1"
      Top             =   360
      Width           =   495
   End
   Begin SpinButton Spin1
      BackColor       =   &H0080FF80&
      ForeColor       =   &H00FFFF00&
      Height          =   495
      Index           =   0
      Left            =   360
      LightColor      =   &H00FFFF80&
      ShadeColor      =   &H00FFFF80&
      ShadowForeColor =   &H0080FF80&
      SpinBackColor   =   &H0080FF80&
      SpinForeColor   =   &H80000008&
      TdThickness     =   1
      Top             =   720
      Width           =   495
   End
   Begin Label Label2
      Alignment       =   2  'Center
      BackStyle       =   0  'Transparent
      Caption         =   "SSC"
      Height          =   255
      Left            =   1080
      TabIndex        =   7
      Top             =   1200
```

```
FORM2.FRM - 19
         Width             =   495
      End
      Begin Label Label1
         Alignment         =   2   'Center
         BackStyle         =   0   'Transparent
         Caption           =   "FSC"
         Height            =   255
         Index             =   0
         Left              =   360
         TabIndex          =   6
         Top               =   1200
         Width             =   495
      End
   End
   Begin Line Line2
      X1                =   2160
      X2                =   3360
      Y1                =   3120
      Y2                =   3600
   End
   Begin Line Line1
      X1                =   2160
      X2                =   3360
      Y1                =   3120
      Y2                =   3600
   End
End
```


Orbit Diagnostic Operating System
Assay Cube
CubeSel
Specific Assay
AssaySel
Test Control
Initialize
Start Test
Counts
Machine Control
Calibrate
Manual Adjust
Exit
Cytometer Status
Flow Control
Sheath Fluid
Waste Water
Pressure
Sample Identification
Operator Instructions
OK
Results
Clear
Save
Print


Photo Multiplier
Text
FSC
SSC
Amplifier
Text
Text
Text
Text
Text
FSC
SSC
Mode
FSC SSC FL1 FL2 FL3
Log
Linear
Compensation
FL1-2
FL2-1
FL2-3
FL3-2
Text
Text
Text
Text
DDM Amp
FLA
FLW
Text
Text
Trigger Level
Text2
FSC
SSC
FL1
FL2
FL3
DDM Select
Status
Laser Volts
Label15
Laser Amps
Label15
Samp. Volts
Label15
DDM
Label15
Set
Reset
Save
Done

```
FORM1.FRM - 1

Option Explicit


Function vbgenproc ()

End Function

Sub AssaySel_Click ()
Dim y As Integer
Dim x As Integer

'find assay record selected

data2.Refresh
Do While Not (AssaySel.Text = data2.Recordset("assay_name"))
    data2.Recordset.MoveNext
Loop

'get number of events required by this assay

events = data2.Recordset("event_count")

'clear the testdef table;  it holds the name of each token (biomolocule
' assayed) and the base green and over under values

For x = 0 To 1023
For y = 0 To 1
testdef(y, x) = 0
Next y
Next x

datatst.Refresh

'now load the new values for this assay

Do While Not datatst.Recordset.EOF
    If AssaySel.Text = datatst.Recordset("assay") Then
    x = datatst.Recordset(2)  'token value
    tkname(x) = datatst.Recordset("token_name")
    For y = 0 To 1
    testdef(y, x) = datatst.Recordset(y + 4) 'load over under green and base gr
en
    Next y

    End If
    datatst.Recordset.MoveNext

Loop

If x = 0 Then MsgBox "There are no Measurement Parameters defined for this assa
", 48, "Assay Select"
lasttst = x

End Sub

Sub CubeSel_Click ()
Dim y As Integer
Dim x As Integer
data2.Refresh
AssaySel.Clear
Do While Not data2.Recordset.EOF
```

```
FORM1.FRM - 2

    If CubeSel.Text = data2.Recordset("cube_name") Then AssaySel.AddItem data2.
ecordset("assay_name")
    data2.Recordset.MoveNext

Loop

y = ReadPanel(tbuf(0))

loadpbuf
For x = 13 To 42
If pbuf(x) <> tbuf(x) Then y = SendPanel(x, pbuf(x))
If y > 30 Then MsgBox "Flow Cytometer is not responding", 48, "Set Flow Cytomet
r"
Next x
datahld.Refresh
x = 0

Do While Not datahld.Recordset.EOF
    If CubeSel.Text = datahld.Recordset("cube_name") Then
    For y = 0 To 7
    hldtab(y, x) = datahld.Recordset(y + 3)
    Next y
    x = x + 1
    End If
    datahld.Recordset.MoveNext

Loop
If x = 0 Then MsgBox "There are no Classification Parameters defined for this c
be", 48, "Cube Select"
lastnode = x


End Sub

Sub Form_Load ()
Dim x As Integer
'x = InitBrd()
data1.Refresh

Do While Not data1.Recordset.EOF
    CubeSel.AddItem data1.Recordset("cube_name")
    data1.Recordset.MoveNext
Loop

state = 0
grid1.ColWidth(0) = 2400
grid1.ColWidth(1) = 2600
For x = 1 To 63
grid1.RowHeight(x) = 500
Next x
grid1.Row = 0
grid1.Col = 0
grid1.Text = "Biomolecule Assayed"
grid1.Col = 1
grid1.Text = "Result of Assay"
End Sub

Sub MCAdjust_Click ()
If form1!CubeSel.Text <> "" Then Form2.Show

End Sub
```

```
FORM1.FRM - 3


Sub MCCalib_Click ()
OpInst.Text = "Load Calibration Beads into FACS"


End Sub

Sub MCEnd_Click ()
End

End Sub

Sub OpInstOK_Click ()
state = state + 1
Select Case state
Case 0

Case 1
OpInst.Text = "Calibration Complete, Select Cube and Assay"
Case Else
OpInst.Text = ""
TCInit.Enabled = True
End Select

End Sub

Sub ResClear_Click ()
Dim x
For x = 1 To 12
grid1.Row = x
grid1.Col = 0
grid1.Text = ""
grid1.Col = 1
grid1.Text = ""
Next x
grid1.Row = 0
TCStart.Enabled = False

End Sub

Sub TCHalt_Click ()

' during development, the halt button displays the results of classification
' by token, **** token 0 is the reject class

Dim x As Integer
Dim y As Integer
Dim z As Integer
Dim measure As Integer

x = DoTest(lbuf(0, 0), hldtab(0, 0), testdef(0, 0), results(0, 0), events, last
ode, lasttst)

For y = 0 To 1023
If results(2, y) <> 0 Then
            grid1.Row = grid1.Row + 1
            grid1.Col = 0
            grid1.Text = "bead " & y
            grid1.Col = 1
            grid1.Text = results(2, y)
End If
```

```
FORM1.FRM - 4

Next y


End Sub

Sub TCInit_Click ()
Dim x As Integer
Dim y As Integer
'x = InitBrd()
If AssaySel.Text = "" Then
OpInst.Text = "You must select a Cube and Assay first!"
Else
For x = 0 To 1023
For y = 0 To 3
results(y, x) = 0
Next y
Next x

OpInst.Text = "Initialization Complete"
TCStart.Enabled = True
TCHalt.Enabled = True
End If


End Sub

Sub TCStart_Click ()
Dim x As Integer
Dim y As Integer
Dim z As Integer
Dim measure As Integer

x = DoTest(lbuf(0, 0), hldtab(0, 0), testdef(0, 0), results(0, 0), events, last
ode, lasttst)

dataint.Refresh
Do While Not dataint.Recordset.EOF
    If AssaySel.Text = dataint.Recordset("assay") Then
        For x = 0 To 4
        resline(x) = dataint.Recordset(x + 2)'move to temp area
        Next x

        z = resline(0)  'token value
        'test*************************************

        results(2, z) = results(2, z) + 1
        results(0, z) = results(0, z) + 1

        If resline(2) = 0 Then
            measure = results(3, z) / results(2, z) 'sum green over total coun
        Else measure = results(1, z) / results(0, z)'over count divided by und
        End If

        If ((measure >= resline(3)) And (measure <= resline(4))) Then
            grid1.Row = grid1.Row + 1
            grid1.Col = 0
            grid1.Text = tkname(z)
            grid1.Col = 1
            grid1.Text = dataint.Recordset("interpretation")
        End If
```

```
FORM1.FRM - 5

    End If
    dataint.Recordset.MoveNext
Loop



graph1.RandomData = 1
graph1.Refresh
state = 0
'AssaySel.Clear
OpInst.Text = "Test Complete"

End Sub

Sub Timer1_Timer ()
Dim x, y, z  As Integer
y = ReadPanel(tbuf(0))

    x = tbuf(38)
    ' set Waste Water indicator
    y = x And 1
    z = 1
    If y = 1 Then z = 0
    WWPic(z).Visible = False
    WWPic(y).Visible = True

    ' set Sheath Fluid
    y = (x And 2) / 2
    z = 1
    If y = 1 Then z = 0
    SFPic(z).Visible = False
    SFPic(y).Visible = True

    'set pressure
    y = (x And &H80) / &H80
    z = 1
    If y = 1 Then z = 0
    PRPic(z).Visible = False
    PRPic(y).Visible = True

    'set flow
    y = x And &H70  'mask bits
    Select Case y
    Case &H10
        FCPic(0).Visible = True
        FCPic(1).Visible = False
        FCPic(2).Visible = False
    Case &H20
        FCPic(0).Visible = False
        FCPic(1).Visible = False
        FCPic(2).Visible = True
    Case &H40
        FCPic(0).Visible = False
        FCPic(1).Visible = True
        FCPic(2).Visible = False
    Case 0
        FCPic(0).Visible = False
        FCPic(1).Visible = False
```

```
FORM1.FRM - 6
        FCPic(2).Visible = True
    End Select


End Sub
```

```
FORM2.FRM - 1

Option Explicit
Dim photo(5)  As Integer
Dim ampnum(5) As Integer

Dim mode(5) As Integer
Dim thresh(5) As Integer
Dim fcomp(5) As Integer
Dim trigsav As Integer


Sub Check3D1_Click (Value As Integer)
pbuf(39) = Value

End Sub

Sub Check3D2_Click (Value As Integer)
pbuf(40) = 2 * (Value + 1)

End Sub

Sub setpvals ()
' this sets the manual adjust screen to reflect
' the current values in pbuf

Dim x As Integer
For x = 0 To 4
txtNumber(x) = pbuf(x + 13)
amp(x) = pbuf(x + 18)
Next x

fscmod(pbuf(23)).Value = True
sscmod(pbuf(24)).Value = True
fl1mod(pbuf(25)).Value = True
fl2mod(pbuf(26)).Value = True
fl3mod(pbuf(27)).Value = True

For x = 29 To 33
If pbuf(x) <> 0 Then
trigval = pbuf(x)
trigger(x - 29).Value = True
End If
Next x

For x = 0 To 3
text1(x) = pbuf(x + 34)
Next x

ddmtxt(0) = pbuf(41)
ddmtxt(1) = pbuf(42)

If pbuf(40) <> 0 Then option1(pbuf(40) - 2).Value = True

stat(0).Caption = Format$(pbuf(10) * .05, "###.00")
stat(1).Caption = Format$(pbuf(11) * .02, "###.00")
stat(2).Caption = Format$(pbuf(12) / 100, "###.00")

If pbuf(39) = 1 Then
stat(3).Caption = "Enabled"
Else stat(3).Caption = "Disabled"
End If
```

FORM2.FRM - 2

```
End Sub

Sub amp_Change (Index As Integer)
If Val(amp(Index).Text) > 999 Then amp(Index).Text = "999"
pbuf(Index + 18) = Val(amp(Index).Text)
amp(Index).Text = Format(pbuf(Index + 18))

End Sub

Sub Command3D1_Click ()
form2.Hide
End Sub

Sub Command3D2_Click ()
Dim x As Integer
Dim y As Integer
Dim z As Integer
y = ReadPanel(tbuf(0))
For x = 13 To 42
If pbuf(x) <> tbuf(x) Then y = SendPanel(x, pbuf(x))
Next x
End Sub

Sub Command3D3_Click ()
Dim y As Integer
Dim x As Integer
loadpbuf
setpvals
For x = 13 To 42
y = SendPanel(x, pbuf(x))
If y > 30 Then MsgBox "Flow Cytometer is not responding", 48, "Set Flow Cytomet
r"
Next x

End Sub

Sub Command3D4_Click ()
savepbuf
End Sub

Sub Command3D5_Click ()
form2.Hide

End Sub

Sub ddmscroll_Change (Index As Integer)
pbuf(Index + 41) = ddmscroll(Index).Value
ddmtxt(Index) = Format(ddmscroll(Index))

End Sub

Sub ddmtxt_Change (Index As Integer)
If Val(ddmtxt(Index)) > 999 Then ddmtxt(Index) = "999"
ddmscroll(Index).Value = Val(ddmtxt(Index))
pbuf(Index + 41) = ddmscroll(Index).Value

End Sub
```

```
FORM2.FRM - 3

Sub fl1mod_Click (Index As Integer, Value As Integer)
pbuf(25) = Index
If Index = 0 Then
    spina(2).Enabled = False
    amp(2).Enabled = False
Else
    spina(2).Enabled = True
    amp(2).Enabled = True
End If

End Sub

Sub fl2mod_Click (Index As Integer, Value As Integer)
pbuf(26) = Index
If Index = 0 Then
    spina(3).Enabled = False
    amp(3).Enabled = False
Else
    spina(3).Enabled = True
    amp(3).Enabled = True
End If

End Sub

Sub fl3mod_Click (Index As Integer, Value As Integer)
pbuf(27) = Index
If Index = 0 Then
    spina(4).Enabled = False
    amp(4).Enabled = False
Else
    spina(4).Enabled = True
    amp(4).Enabled = True
End If

End Sub

Sub Form_Load ()
Dim x As Integer
Dim y As Integer
loadpbuf
setpvals
form2.Caption = form1!CubeSel.Text & " Cube Manual Adjust"


End Sub

Sub fscmod_Click (Index As Integer, Value As Integer)
pbuf(23) = Index
If Index = 0 Then
    spina(0).Enabled = False
    amp(0).Enabled = False
Else
    spina(0).Enabled = True
    amp(0).Enabled = True
End If



End Sub

Sub HScroll1_Change (Index As Integer)
```

FORM2.FRM - 4

```
pbuf(Index + 34) = HScroll1(Index).Value
text1(Index) = Format(HScroll1(Index))
End Sub

Sub Option1_Click (Index As Integer)
pbuf(40) = Index + 2
pbuf(39) = 1
End Sub

Sub Spin1_SpinDown (Index As Integer)
Dim min
pbuf(Index + 13) = pbuf(Index + 13) - 1
If Index = 0 Then min = 0 Else min = 150
If pbuf(Index + 13) < min Then pbuf(Index + 13) = min
txtNumber(Index).Text = Format(pbuf(Index + 13))
End Sub

Sub Spin1_SpinUp (Index As Integer)
Dim max
pbuf(Index + 13) = pbuf(Index + 13) + 1
If Index = 0 Then max = 4 Else max = 999
If pbuf(Index + 13) > max Then pbuf(Index + 13) = max
txtNumber(Index).Text = Format(pbuf(Index + 13))

End Sub

Sub Spina_SpinDown (Index As Integer)
Dim min
pbuf(Index + 18) = pbuf(Index + 18) - 1
min = 100
If pbuf(Index + 18) < min Then pbuf(Index + 18) = min
amp(Index).Text = Format(pbuf(Index + 18))

End Sub

Sub Spina_SpinUp (Index As Integer)
Dim max
pbuf(Index + 18) = pbuf(Index + 18) + 1
max = 999
If pbuf(Index + 18) > max Then pbuf(Index + 18) = max
amp(Index).Text = Format(pbuf(Index + 18))

End Sub

Sub sscmod_Click (Index As Integer, Value As Integer)
pbuf(24) = Index
If Index = 0 Then
    spina(1).Enabled = False
    amp(1).Enabled = False
Else
    spina(1).Enabled = True
    amp(1).Enabled = True
End If

End Sub

Sub Text1_Change (Index As Integer)
If Val(text1(Index)) > 999 Then text1(Index) = "999"
HScroll1(Index).Value = Val(text1(Index))
pbuf(Index + 34) = HScroll1(Index).Value
```

```
FORM2.FRM - 5

End Sub

Sub trigger_Click (Index As Integer, Value As Integer)
Dim x
For x = 29 To 33
pbuf(x) = 0
Next x
pbuf(Index + 29) = Val(trigval.Text)
trigsav = Index
pbuf(28) = 128 + Index
End Sub

Sub trigval_Change ()
If Val(trigval.Text) > 999 Then trigval.Text = "999"
VScroll1.Value = Val(trigval.Text)
pbuf(trigsav + 29) = VScroll1.Value

End Sub

Sub VScroll1_Change ()
trigval.Text = Format(VScroll1.Value)
pbuf(trigsav + 29) = VScroll1.Value

End Sub
```

```
MODULE1.BAS - 1

Option Explicit
Global pbuf(64) As Integer
Global tbuf(64) As Integer

Global lbuf(8, 32) As Integer
Global hldtab(7, 1023) As Integer
Global results(3, 1023) As Long
Global testdef(1, 1023) As Integer
Global resline(5) As Integer
Global tkname(1024) As String
Global state As Integer
Global lastnode As Integer
Global lasttst As Integer
Global events As Long

Declare Function DoTest Lib "c:\msvc\bin\orbit.dll" (lbuf As Integer, hldtab As
Integer, testdef As Integer, results As Long, ByVal events&, ByVal lastnode%, B
Val lasttst%) As Integer
Declare Function InitBrd Lib "c:\msvc\bin\orbit.dll" () As Integer
Declare Function SendPanel Lib "c:\msvc\bin\orbit.dll" (ByVal parm%, ByVal panv
l%) As Integer
Declare Function ReadPanel Lib "c:\msvc\bin\orbit.dll" (pbuf As Integer) As Int
ger
Declare Function ReadList Lib "c:\msvc\bin\orbit.dll" (lbuf As Integer) As Inte
er


Sub loadpbuf ()
' for a selected cube, the data base values are loaded into pbuf

Dim x As Integer
form2!facset.Refresh
Do While form2!facset.Recordset.EOF = False
If form1!CubeSel.Text = form2!facset.Recordset(0) Then Exit Do
form2!facset.Recordset.MoveNext
Loop

For x = 0 To 42
pbuf(x) = form2!facset.Recordset(x + 1)
Next x

End Sub

Sub savepbuf ()
' This saves the current values in pbuf to the data base

Dim x As Integer
form2!facset.Refresh
Do While form2!facset.Recordset.EOF = False
If form1!CubeSel.Text = form2!facset.Recordset(0) Then Exit Do
form2!facset.Recordset.MoveNext
Loop
form2!facset.Recordset.Edit
For x = 0 To 42
form2!facset.Recordset(x + 1) = pbuf(x)
Next x
form2!facset.Recordset.Update
End Sub
```

```
MODULE2.BAS - 1
Option Explicit
```

```
////////////////////////////////////////////////////////
//
//  MDXDLL.DLL source.
//
////////////////////////////////////////////////////////

#include <windows.h>
#include "mdxdll.h"

int WINAPI _export InitBrd(void)
{
__asm
        {
        mov     dx,brdctrl
        mov     ax,20h
        out     dx,ax
        mov     dx,brdctrl
        in      ax,dx
        or      ax,outfifoclr   ;set both fifos to clear
        or      ax,infifoclr
        out     dx,ax

        mov     dx,inpxfer
        mov     ax,0    ;set xfer count to 0
        out     dx,ax

        mov     dx,prtctrl
        mov     ax,0
        or      ax,ctl0 ;set control lines to idle
        or      ax,ctl1
        out     dx,ax


        mov     dx,prtctrl
        in      ax,dx
        mov     bx,preset       ;reset facs
        not     bx
        and     ax,bx
        out     dx,ax

        or      ax,preset       ;toggle
        out     dx,ax

        mov     dx,brdsts
        in      ax,dx

        mov     dx,inpxfer
        mov     ax,0
        out     dx,ax
        mov     dx,prtctrl
        mov     ax,6    ;assert ctl0 &1
        out     dx,ax

        // asm code here

        }

//SendPanel(ddmena,1);
return(0);
}

int WINAPI _export SendPanel(int parm, int PanVal)
```

```
{
int i = 0;

__asm
        {
        mov     dx,brdctrl
        in      ax,dx
        mov     bx,inen
        not     bx
        and     ax,bx
        or      ax,outen
        out     dx,ax

        mov     dx,prtctrl
        in      ax,dx
        mov     bx,ctl0
        not     bx
        and     ax,bx
        out     dx,ax

        mov     dx,brdctrl
        in      ax,dx
        or      ax,outfifoclr
        out     dx,ax
        mov     bx,outfifoclr
        not     bx
        and     ax,bx
        out     dx,ax

        mov     dx,bdata
        mov     ax,1
        out     dx,ax

        mov     bx,parm         ;get code
        shl     bx,10
        mov     ax,PanVal       ;get panel value
        and     ax,3ffh
        or      ax,bx
        out     dx,ax   ;send it out

        mov     dx,dlyctrl
        mov     ax,0160h
        out     dx,ax

        mov     dx,brdctrl
        in      ax,dx
        or      ax,hshken
        out     dx,ax
        mov     cx,4000

wsend:  mov     dx,brdsts
        in      ax,dx
        mov     bx,ax
        and     bx,1    ;1=empty fifo
        loope   wsend
        and     ax,40h  ;bit 6 = 1 = xfer pending
        jnz         wsend
        mov     dx,prtctrl
        in      ax,dx
        or      ax,ctl0
        out     dx,ax
```

```
        mov     dx,brdctrl
        in      ax,dx
        mov     bx,outen
        or      bx,hshken
        not     bx
        and     ax,bx
        out     dx,ax

        mov     dx,dlyctrl
        mov     ax,0101h
        out     dx,ax
        // asm code here

        }
return(i);

}


int WINAPI _export ReadPanel(int __far *pbuf)
{
int far* pbufptr = pbuf;

__asm
        {
        mov     dx,brdctrl
        in      ax,dx
        or      ax,outfifoclr   ;set both fifos to clear
        or      ax,infifoclr
        out     dx,ax

        mov     dx,inpxfer
        mov     ax,0    ;set xfer count to 0
        out     dx,ax

        mov     dx,prtctrl
        mov     ax,0
        or      ax,ctl0 ;set control lines to idle
        or      ax,ctl1
        out     dx,ax


        mov     dx,prtctrl
        in      ax,dx
        or      ax,preset       ;toggle
        out     dx,ax

lowlp:  mov     dx,prtsts
        in      ax,dx
        mov     bx,ax
        and     ax,stsl ;has stsl gone low yet?
        jz      waithigh        ;yes, wait for it to go back high
        jmp     lowlp   ;***************

waithigh:

        mov     dx,brdctrl      ;no, panel coming so set up read
        in      ax,dx
        and     al,0eeh ;handsk off out off
        or      ax,8    ;clr in fifo
        out     dx,ax
```

```
        and     al,0f7h  ;fifo clr off
        out     dx,ax

        mov     dx,inpxfer      ;set up xfer count
        mov     ax,43
        out     dx,ax

        mov     dx,prtctrl
        in      ax,dx
        mov     bx,ctl0
        not     bx
        and     ax,bx
        out     dx,ax   ;set ctl0 to request panel data

        mov     dx,dlyctrl
        mov     ax,101h
        out     dx,ax

        mov     dx,brdctrl
        in      ax,dx
        or      ax,hshken       ;initiate handshake
        out     dx,ax

waitdata:
        mov     dx,brdsts
        mov     bx,xferincom
        or      bx,xferpend
        in      ax,dx
        and     ax,bx   ;wait until xfer not pending
        jnz     waitdata        ;or incomplete

        mov     dx,prtctrl  ;done
        in      ax,dx
        or      ax,ctl0 ;stop
        out     dx,ax
        mov     dx,brdctrl
        in      ax,dx
        and     al,0ebh ;turn off handshake and input enable
        out     dx,ax
        push    di
        push    es
        mov     cx,43
        les     di,pbufptr
        mov     dx, bdata
readlp:
        in      ax,dx   ;read data from fifo
        mov     bx,ax
        and     ax,3ffh
        and     bx,0fc00h
        shr     bx,9    ;right 10 * 2
        mov     word ptr es:[di+bx],ax
        loop    readlp
        pop             es
        pop             di

        }
return(0);
}

int WINAPI _export DoTest(int __far *lbuf, int __far *hldtab,
                          int __far *testdef, long __far *results,
                          long events, int lastnode, int lasttst)
```

```
{

int far* lbufptr = lbuf;
int far* hldptr = hldtab;
int far* testptr = testdef;
long far* resptr = results;

int x,y,z,bdi,goodcnt,token = 0;

while (events > 0)
        {
        goodcnt = ReadList(lbufptr);
        if (goodcnt > 0)
                {
                for (bdi = 0; bdi < goodcnt; bdi++)
                        {
                        x = y = 0;
                        // orange correction goes here for lbufptr[bdi][fl2]
                        // x is the current node in the hld table
                        token = -1;
                        while (token < 0)
                                {
                                z = lbufptr[bdi * 8 + (hldptr[x * 8])];
                                // z is the value of the parameter under test
                                if ((z >= hldptr[x * 8 + 2]) && (z <= hldptr[x * 8 + 3])) //low
 val && high val
                                        {
                                        if (hldptr[x * 8 + 4] == 0) token = hldptr[x * 8 + 6]; /
 / 0 means done, get token true
                                        else x = hldptr[x * 8 + 4]; // get node true
                                        }
                                else
                                        {
                                        if (hldptr[x * 8 + 5] == 0) token = hldptr[x * 8 + 7];
 / get token false
                                        else x = hldptr[x * 8 + 5]; // get node false
                                        }
                                }
                                events--;
                                z = lbufptr[bdi * 8 + 2]; // z is FL1
                                if (z < testptr[token * 2 + 1]) resptr[token * 4]++; //inc unde
  count
                                if (z > testptr[token * 2 + 1]) resptr[token * 4 + 1]++; //inc
 ver count
                                resptr[token * 4 + 2]++;  // inc total count
                                resptr[token * 4 + 3] += z;  //sum FL1
                        }
                }

        }


  __asm
        {
         nop
        }

 return(0);
 }

 int WINAPI _export ReadList(int __far *lbuf)
```

```
{
int far* lbufptr = lbuf;
int i = 0;

__asm
        {
        ;*****************************************************************
        ;                read list mode
        ;*****************************************************************

        mov     dx,brdctrl
        in      ax,dx
        or      ax,outfifoclr   ;set both fifos to clear
        or      ax,infifoclr
        out     dx,ax

        mov     dx,inpxfer
        mov     ax,0    ;set xfer count to 0
        out     dx,ax

        mov     dx,prtctrl
        mov     ax,0
        or      ax,ctl0 ;set control lines to idle
        or      ax,ctl1
        out     dx,ax


        mov     dx,prtctrl
        in      ax,dx
        or      ax,preset       ;toggle
        out     dx,ax

lowlpl: mov     dx,prtsts
        in      ax,dx
        mov     bx,ax
        and     ax,sts0 ;has sts0 gone low yet?
        jz      waithighl       ;yes, wait for it to go back high
        jmp     lowlpl  ;  ********

waithighl:

        mov     dx,brdctrl      ;no, list mode coming so set up read
        in      ax,dx
        and     al,0aeh ;handsk off out off
        or      ax,8    ;clr in fifo
        out     dx,ax

        and     al,0f7h  ;fifo clr off
        out     dx,ax

        push    es
        push    di
        push    ds
        push    si
        les     di,lbufptr


        mov     dx,inpxfer      ;set up xfer count
        mov     ax,120  ;7 vals + chk sum * 15 events
        out     dx,ax

        mov     dx,prtctrl
        in      ax,dx
```

```
        mov     bx,ctl1
        not     bx
        and     ax,bx
        out     dx,ax   ;set ctl1 to request list data

        mov     dx,dlyctrl
        mov     ax,101h
        out     dx,ax

        mov     dx,brdctrl
        in      ax,dx
        or      ax,hshken       ;initiate handshake
        out     dx,ax

waitdata1:
        mov     dx,brdsts
        mov     bx,xferincom
        or      bx,xferpend
        in      ax,dx
        and     ax,bx   ;wait until xfer not pending
        jnz     waitdata1       ;or incomplete
        mov     dx,bdata
        mov     cx,120  ;number of words to read
chunk:  in      ax,dx
        stosw
        loop    chunk
        mov     dx,prtctrl  ;done
        in      ax,dx
        or      ax,ctl1 ;stop
        out     dx,ax
        mov     dx,brdctrl
        in      ax,dx
        and     al,0ebh ;turn off handshake and input enable
        out     dx,ax


        les     di,lbufptr
        lds     si,lbufptr
        mov     cx,15   ;up to 15 good events could be present

        ;check alignment

        mov     dx,0    ;count of good records


        test    word ptr [si],0e000h
        jz      alnok
        dec     cx      ;only 14 max possible now
        push    cx
        mov     cx,7    ;seven other possible alignments
alnlp:  lodsw
        test    ax,0e000h
        jz      nowalnok        ;now have alignment
        loop    alnlp
        pop     cx
        jmp     badbuff
nowalnok:
        sub     si,2    ;repoint to good align
        pop     cx      ;restore loop count
alnok:  push    cx
        mov     cx,7
        mov     bx,0
sumchk: lodsw
```

```
        add     bx,ax
        loop    sumchk
        lodsw
        cmp     bx,ax   ;is check sum good
        jne     nocopy
        sub     si,16   ;repoint to start of rec
        mov     cx,8
clncpy: lodsw
        and     ax,3ffh ;knock off parm number
        stosw
        loop    clncpy
        inc     dx      ;count good records
nocopy: pop     cx
        loop    alnok
badbuff:
        mov     i,dx

        pop     si
        pop     ds
        pop     di
        pop     es




        }
return(i);
}


////////////////////////////////////////////////////////
//
//  DLL Initialization and exit.
//
int WINAPI LibMain(HANDLE hInst, WORD wDataSeg,
                   WORD cbHeapSize, LPSTR lpszCmdLine)
{
  hModInst = hInst;
  if (cbHeapSize != 0)
    UnlockData(0);
  return 1;
}


int _export WINAPI WEP(int nParam)
{
  return 1;
}
```

```
////////////////////////////////////////////////////////
//
//  Include file for MDXDLL.DLL
//
////////////////////////////////////////////////////////

#ifdef __cplusplus
extern "C" {
#endif


////////////////////////////////////////////////////////
//
// Define DEBUG macros.
//
#if defined(_DEBUG) && !defined(_AFX)

char _sz_ASSERT[255];

#define ASSERT(a)  if(!(a)) { wsprintf(_sz_ASSERT,  \
   "assertion failed in file %s at line %d\r\n",      \
   (LPSTR)(__FILE__),__LINE__); OutputDebugString(_sz_ASSERT); }
#define TRACE(a)   OutputDebugString(a"\r\n")
#define TRACE1(a,b)     { wsprintf(_sz_ASSERT,a"\r\n",(int)(b)); \
   OutputDebugString(_sz_ASSERT); }
#define TRACE2(a,b,c)  { wsprintf(_sz_ASSERT,a"\r\n",(int)(b),(int)(c)); \
   OutputDebugString(_sz_ASSERT); }

#endif

#if !defined(_DEBUG) && !defined(_AFX)

#define ASSERT(a)
#define TRACE(a)
#define TRACE1(a,b)
#define TRACE2(a,b,c)

#endif

////////////////////////////////////////////////////////
//
// From Windowsx.h
//
#define     GlobalPtrHandle(lp)          \
        ((HGLOBAL)LOWORD(GlobalHandle(SELECTOROF(lp))))

#define     GlobalLockPtr(lp)        \
        ((BOOL)SELECTOROF(GlobalLock(GlobalPtrHandle(lp))))
#define     GlobalUnlockPtr(lp)      \
        GlobalUnlock(GlobalPtrHandle(lp))

#define     GlobalAllocPtr(flags, cb)    \
        (GlobalLock(GlobalAlloc((flags), (cb))))
#define     GlobalReAllocPtr(lp, cbNew, flags)  \
        (GlobalUnlockPtr(lp), GlobalLock(GlobalReAlloc(GlobalPtrHandle(lp) , (cbNew), (
lags))))
#define     GlobalFreePtr(lp)        \
        (GlobalUnlockPtr(lp), (BOOL)GlobalFree(GlobalPtrHandle(lp)))

////////////////////////////////////////////////////////
```

```
//
// Global variables.
//

HINSTANCE hModInst;      // module handle.


// Cytomation board definitions

#define base     0x240
#define bdata    base + 0
#define brdctrl base + 2
#define brdsts   base + 4
#define prtctrl base + 6
#define prtsts   base + 8
#define dlyctrl base + 0x0a
#define inpxfer base + 0x0c

        //facs codes

#define ddmena  39

        //board control

#define outen   1
#define outfifoclr  2
#define inen    4
#define infifoclr   8
#define hshken      0x10

        //board status

#define outfifoe    1
#define outfifoh    2
#define outfifof    4
#define infifoe     8
#define infifoh     0x10
#define infifof     0x20
#define xferpend    0x40
#define xferincom   0x80

        //port control

#define preset      1
#define ctl0        2
#define ctl1        4

        //port status

#define eir         1
#define pets        2
#define sts0        4
#define sts1        8

////////////////////////////////////////////////////////////
//
// Exported Functions.
//

////////////////////////////////////////////////////////////
```

```
//
// Local Functions.
//

int WINAPI _export SendPanel(int parm, int PanVal);
int WINAPI _export ReadPanel(int pbuf[]);
int WINAPI _export ReadList(int lbuf[]);
int WINAPI _export InitBrd (void);
int WINAPI _export DoTest(int lbuf[], int hldtab[],
                          int testdef[], long results[],
                          long events, int lastnode, int lasttst);

#ifdef __cplusplus
}
#endif
```

```
;-----------------------------------------------------------
;
; Module Definition file for MDXDLL.DLL
;
;-----------------------------------------------------------
LIBRARY   MDXDLL

DESCRIPTION  'MDXDLL Library'

EXETYPE   WINDOWS

CODE      PRELOAD MOVABLE
DATA      PRELOAD MOVABLE

HEAPSIZE  1024

EXPORTS
  WEP                 @1 RESIDENTNAME
  InitBrd             @2
  SendPanel           @3
  ReadPanel           @4
  ReadList            @5
  DoTest              @6
```

WHAT IS CLAIMED IS:

1. A method of preparing a beadset capable of detecting a plurality of analytes in a single fluid sample by flow cytometric analysis comprising:

(a) obtaining a plurality of subsets of beads wherein the beads in each subset are sufficiently homogeneous with respect to at least three selected classification parameters (C_1, C_2, C_3 ... C_n) and sufficiently different in at least one of said classification parameters from beads in any other subset so that the profile of classification parameter values within each subset detectable by flow cytometry is unique;

(b) coupling the beads within each subset to a reactant that will specifically react with a given analyte of interest in a fluid to be tested; and

(c) mixing the subsets of beads to produce a beadset, wherein the subset identity and therefore the reactant to which the bead has been coupled is identifiable by flow cytometry based on the unique classification parameter profile of the beads.

2. A beadset capable of detecting a plurality of analytes in a single fluid sample by flow cytometric analysis comprising a plurality of subsets of beads wherein:

(a) the beads in each subset are sufficiently homogeneous with respect to at least three selected classification parameters (C_1, C_2, C_3 ... C_n) and sufficiently different in at least one of said classification parameters from beads in any other subset so that the profile of classification parameter values within each subset detectable by flow cytometry is unique;

(b) wherein the beads within each subset are coupled to a reactant that will specifically react with a given analyte of interest in a fluid to be tested; and

(c) wherein said subsets have been mixed to produce the beadset, characterized in that the subset identity and therefore the reactant to which the bead has been coupled is identifiable based on the unique classification parameter profile of the bead.

3. A method of flow cytometric analysis capable of detecting a plurality of analytes of interest in a single fluid sample comprising:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to a reactant that will specifically react with a selected analyte of interest in a fluid to be tested;

(b) mixing, to produce a reacted bead sample, the beadset with the fluid to be tested under conditions that will allow reactions between analytes of interest in the fluid and the reactants on the beads in said set, wherein a reaction between a reactant

and an analyte of interest on a bead causes a change in the value of a fluorescent signal (F_m) emitted from said bead;

(c) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(d) identifying the subset to which each bead belongs and therefore the reactant on the bead as a function of the unique profile of classification parameter values; and

(e) detecting the presence or absence of a particular analyte of interest in said sample as a function of the identification in step (d) and a change in the F_m values of the beads in each of said subsets in the reacted fluid sample relative to the F_m values of the beads in each of said subsets not reacted with said fluid.

4. A method of flow cytometric analysis capable of detecting a plurality of analytes of interest in a single fluid sample comprising:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to a reactant that will specifically react with a selected analyte of interest in a fluid to be tested;

(b) mixing, to produce a reacted bead sample, the beadset with the fluid to be tested under conditions that will allow reactions between analytes of interest in the fluid and the reactants on the beads in said set;

(c) mixing with the reacted bead sample a fluorescent label under conditions such that said label will bind to and thereby increase the value of a fluorescent signal F_m emitted from said bead;

(d) analyzing the reacted sample containing the fluorescent label by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(e) identifying the subset to which each bead belongs and therefore the reactant on the bead as a function of the unique profile of classification parameter values; and

(f) detecting the presence or absence of a particular analyte of interest in said sample as a function of the identification in step (e) and an increase in the F_m values of the beads in each of said subsets in the reacted fluid sample relative to the F_m values of the beads in each of said subsets not reacted with said fluid.

5. A method of flow cytometric analysis capable of detecting a plurality of analytes of interest in a single sample comprising:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said

classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique,

(2) are coupled to a reactant that will specifically react with a selected analyte of interest in a fluid to be tested, and

(3) are reacted with a fluorescently labeled compound which competes with said analyte for reaction with said reactant;

(b) mixing, to produce a reacted bead sample, the beadset with the fluid to be tested under conditions that will allow reactions between analytes of interest in the fluid and the reactants on the beads in said set and thereby to allow the analytes to competitively inhibit or displace the fluorescently labeled compounds from said beads, resulting in a decrease in a fluorescent signal F_m emitted from a bead with which an analyte of interest in the fluid has reacted;

(c) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(d) identifying the subset to which each bead belongs and therefore the reactant on the bead as a function of the unique profile of classification parameter values; and

(e) detecting the presence or absence of a particular analyte of interest in said sample as a function of the identification in step (d) and an increase in the F_m values of the beads in each of said subsets in the reacted fluid sample relative to the F_m values of the beads in each of said subsets not reacted with said fluid.

6. The method of any one of claims 3, 4, and 5 wherein C_1, C_2, and C_3 are each different and are selected from the group consisting of forward light scatter, side light scatter and fluorescence.

7. The method of any one of claims 3, 4, and 5 wherein n is greater than or equal to 4 and C_1 is forward angle light scatter, C_2 is side angle light scatter, C_3 is fluorescence at a first wavelength and C_4 is fluorescence at a second wavelength.

8. The method of claim 7 wherein said first wavelength is red and said second wavelength is orange.

9. The method of claim 7 wherein said first wavelength is red, said second wavelength is orange, and the wavelength of said F_m signal is green.

10. The method of claim 3 wherein said analytes of interest are antigens and said reactants are antibodies specifically reactive with said antigens.

11. The method of claim 3 wherein said analytes of interest are antibodies and said reactants are antigens specifically reactive with said antibodies.

12. The method of claim 3 wherein said analytes of interest are antigens selected from the group consisting of bacterial, viral, fungal, mycoplasmal, rickettsial, chlamydial and protozoal antigens and said reactants are antibodies specifically reactive with said antigens.

13. The method of claim 3 wherein said reactants are antigens selected from the group consisting of bacterial, viral, fungal, mycoplasmal, rickettsial, chlamydial and protozoal

antigens and said analytes of interest are antibodies specifically reactive with said antigens.

14. The method of any one of claims 10 or 12 wherein said antigens are antigens borne by pathogenic agents responsible for sexually transmitted disease.

15. The method of any of claims 10 or 12 wherein said antigens are antigens borne by pathogenic agents responsible for a pulmonary disorder.

16. The method of any of claims 10 or 12 wherein said antigens are antigens borne by pathogenic agents responsible for a gastrointestinal disorder.

17. The method of claim 3 wherein said analytes of interest are substances of abuse.

18. The method of claim 3 wherein said analytes of interest are therapeutic drugs.

19. The method of claim 3 wherein said analytes of interest are antigens or antibodies associated with one or more selected pathological syndromes.

20. The method of claim 19 wherein said syndromes are selected from the group consisting of malignancy, allergy, autoimmune diseases, and blood borne viruses.

21. The method of claim 19 wherein at least one said syndrome is a cardiovascular disorder.

22. The method of claim 3 wherein said analytes of interest are selected from the group consisting of analytes testing for pregnancies and hormones.

23. The method of claim 3 wherein said fluorescent signal is emitted from fluoresceinated antibodies specific for antibodies coupled to said beads in said set.

24. The method of claim 3 wherein said fluorescent signal is emitted from a fluoresceinated compound specifically reactive with an immunoglobulin molecule.

25. The method of claim 3 wherein said fluorescent signal is emitted from an agent selected from the group consisting of a fluoresceinated anti-immunoglobulin antibody or a specifically reactive fragment thereof, fluoresceinated protein A, and fluoresceinated protein G.

26. The method of claim 19 wherein said analyte comprises autoantibodies and said antigens comprise oligopeptide epitopes reactive with said autoantibodies, said fluorescent labels comprise fluorescent monoclonal antibodies reactive with said epitopes and wherein the presence of the analyte autoantibodies is detected as a result in a decrease of F_m.

27. The method of claim 3 wherein said analytes are enzymes, said reactants are fluorescently labeled substrates for said enzymes, said change in F_m results from cleavage of said substrates from said beads.

28. The method of claim 27 wherein said enzymes are selected from the groups consisting of proteases, glycosidases, nucleotidases, oxidoreductases, hydrolyases, esterases, convertases, ligases, transferases, phosphorylases, lyases, lipases, peptidases, dehydrogenases, oxidases, phospholipases, decarboxylases, invertases, aldolases, transaminases, synthetases, and phosphotases.

29. The method of claim 3 wherein the fluid to be tested is selected from the group consisting of plasma, serum, tears, mucus, saliva, urine, pleural fluid, spinal fluid and gastric fluid, sweat, semen, vaginal secretions, fluid from ulcers and other surface eruptions, blisters, and abscesses, and extracts of tissues including biopsies of normal, malignant, and suspect tissues.

30. A method of flow cytometric analysis for detection of immunoglobulins in a fluid sample comprising the steps of:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to an immunoglobulin that corresponds to the immunoglobulin to be assayed for in the fluid sample;

(b) obtaining a fluorescently labeled immunoglobulin-binding reagent capable of reacting with the immunoglobulins to be detected;

(c) mixing, to produce a reacted bead sample, the beadset with the fluid sample to be tested and the fluorescently labeled immunoglobulin-binding reagent under conditions that will allow competitive binding reactions between the

immunoglobulin-binding reagent and immunoglobulin in the fluid to be tested and between the immunoglobulin-binding reagent and the immunoglobulin on the beads in said set, wherein a reaction between a bead-bound immunoglobulin and the fluorescently labeled immunoglobulin-binding reagent causes an increase in the value of a fluorescent signal (F_m) emitted from said bead;

(d) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(e) identifying the subset to which each bead belongs and therefore the immunoglobulin on the bead as a function of the unique profile of classification parameter values; and

(f) detecting a corresponding immunoglobulin in said sample as a function of the identification in step (e) and a change in the F_m values of the beads in each of said subsets in the reacted fluid sample relative to the F_m values of the beads in each of said subsets not reacted with said fluid.

31. The method of claim 30 wherein said immunoglobulins to be detected are immunoglobulins belonging to different immunoglobulin classes.

32. The method of claim 31 wherein said classes are selected from the group consisting of IgG, IgM, IgA, and IgE.

33. The method of claim 32 wherein said immunoglobulins to be detected are immunoglobulins belonging to different immunoglobulin sub-classes.

34. The method of claim 33 wherein said subclasses are selected from the group consisting of human IgG_1, IgG_2, IgG_3, IgG_4, IgA_1, and IgA_2.

35. A method of flow cytometric analysis for detection of immunoglobulin specific for a particular epitope of interest in a sample comprising the steps of:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to a monoclonal antibody preparation which is specific for an epitope that is the same epitope as that epitope which binds to an immunoglobulin to be assayed for;

(b) obtaining a plurality of fluorescently labeled reagents wherein each of said reagents bears an epitope to which the monoclonal antibody preparation coupled to the beads within a subset binds;

(c) mixing, to produce a reacted bead sample, the beadset with the fluid sample to be tested and the fluorescently labeled reagents under conditions that will allow competitive binding reactions between the fluorescently labeled reagents and immunoglobulin in the fluid to be tested and between the fluorescently labeled

reagents and the monoclonal antibodies on the beads wherein a reaction between a bead-bound antibody and the fluorescently labeled reagent causes an increase in the value of a fluorescent signal (F_m) emitted from said bead;

(d) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(e) identifying the subset to which each bead belongs and therefore the monoclonal antibody on the bead as a function of the unique profile of classification parameter values; and

(f) detecting the presence or absence of an immunoglobulin in said sample specific for said particular epitope as a function of the identification in step (e) and a change in the F_m values of the beads in each of said subsets in the reacted fluid sample relative to the F_m values of the beads in each of said subsets not reacted with said fluid.

36. The method of claim 35 wherein the epitopes are epitopes located on viral antigens.

37. The method of claim 36 wherein said viral antigen is an antigen from HIV.

38. A method of flow cytometric analysis for detection of analytes commonly elevated in pregnancy in a fluid sample comprising the steps of:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently

different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to an antibody which is specific for an analyte to be assayed for in the fluid sample;

(b) obtaining a plurality of preparations of antibody molecules wherein each of said preparations contains fluorescently labeled antibodies specific for an analyte to be assayed for in the fluid sample;

(c) mixing, to produce a reacted bead sample, the beadset with the fluid sample to be tested and the fluorescently labeled antibodies under conditions that will allow binding reactions between the antibody that is coupled to the bead, the analyte of interest in the fluid to be tested, and the fluorescently labeled antibodies so as to bind said fluorescent antibodies to said beads though binding to said enzymes which are in turn bound to said beads though said bead-bound antibodies and wherein a bridging reaction between a bead-bound antibody, the analyte to which that antibody binds, and the fluorescently labeled antibody specific for said enzyme causes an increase in the value of a fluorescent signal (F_m) emitted from said bead;

(d) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(e) identifying the subset to which each bead belongs and therefore the antibody on the bead as a function of the unique profile of classification parameter values; and

(f) detecting the analyte in said sample as a function of the identification in step (e) and a change in the F_m values of the beads in each of said subsets in the reacted fluid sample relative to values of the beads in each of said subsets not reacted with said fluid.

39. The method of claim 38 wherein said analytes are selected from the group consisting of human chorionic gonadotropin, alpha fetoprotein, and 3' estradiol.

40. A method of flow cytometric analysis for determining the epitope to which a monoclonal antibody binds comprising the steps of:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to a peptide which provides a given epitope;

(b) obtaining a fluorescently labeled monoclonal antibody of interest;

(c) mixing, to produce a reacted bead sample, the beadset with the fluorescently labeled monoclonal antibody under conditions that will allow binding reactions

between the bead-bound peptide which provides the epitope to which the monoclonal antibody is capable of binding and said monoclonal antibody, wherein a reaction between a bead-bound peptide and the fluorescently labeled monoclonal antibody causes an increase in the value of a fluorescent signal (F_m) emitted from said bead;

(d) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(e) identifying the subset to which each bead belongs and therefore the peptide on said bead as a function of the unique profile of classification parameter values; and

(f) detecting the particular epitope to which the monoclonal antibody binds as a function of the identification in step (e) and a change in the F_m values of the beads in each of said subsets in the sample relative to the F_m values of beads not reacted with said monoclonal antibody.

41. The method of claim 40 where said peptides are from 2 -100 amino acids in length.

42. A method of flow cytometric assay for antibodies reactive with given pathogens of interest in a fluid sample comprising the steps of:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said

classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to an antigen derived from one of said pathogens of interest :

(b) obtaining a fluorescently labeled immunoglobulin-reactive reagent;

(c) mixing, to produce a reacted bead sample, the beadset with the fluid sample and the fluorescently labeled immunoglobulin-reactive reagent under conditions that will allow binding reactions between the bead-bound antigen and antibody in said sample and the fluorescently labeled immunoglobulin-reactive reagent wherein a reaction between a bead-bound antigen, antibody in said fluid sample and the fluorescently labeled reagent causes an increase in the value of a fluorescent signal (F_m) emitted from said bead;

(d) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(e) identifying the subset to which each bead belongs and therefore the peptide on said bead as a function of the unique profile of classification parameter values; and

(f) detecting the particular epitope to which the monoclonal antibody binds as a function of the identification in step (e) and a change in the F_m values of the beads in each of said subsets in the sample relative to the F_m values of beads not reacted with said fluid sample.

43. The method of claim 42 wherein said antigens comprise one or more of the following antigens: *Toxoplasma gondii*, Rubella virus, Cytomegalovirus, and Herpes Simplex virus.

44. The method of claim 43 wherein said fluorescently labeled immunoglobulin-reactive reagent is anti-Human IgG.

45. The method of claim 43 wherein said fluorescently labeled immunoglobulin-reactive reagent is anti-Human IgM

46. A method of flow cytometric assay for antibodies reactive with allergens of interest in a fluid sample comprising the steps of:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to an antigen derived from an allergen of interest;

(b) obtaining a fluorescently labeled IgE reactive reagent;

(c) mixing, to produce a reacted bead sample, the beadset with the fluid sample and the fluorescently labeled reagent under conditions that will allow binding

reactions between the bead-bound allergen and antibody in said sample and the fluorescently labeled IgE reactive reagent wherein a reaction between a bead-bound allergen, antibody in said fluid sample and the fluorescently labeled IgE-reactive reagent causes an increase in the value of a fluorescent signal (F_m) emitted from said bead;

(d) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(e) identifying the subset to which each bead belongs and therefore the allergen on said bead as a function of the unique profile of classification parameter values; and

(f) detecting the particular epitope to which the monoclonal antibody binds as a function of the identification in step (e) and a change in the F_m values of the beads in each of said subsets in the sample relative to the F_m values of beads not reacted with said fluid sample.

47. The method of claim 46 wherein said allergen comprise one or more of the following antigens: Junegrass, Red Top, Brome, Orchard, Timothy, Rye, Fesque, What, Quack, Bermuda, Johnson, Canary, Velvet, Saltgrass, Bahia, and Vernal.

48. The method of claim 46 wherein said fluorescently labeled IgE reactive reagent is anti-human IgE

49. The method of claim 46 wherein said fluorescently labeled IgE reactive reagent is anti-canine IgE.

50. A method of flow cytometric analysis capable of quantitating the concentration of an analyte of interest in a fluid sample comprising:

(a) obtaining a beadset comprising a plurality of subsets of beads, wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1 , C_2, C_3, ...C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to a reactant that will specifically react with the selected analyte of interest in the sample to be tested; and

wherein the beads in a plurality of said subsets are coupled to the same reactant but at concentrations which differ among said subsets;

(b) mixing, to produce a reacted bead sample, the beadset with the fluid sample to be tested under conditions that will allow reactions between the analyte of interest in the fluid sample and the reactants on the beads in said set, wherein a reaction between a reactant and an analyte of interest on a bead causes a change in the value of a fluorescent signal (F_m) emitted from said bead;

(c) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(d) identifying the subset to which each bead belongs and therefore the concentration of reactant with which the bead was coupled as a function of the unique profile of classification parameter values; and

(e) detecting the concentration of the analyte of interest in said sample as a function of the identification in step (d) and the F_m values of the beads in each of said subsets relative to the F_m values of a second set of the beads in each of said subsets, wherein said beads in said second set have not been reacted with said fluid sample but have been reacted with a known concentration of the analyte of interest.

51. A method of flow cytometric analysis capable of quantitating the concentration of an analyte of interest in a fluid sample comprising:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1 , C_2, C_3, ...C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to a reactant that will specifically react with the analyte of interest in the sample to be tested; and

wherein the beads in a plurality of said subsets are coupled to the same reactant but at concentrations which differ among said subsets;

(b) mixing, to produce a reacted bead sample, the beadset with a fluorescently labeled competitive inhibitor of the reaction between the analyte of interest and the reactant on the beads and with the fluid sample under conditions that will allow reactions between the analyte of interest in the fluid sample and the reactants on the beads in said set, wherein a reaction between an analyte of interest and a reactant on a bead causes a decrease in the value of a fluorescent signal (F_m) emitted from said bead;

(c) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(d) identifying the subset to which each bead belongs and therefore the concentration of reactant with which the bead was coupled as a function of the unique profile of classification parameter values;

(e) assigning a bead subset value to each bead subset with correlates relatively with the concentration of analyte with which the bead subset was coupled;

(f) determining an inter-bead subset slope from a plot of mean F_m for each bead subset versus bead subset value to produce an inter-bead subset slope; and

(g) determining the concentration of the analyte of interest in the sample by interpolation of the slope determined in step (f) into a standard assay curve wherein the inter-bead subset slopes of beads incubated with known

concentrations of the analyte of interest are plotted against the log of the known concentration of the analyte of interest.

52. A method of generating a multiplexed standard assay curve for use in quantitating the concentration of an analyte of interest in a fluid sample comprising the steps of:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in said subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1 , C_2, C_3, ...C_n) values and sufficiently different from beads in any other subset in at least one of said classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique; and

(2) are coupled to a reactant that will specifically react with a selected analyte of interest in a fluid to be tested; and

(3) wherein the beads in a plurality of said subsets are coupled to the same reactant but at concentrations which differ among said subsets;

(b) mixing, to produce a reacted bead sample, the beadset with a fluorescently labeled competitive inhibitor of the analyte of interest and a known concentration of the analyte of interest under conditions that will allow reactions between the analyte of interest in the fluid and the reactants on the beads in said set, wherein a reaction between a reactant and an analyte of interest on a bead causes a decrease in the value of a fluorescent signal (F_m) emitted from said bead;

(c) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(d) identifying the subset to which each bead belongs and therefore the concentration of reactant with which the bead was coupled as a function of the unique profile of classification parameter values; and

(e) assigning a bead subset value to each bead subset with correlates relatively with the concentration of analyte with which the bead subset was coupled; and

(f) determining an inter-bead subset slope from a plot of mean F_m for each bead subset versus bead subset value; and

(g) repeating steps (a) - (f) at least one time but with a known concentration of analyte of interest that differs from said concentration of analyte of interest employed in any other step (b); and

(h) plotting to produce a standard curve the inter-bead subset slopes at each known concentration of analyte of interest against the log of each known concentration of analyte of interest.

53. A method for flow cytometric analysis to detect a plurality of nucleic acid analytes of interest in a single sample comprising:

(a) obtaining a beadset comprising a plurality of subsets of beads wherein the beads in each subset;

(1) are sufficiently homogeneous with respect to each of at least three selected classification parameter (C_1, C_2, C_3 ... C_n) values and sufficiently different from beads in any other subset in at least one of said

classification parameter values so that the profile of classification parameter values within each subset detectable by flow cytometry is unique,

(2) are coupled to a nucleic acid that will specifically hybridize with a selected nucleic acid analyte of interest in a fluid to be tested,

(3) are reactive with a fluorescently labeled nucleic acid probe which competes with said nucleic acid analyte for hybridization with said nucleic acid coupled to the bead;

(b) mixing, to produce a reacted bead sample, the beadset with the fluid to be tested under conditions that will allow hybridization between nucleic acid analytes of interest in the fluid and the nucleic acids coupled to the beads in said beadset and thereby to allow the nucleic acid analytes in said fluid to inhibit hybridization between the fluorescently labeled nucleic acids with the nucleic acids coupled to said beads, resulting in a decrease in a fluorescent signal F_m emitted from a bead with which a nucleic acid analyte of interest in the fluid has reacted;

(c) analyzing the reacted sample by flow cytometry to determine the classification parameter value profile and an F_m value of each bead analyzed;

(d) identifying the subset to which each bead belongs and therefore the reactant on the bead as a function of the unique profile of classification parameter values; and

(e) detecting the presence or absence of a particular analyte of interest in said sample as a function of the identification in step (d) and an increase in the F_m values of

the beads in each of said subsets in the reacted fluid sample from the F_m values of the beads in each of said subsets not reacted with said fluid.

54. The method of claim 3 wherein said analytes are enzymes, said reactants are fluorescent molecules which upon reaction with the enzyme lose fluorescence, said change in F_m results from alteration of said substrates attached to said beads.

55. The method of claim 3 wherein said analytes are enzymes, said reactants are non-fluorescent molecules which upon reaction with the enzyme become fluorescent, and said change in F_m results from alteration of said substrates attached to said beads.

56. The method of claim 3 wherein said analytes are convertases which produce active enzymes from inactive precursors, said reactants are inactive precursors that are converted to active enzyme which in turn are reactants of fluorescently labeled substrates for said newly activated enzymes, and said change in F_m results from cleavage of said substrates from said beads.

57. The method of claim 5 wherein said analytes are enzymes, said reactants are molecules attached to a bead which, upon reaction with the enzyme, become ligates for a fluorescently labeled ligand, and wherein said change in F_m results from reaction of the new ligate with the fluorescently labeled ligand.

58. The method of claim 3 wherein said analyte is a cofactor which produces an active enzyme from an inactive apo-enzymes, said reactant is a fluorescently labeled substrate for said activated enzyme, and said change in F_m results from cleavage of said substrate from said active enzyme.

59. A method of processing a plurality of data signals generated by a flow cytometer in real-time, each said data signal being associated with a specific cytometric target and encoding a forward light scatter value, a side light scatter value, a red fluorescence value, an orange fluorescence value, and a green fluorescence value, comprising:

(a) receiving a data signal;

(b) extracting from said data signal (1) a forward light scatter value, (2) a side light scatter value, (3) a red fluorescence value, (4) an orange fluorescence value, and (5) a green fluorescence value;

(c) classifying said cytometric target into one of a plurality of classes, referred to as an identified class, said classification being a function of said extracted (1) forward light scatter value, (2) side light scatter value, (3) red fluorescence value, and (4) orange fluorescence value;

(d) incrementing a class-count value associated with said identified class, said class-count value encoding the number of cytometric targets classified as belonging to said identified class;

(e) accumulating a green-fluorescence-sum value associated with said identified class, said green fluorescence sum-value encoding an arithmetic sum of said extracted green fluorescence value for all cytometric targets classified as belonging to said identified class;

(f) repeating the operations described in paragraphs (a) through (e) for subsequent data signals;

(g) generating, for each of said plurality of classes, one or more outcome-description signals encoding textual information correlated with the class-count value and with the green-fluorescence-sum value for said class; and

(h) displaying said textual information.

60. The method of claim 59 wherein said specific cytometric target is an appropriately labeled bead.

61. The method of claim 60 wherein each one of said plurality of classes is associated with one bead subset, said bead subset formed in accordance with claim 2.

62. The method of claim 59 wherein the operation of paragraph © is further comprised of performing a reasonableness test on said cytometric target's identified class, said reasonableness test being a function of one or more of said identified class' (1) forward light scatter value, (2) side light scatter value, (3) red fluorescence value, and (4) orange fluorescence value.

63. A method of processing a plurality of data signals generated by a flow cytometer, each said data signal being associated with a specific flow cytometric target and encoding a plurality of classification parameter values and one or more measurement parameter values, comprising:

(a) receiving a data signal;

(b) extracting said plurality of classification parameter values and said one or more measurement parameter values from said data signal;

(c) classifying said cytometric target into one of a plurality of classes, referred to as an identified class, said classification being a function of said plurality of extracted classification parameter values;

(d) incrementing a class-count value associated with said identified class;

(e) accumulating each of said one or more extracted measurement values into one or more respective accumulation-values for said identified class;

(f) repeating the operations described in paragraphs (a) through (e) for subsequent data signals;

(g) generating, for each of said plurality of classes, one or more outcome-description signals encoding information correlated with the class-count and with the one or more accumulation-sum values for said class; and

(h) displaying said textual information.

64. The method of claim 63 wherein said processing in performed in real-time.

65. The method of claim 64 wherein each one of said plurality of classes is associated with one bead subset, said bead subset formed in accordance with claim 2.

66. The method of claim 63 wherein said specific cytometric target is an appropriately labeled bead.

67. The method of claim 63 wherein said data signal encodes a plurality of classification parameter values selected from the group consisting of forward light scatter, side light scatter, red fluorescence, and orange fluorescence.

68. The method of claim 63 wherein said data signal encodes one or more measurement parameter values selected from the group consisting of orange fluorescence and green fluorescence.

69. The method of claim 63 wherein said one or more outcome-description signals encodes textual information.

70. The method of claim 63 wherein each of said one or more outcome-description signals is determined by either an OVER-UNDER test or a SHIFT test.

71. An machine readable assay database, stored in a storage device, for the processing of flow-cytometric measurement data comprising:

(a) an assay definition table, said assay definition table encoding (1) one or more measurement subset token identifiers, (2) for each subset token identifier, one or more baseline measurement parameter values, and (3) for each subset token identifier, an interpretation test-type token;

(b) a discriminant function table, said discriminant function table encoding a classification decision tree based on one or more classification measurement parameters, said one or more classification measurement parameters encoded in said flow-cytometric measurement data;

(c) an interpretation table, said interpretation table encoding textual assay outcome-description information; and

(d) a results table, said results table capable of encoding statistical accumulation of real-time flow-cytometric measurement data.

72. A method of processing a plurality of data signals, in real-time, generated by a diagnostic device, each of said plurality of data signals being associated with a specific diagnostic target and encoding a plurality of classification parameter values and one or more measurement parameter values, comprising:

(a) receiving a data signal;

(b) extracting said plurality of classification parameter values and said one or more measurement parameter values from said data signal;

(c) classifying said diagnostic target into one of a plurality of classes, referred to as an identified class, said classification being a function of said plurality of extracted classification parameter values;

(d) incrementing a class-count value associated with said identified class;

(e) accumulating each of said one or more extracted measurement values into one or more respective accumulation-values for said identified class;

(f) repeating the operations described in paragraphs (a) through (e) for subsequent data signals;

(g) generating, for each of said plurality of classes, one or more outcome-description signals encoding information correlated with the class-count and with the one or more accumulation-sum values for said class; and

(h) displaying said outcome-description signals.

73. The method of claim 72 wherein said diagnostic device is selected from the group consisting of a flow cytometer and a cell sorter.

74. A program storage device that is readable by a computer, said program storage device having encoded therein a program of instructions that includes instructions for executing the method steps of a specified one of claims 59, 63, 71, and 72.

75. A method for flow cytometric analysis to detect genetic mutations in a DNA comprising:

(a) obtaining beads coupled to an oligonucleotide molecule designed to hybridize with a selected PCR product of interest;

(b) mixing the beads with said PCR product under conditions that will allow hybridization between said PCR product and the oligonucleotide coupled to the beads and thereby to allow the PCR product to inhibit hybridization between a fluorescently labeled nucleic acid probe that is completely complementary to said oligonucleotide coupled to said beads;

(c) adding said fluorescent probe to the mixture;

(d) analyzing the reacted sample by flow cytometry to determine the fluorescence of each bead analyzed; and

(e) detecting the genetic mutation or absence thereof as a result of the degree of fluorescence on the beads.

76. A method to detect a genetic mutations in a DNA comprising:

(a) obtaining beads coupled to an oligonucleotide molecule, said oligonucleotide molecule designed to hybridize with a selected PCR product of interest;

(b) mixing said beads with said PCR product, under conditions that will allow hybridization between said PCR product and the oligonucleotide coupled to the beads, to form a reacted mixture;

(c) adding a fluorescent probe to said reacted mixture;

(d) determining the fluorescence of the beads by flow cytometry; and

(e) detecting the genetic mutation, or absence thereof, as a result of the degree of the determined fluorescence on the beads.

77. A method of detecting a genetic mutation in a DNA comprising the steps of:

(a) selecting an oligonucleotide probe for said genetic mutation;

(b) preparing a fluorescent DNA probe complementary to the oligonucleotide probe coupling said selected probe to each one of a plurality of beads to form a bead aliquot;

(c) selecting PCR primers to amplify a region of said DNA corresponding to said selected probe;

(d) amplifying said genetic mutation by PCR to form PCR products;

(e) mixing said bead aliquot, said PCR products and said fluorescent probe to form a mixture;

(f) incubating said mixture to promote under competitive hybridization conditions;

(g) measuring the fluorescence said beads by flow cytometry; and

(h) detecting said genetic mutation, or absence thereof, as a function of the measured fluorescence of said beads.

78. The method of claim 77 wherein said genetic mutation is selected from the group consisting of mutations in MEN2a, MEN2b, MEN1, ret proto-oncogene, LDL receptor, NF1, NF type 2, BRCA1, BRCA2, BRCA3, APC, adenosine deaminase, XPAC, ERCC6 excision repair gene, fmr1, Duchenne's muscular dystrophy gene, myotonic dystrophy

protein kinase, androgen receptor, Huntington's, HPRT, apolipoprotein E, HEXA, steroid 2-hydroxylase, angiotensin, hNMLH1, 2 mismatch repair, APC, Rb, p53, bcr/abl, bcl-2 gene, chromosomes 11 to 14 and chromosomes 15 to 17 gene transpositions, and genes encoding ion transporters.

79. The method of claim 77 wherein said oligonucleotide probe has a length of between 5 and 500 nucleotides.

80. The method of claim 77 wherein said PCR primers are designed to amply a region of said DNA corresponding to said oligonucleotide probe.

81. The method of claim 77 wherein said fluorescent probe is selected from the group consisting of DNA sequences complementary to wild-type or mutant sequences coupled to the beads.

82. A kit for detection of a genetic mutation in a DNA comprising:

(a) a first container comprising beads coupled to an oligonucleotide designed to hybridize with a selected PCR product of interest;

(b) a second container comprising a PCR primer designed to amplify a section of DNA complementary to said oligonucleotide; and

(c) a third container comprising a fluorescent labeled DNA probe capable of selectively hybridizing said oligonucleotide.

83. The kit of claim 82, wherein said genetic mutation is selected from the group consisting of mutations in MEN2a, MEN2b, MEN1, ret proto-oncogene, LDL receptor, NF1, NF type 2, BRCA1, BRCA2, BRCA3, APC, adenosine deaminase, XPAC, ERCC6 excision

repair gene, fmr1, Duchenne's muscular dystrophy gene, myotonic dystrophy protein kinase, androgen receptor, Huntington's, HPRT, apolipoprotein E, HEXA, steroid 2-hydroxylase, angiotensin, hNMLH1, 2 mismatch repair, APC, Rb, p53, bcr/abl, bcl-2 gene, chromosomes 11 to 14 and chromosomes 15 to 17 gene transpositions, and genes encoding ion transporters.

84. The kit of claim 82 wherein said fluorescent labeled DNA probe has a length of between 5 and 500 nucleotides.

85. The method claim 3 where the analytes of interest are DNA segments, the reactant on the bead are DNA segment capable of specifically hybridizing to said analytes, and the fluorescent label is a fluorescent DNA segment also capable of specifically hybridizing with said reactant to compete with the hybridization of said reactant to said label.

Fig. 1

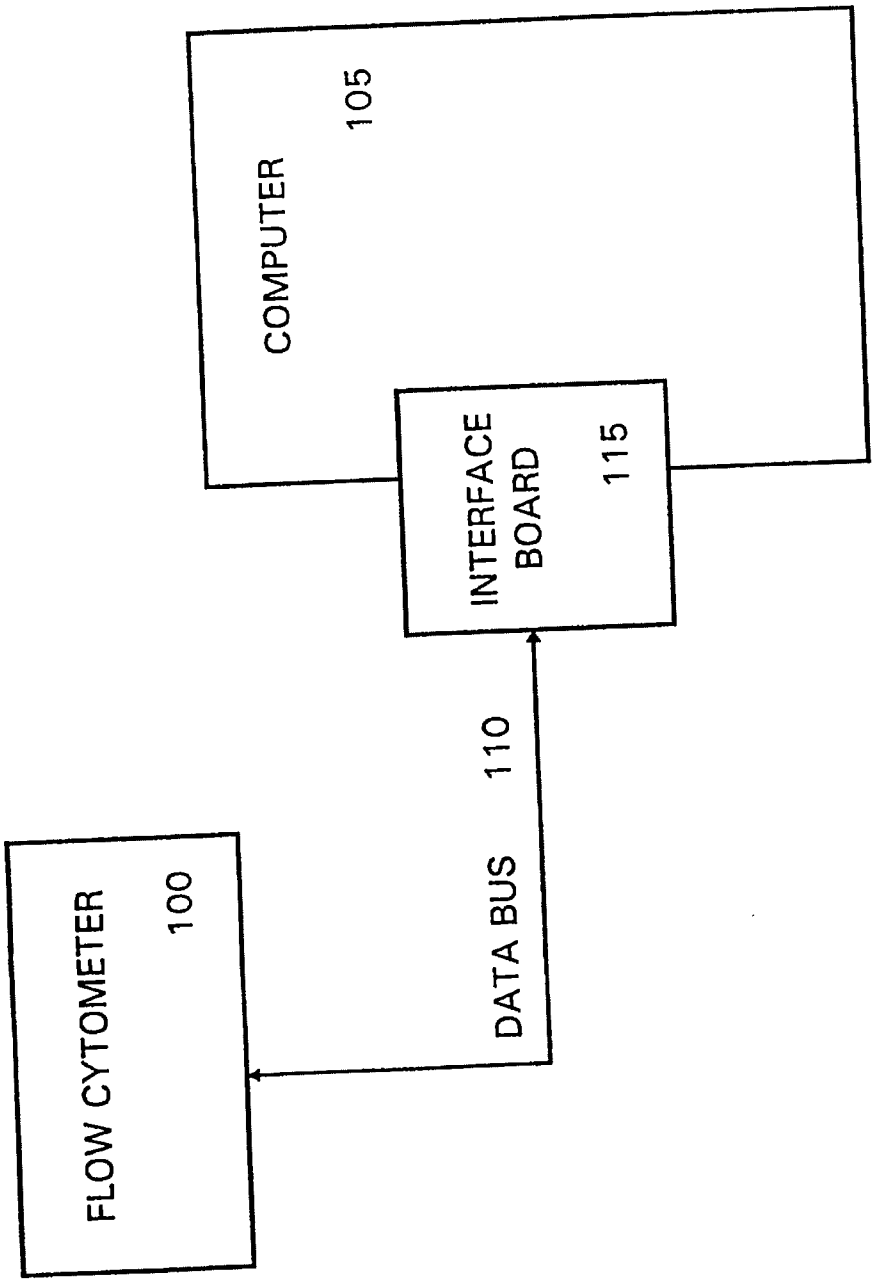

COMPUTER
105
INTERFACE BOARD
115
DATA BUS
110
FLOW CYTOMETER
100

Fig. 2

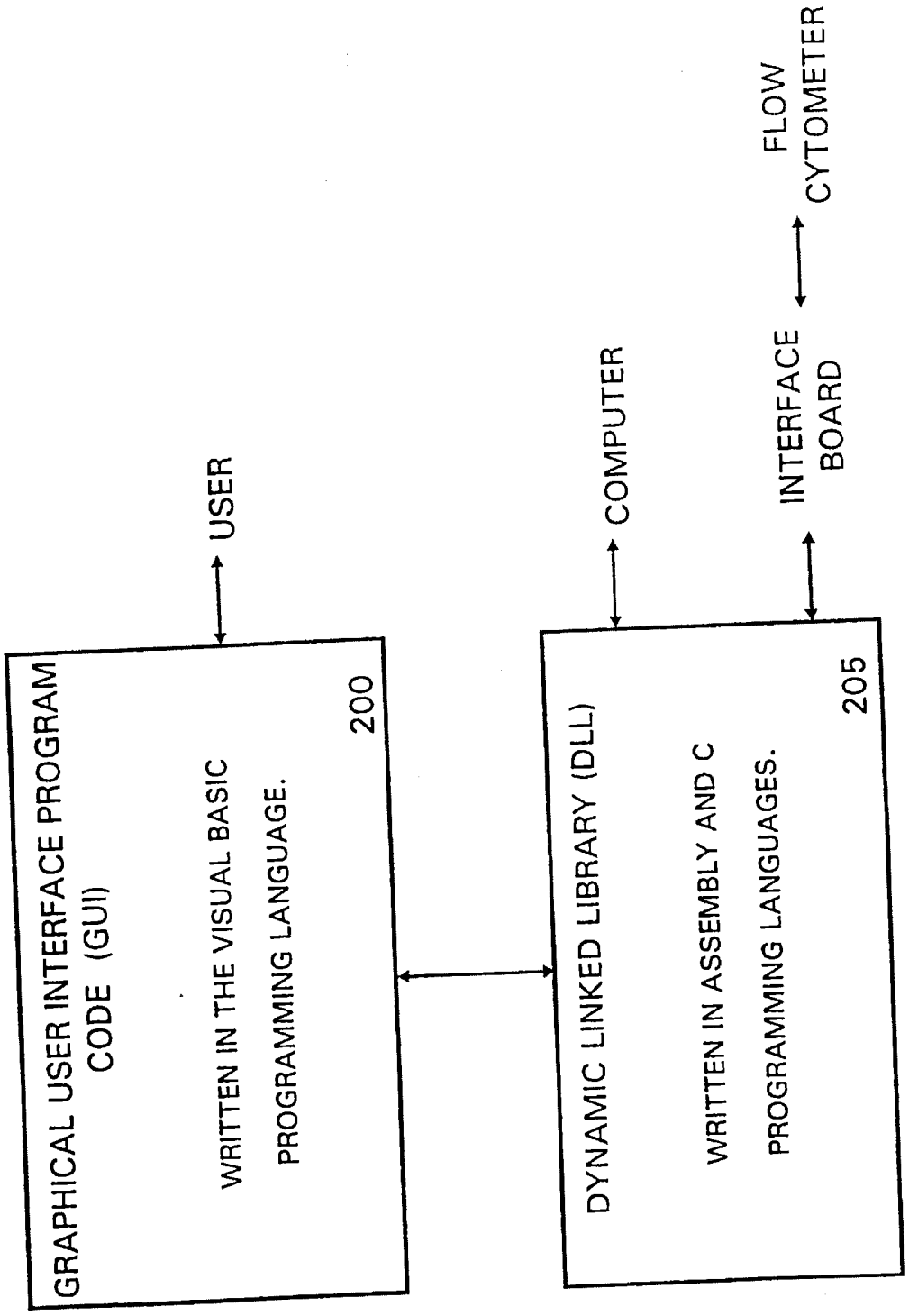

USER
GRAPHICAL USER INTERFACE PROGRAM CODE (GUI)
WRITTEN IN THE VISUAL BASIC
PROGRAMMING LANGUAGE.
200
COMPUTER
INTERFACE BOARD
FLOW CYTOMETER
DYNAMIC LINKED LIBRARY (DLL)
WRITTEN IN ASSEMBLY AND C
PROGRAMMING LANGUAGES.
205

Fig. 3

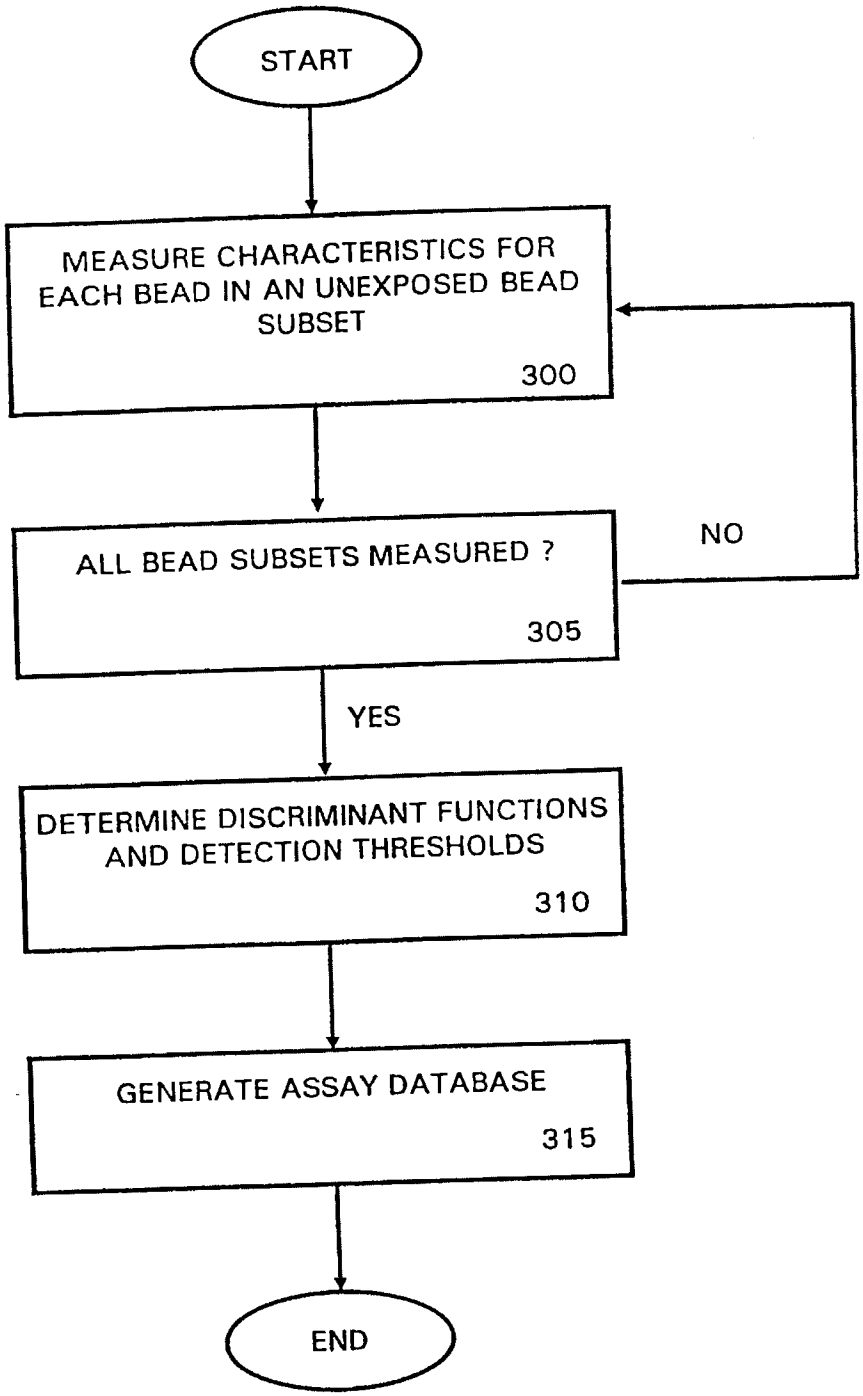

START
MEASURE CHARACTERISTICS FOR EACH BEAD IN AN UNEXPOSED BEAD SUBSET
300
ALL BEAD SUBSETS MEASURED ?
305
NO
YES
DETERMINE DISCRIMINANT FUNCTIONS AND DETECTION THRESHOLDS
310
GENERATE ASSAY DATABASE
315
END

Fig. 4

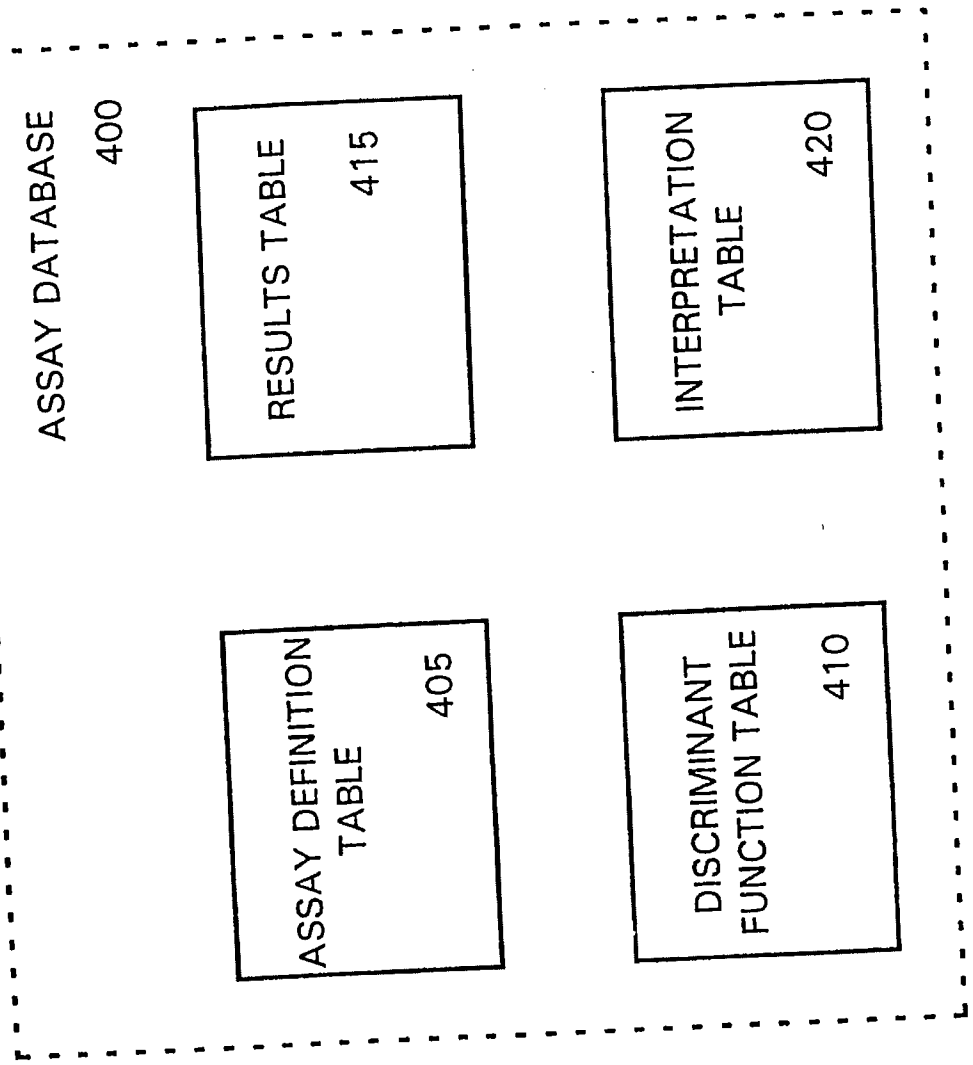

ASSAY DATABASE
400
RESULTS TABLE
415
INTERPRETATION TABLE
420
ASSAY DEFINITION TABLE
405
DISCRIMINANT FUNCTION TABLE
410

Fig. 5

SAMPLE BASELINE DATA ACQUISITION

ANALYTE	C_1: SIDE LIGHT SCATTER	C_2: FORWARD LIGHT SCATTER	C_3: ORANGE FLUORESCENCE	C_4: RED FLUORESCENCE	F_1: GREEN FLUORESCENCE
1	$\mu = 560$ $\sigma = 5.1$	*	$\mu = 0$	$\mu = 0$	$\mu = 170$ $\sigma = 1.3$
2	$\mu = 579$ $\sigma = 5.1$	*	$\mu = 48$ $\sigma = 0.69$	$\mu = 368$ $\sigma = 1.92$	$\mu = 170$ $\sigma = 1.3$
3	$\mu = 519$ $\sigma = 4.56$	*	$\mu = 98$ $\sigma = 0.99$	$\mu = 550$ $\sigma = 2.35$	$\mu = 170$ $\sigma = 1.3$
4	$\mu = 519$ $\sigma = 4.56$	*	$\mu = 0$	$\mu = 527$ $\sigma = 2.30$	$\mu = 170$ $\sigma = 1.3$

* Not used in illustrative example

Fig. 6

SAMPLE ASSAY DEFINITION TABLE

ASSAY NAME	SUBSET TOKEN	SUBSET NAME	F_1's BASE VALUE	F_1's STANDARD DEVIATION	TEST TYPE TOKEN
test_assay	46	KRAS CODON 46 WILDTYPE	170	1.3	0
test_assay	21	KRAS CODON 21 MUTANT	170	1.3	0
test_assay	50	KRAS CODON 50 MUTANT	170	1.3	0
test_assay	5	KRAS CODON 5 MUTANT	170	1.3	0

Fig. 7

SAMPLE DISCRIMINANT FUNCTION TABLE

ASSAY NAME	ROW ID	PARAMETER	LOW VALUE	HIGH VALUE	TRUE ROW ID	FALSE ROW ID	TRUE TOKEN	FALSE TOKEN
test_assay	0	C_1	500	620	1	0	0	0
test_assay	1	C_4	0	514	2	3	0	0
test_assay	2	C_4	0	200	0	0	50	5
test_assay	3	C_3	0	98	0	0	46	21

Fig. 8

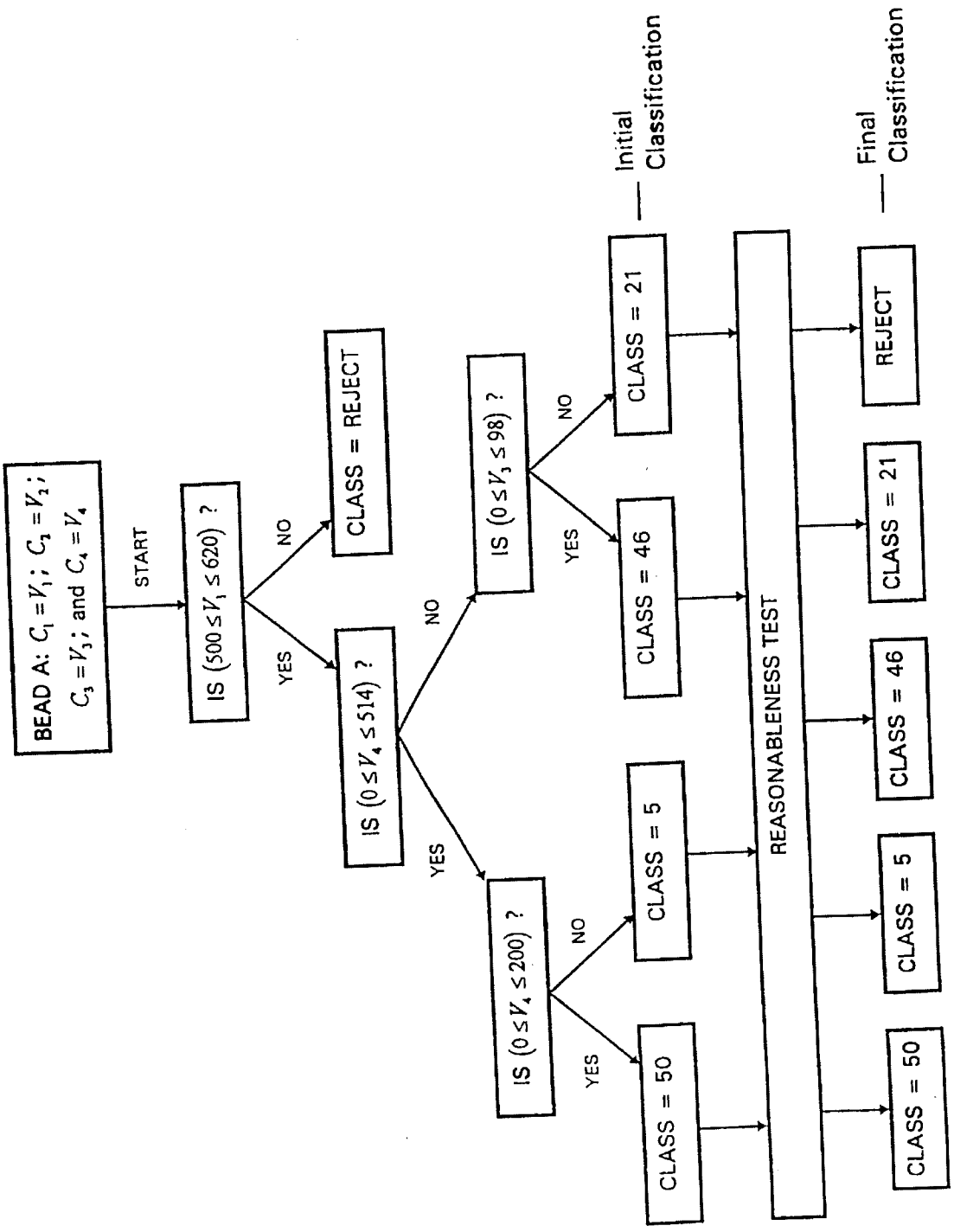

BEAD A: $C_1 = V_1$; $C_2 = V_2$; $C_3 = V_3$; and $C_4 = V_4$
START
IS $(500 \leq V_1 \leq 620)$?
NO
CLASS = REJECT
YES
IS $(0 \leq V_4 \leq 514)$?
NO
IS $(0 \leq V_3 \leq 98)$?
NO
CLASS = 21
YES
CLASS = 46
YES
IS $(0 \leq V_4 \leq 200)$?
NO
CLASS = 5
YES
CLASS = 50
Initial Classification
REASONABLENESS TEST
Final Classification
REJECT
CLASS = 21
CLASS = 46
CLASS = 5
CLASS = 50

Fig. 9

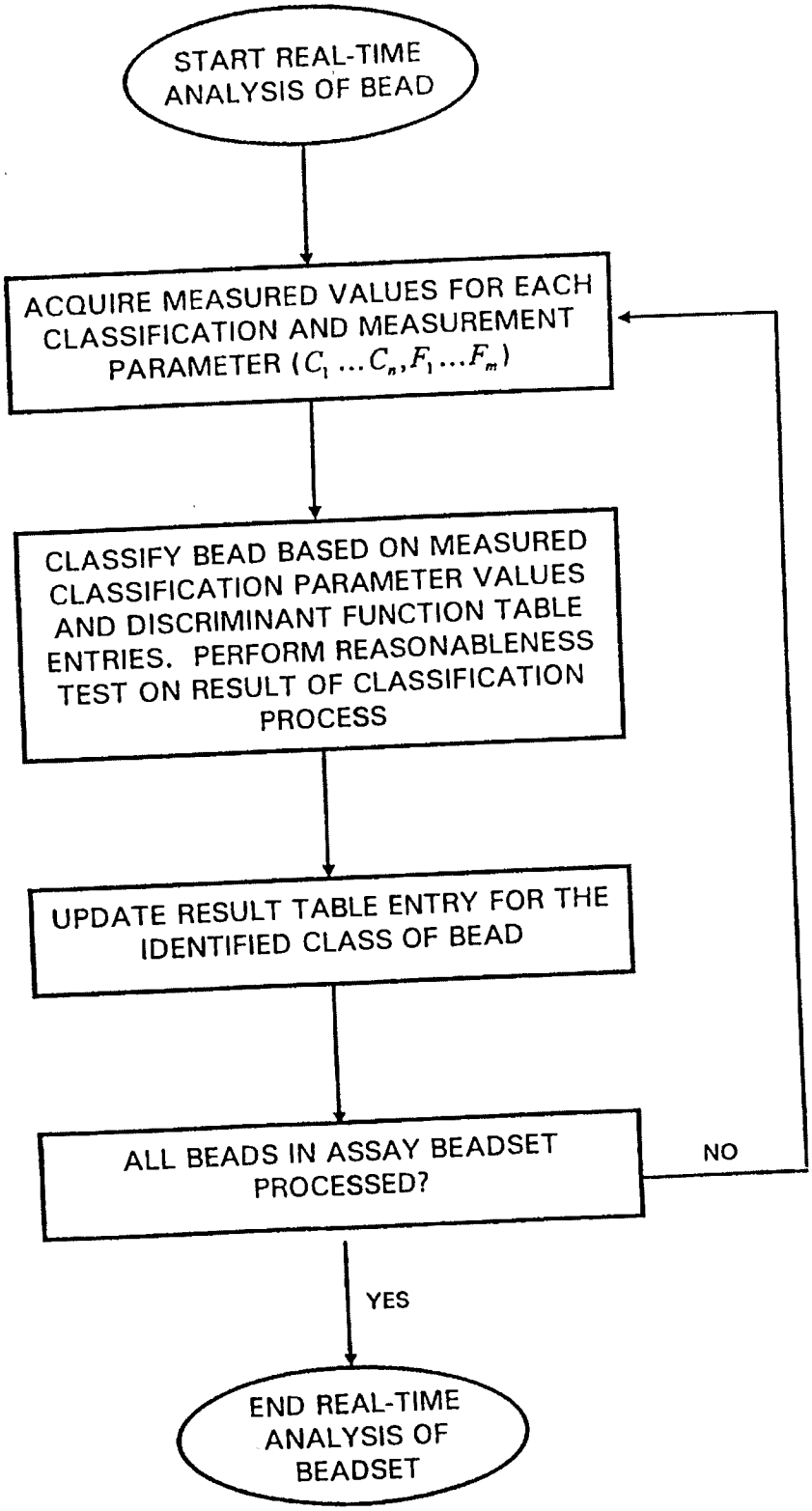

START REAL-TIME ANALYSIS OF BEAD
ACQUIRE MEASURED VALUES FOR EACH CLASSIFICATION AND MEASUREMENT PARAMETER ($C_1 \ldots C_n, F_1 \ldots F_m$)
CLASSIFY BEAD BASED ON MEASURED CLASSIFICATION PARAMETER VALUES AND DISCRIMINANT FUNCTION TABLE ENTRIES. PERFORM REASONABLENESS TEST ON RESULT OF CLASSIFICATION PROCESS
UPDATE RESULT TABLE ENTRY FOR THE IDENTIFIED CLASS OF BEAD
ALL BEADS IN ASSAY BEADSET PROCESSED?
NO
YES
END REAL-TIME ANALYSIS OF BEADSET

Fig. 10

SAMPLE RESULTS TABLE

ASSAY NAME	SUBSET TOKEN	COUNT	SUM	OVER COUNT	UNDER COUNT
test_assay	46	1,000	71,111	*	*
test_assay	21	1,000	90,000	*	*
test_assay	50	1,000	1,700,000	*	*
test_assay	5	1,000	70,000	*	*

* Not used in illustrative example

449

SAMPLE INTERPRETATION TABLE

Fig. 11

450

ASSAY NAME	SUBSET TOKEN	OUTCOME ID	TEST-TYPE TOKEN	LOW VALUE	HIGH VALUE	INTERPRETATION
test_assay	5	1	1	10	667	Identical complementary strand
test_assay	5	2	1	668	970	Similar oligo
test_assay	5	3	1	971	2,000	Not found in sample
test_assay	21	1	1	10	667	Identical complementary strand
test_assay	21	2	1	668	970	Similar oligo
test_assay	21	3	1	971	2,000	Not found in sample
test_assay	50	1	1	10	667	Identical complementary strand
test_assay	50	2	1	668	970	Similar oligo
test_assay	50	3	1	971	2,000	Not found in sample
test_assay	46	1	1	10	667	Identical complementary strand
test_assay	46	2	1	558	970	Similar oligo
test_assay	46	3	1	971	2,000	Not found in sample

Fig. 12

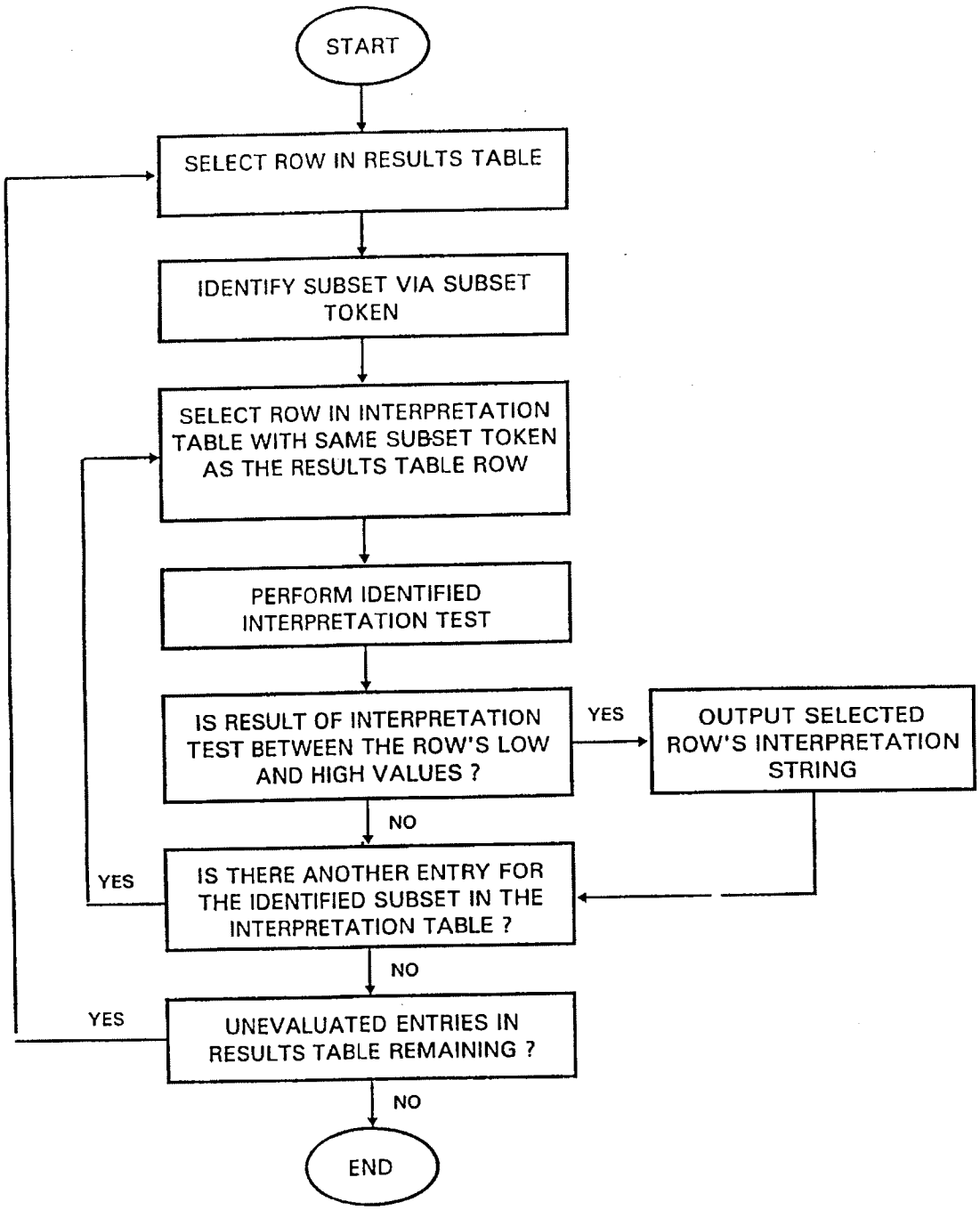

START
SELECT ROW IN RESULTS TABLE
IDENTIFY SUBSET VIA SUBSET TOKEN
SELECT ROW IN INTERPRETATION TABLE WITH SAME SUBSET TOKEN AS THE RESULTS TABLE ROW
PERFORM IDENTIFIED INTERPRETATION TEST
IS RESULT OF INTERPRETATION TEST BETWEEN THE ROW'S LOW AND HIGH VALUES ?
YES
OUTPUT SELECTED ROW'S INTERPRETATION STRING
NO
IS THERE ANOTHER ENTRY FOR THE IDENTIFIED SUBSET IN THE INTERPRETATION TABLE ?
YES
NO
UNEVALUATED ENTRIES IN RESULTS TABLE REMAINING ?
YES
NO
END

Fig. 13a

Assay Definition Table

Antigen	Microsphere identified via dye color	Antigen-Microsphere Complex	Fluorescent Antibody (F_m)	Subset Token
LH	1 µm Crimson Beads	LH—Crimson Beads	anti-LH	18
TSH	1 µm Dark Red Beads	TSH—Dark Red Beads	anti-TSH	45
IgA	3 µm Clear Beads	IgA — Clear Beads	anti-Ig	50

Fig. 13b

Baseline Data Table

Bead Subset dye color (token)	C_1 mean side light scatter	C_2 mean orange fluorescence	C_3 mean red fluorescence	F_m mean green fluorescence
Crimson (18)	38.46	1.71	30.69	9.15
Dark Red (45)	40.37	0.00	88.62	5.44
Clear (50)	167	0.00	0.00	7.83

Fig. 13c

Discriminant Function Table

Row ID	Parameter	Low Value	High Value	True Node	False Node	True Token	False Token
0	C_3	0	5	3	1	0	0
1	C_1	400	420	2	0	0	0
2	C_3	0	454	0	0	18	45
3	C_1	560	580	0	0	50	0

Fig. 13d

Interpretation Table

— $F_m \pm 1.5$ log(baeline F_m value) —

Subset Token	Outcome ID	Test-Type Token	Low Value	High Value	Interpretation
18	1	1	289.35	-	Anti-LH found.
18	2	1	-	289.35	Anti-LH not found.
45	1	1	172.03	-	Anti-TSH found.
45	2	1	-	172.03	Anti-TSH not found.
50	1	1	247.61	-	Anti-IgA found.
50	2	1	-	247.61	Anti-IgA not found.

455

Fig. 13e

Results Table

Sample	Antibody Present	Measured F_m Subset 18	Measured F_m Subset 45	Measured F_m Subset 50
1	Anti—IgA	70	10	6049
2	anti—LH	132	442	180
3	anti—TSH	2124	108	182
4	anti—IgA + anti—TSH	2152	115	5917

456

Fig. 14

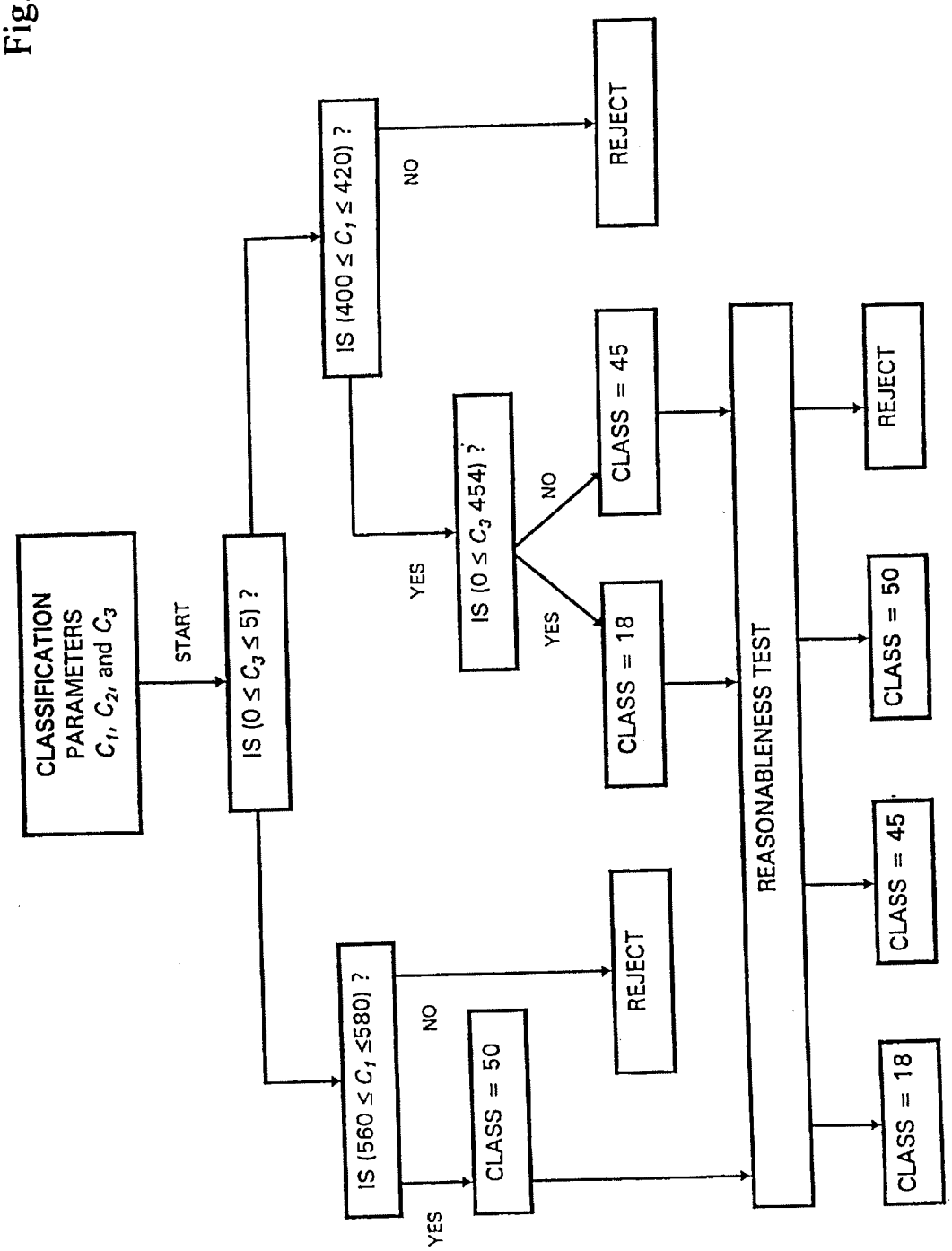

CLASSIFICATION PARAMETERS C_1, C_2, and C_3
START
IS $(0 \leq C_3 \leq 5)$?
IS $(400 \leq C_1 \leq 420)$?
NO
REJECT
YES
IS $(0 \leq C_3 \leq 454)$?
NO
CLASS = 45
YES
CLASS = 18
IS $(560 \leq C_1 \leq 580)$?
NO
REJECT
YES
CLASS = 50
REASONABLENESS TEST
REJECT
CLASS = 50
CLASS = 45
CLASS = 18

Fig. 15a

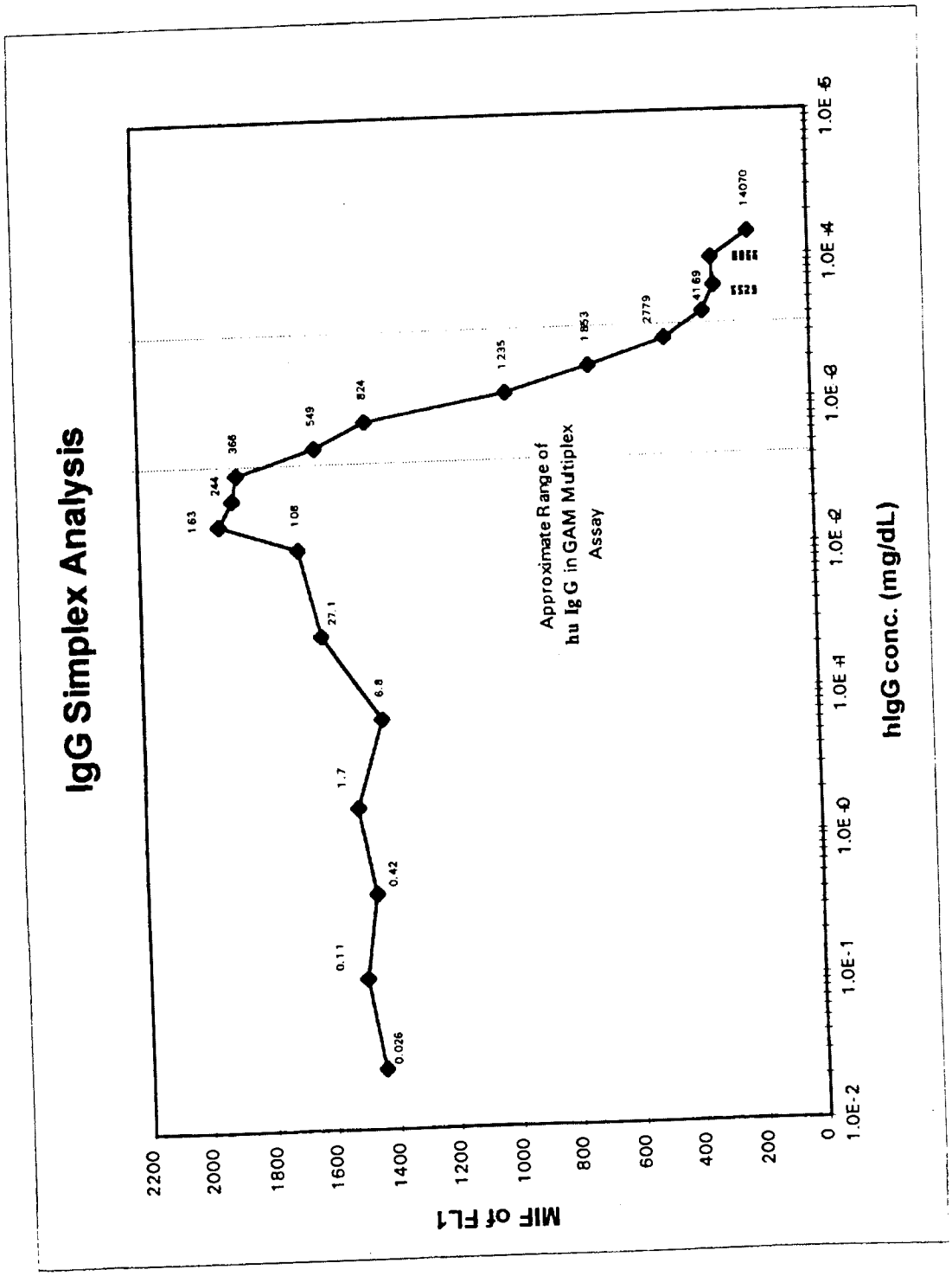
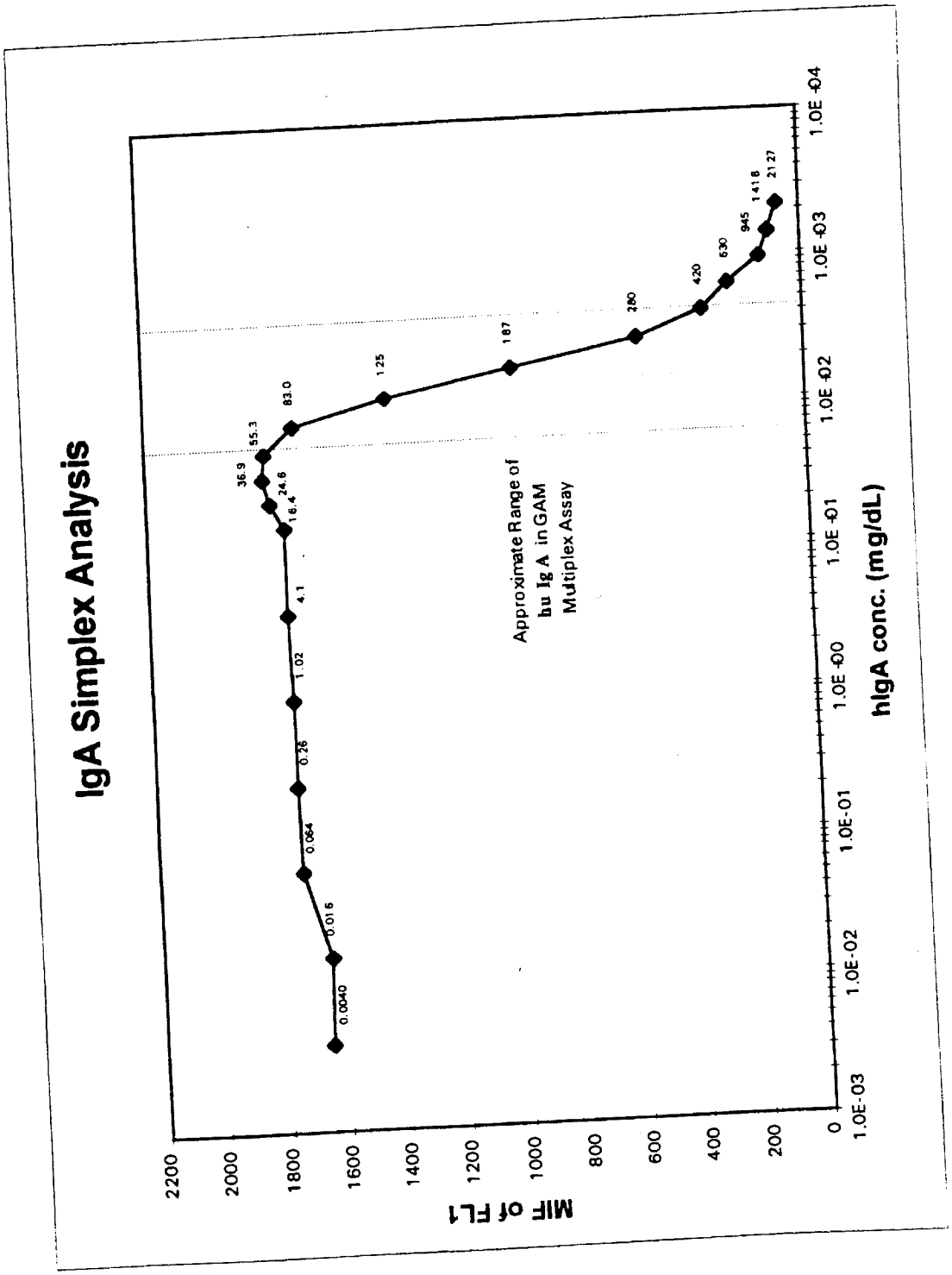

IgG Simplex Analysis
MIF of FL1
hIgG conc. (mg/dL)
Approximate Range of hu Ig G in GAM Multiplex Assay
0.026
0.11
0.42
1.7
6.8
27.1
108
163
244
366
549
824
1235
1853
2779
4169
14070
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1.0E-2
1.0E-1
1.0E-0
1.0E+1
1.0E+2
1.0E+3
1.0E+4
1.0E+5

Fig. 15b


IgA Simplex Analysis
MIF of FL1
hIgA conc. (mg/dL)
Approximate Range of
hu Ig A in GAM
Multiplex Assay
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1.0E-03
1.0E-02
1.0E-01
1.0E-00
1.0E-01
1.0E-02
1.0E-03
1.0E-04
0.0040
0.016
0.064
0.26
1.02
4.1
16.4
24.6
36.9
55.3
83.0
125
187
280
420
630
945
1418
2127

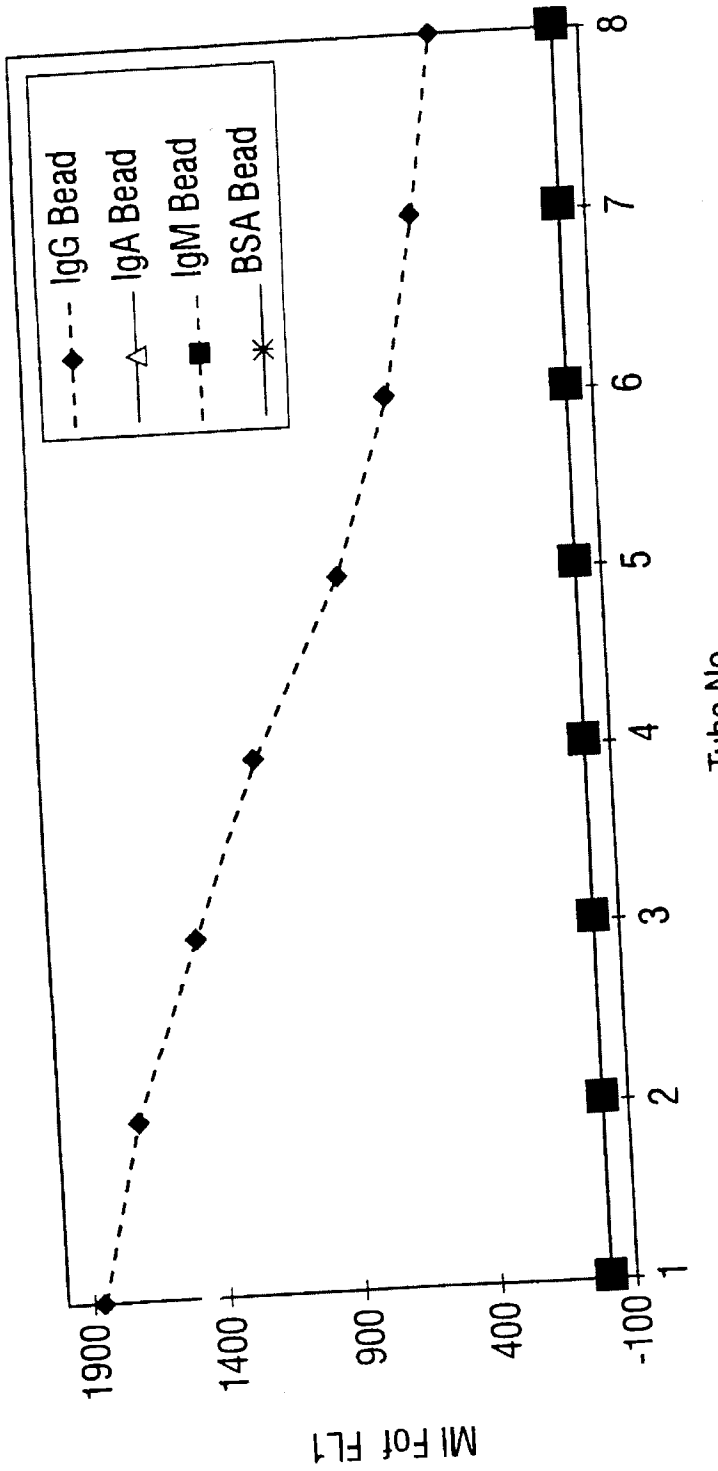

IgG Bead
IgA Bead
IgM Bead
BSA Bead
1900
1400
900
400
-100
MI For FL1
1
2
3
4
5
6
7
8
Tube No.

FIG. 16A

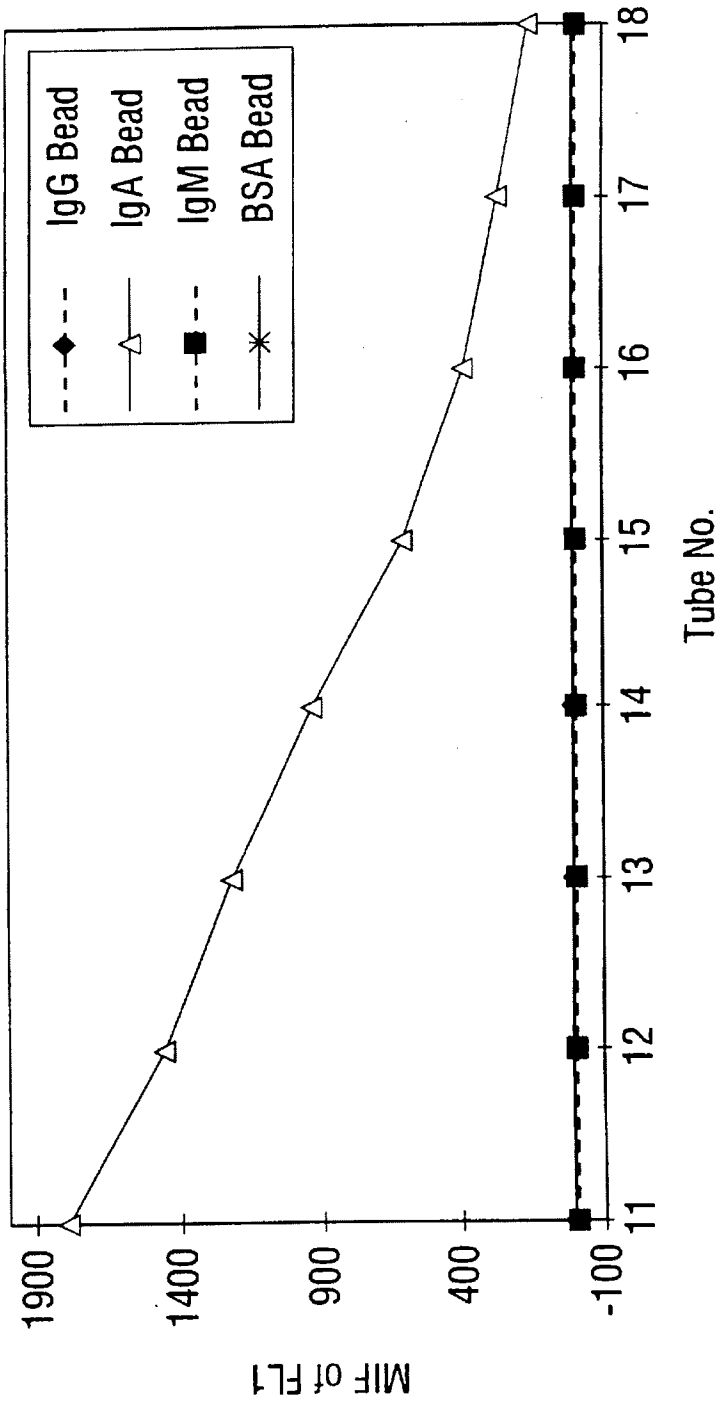

IgG Bead
IgA Bead
IgM Bead
BSA Bead
MIF of FL1
1900
1400
900
400
-100
11
12
13
14
15
16
17
18
Tube No.

FIG. 16B

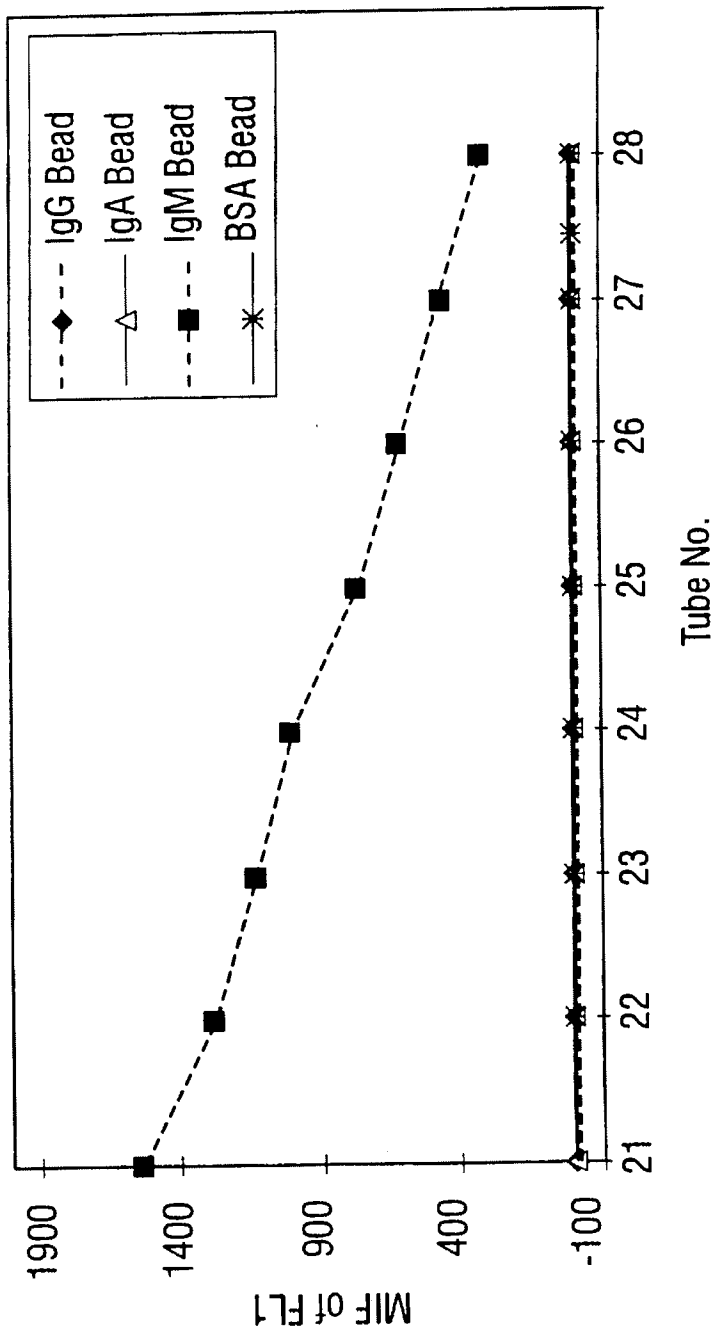

IgG Bead
IgA Bead
IgM Bead
BSA Bead
MIF of FL1
1900
1400
900
400
-100
21
22
23
24
25
26
27
28
Tube No.

FIG. 16C

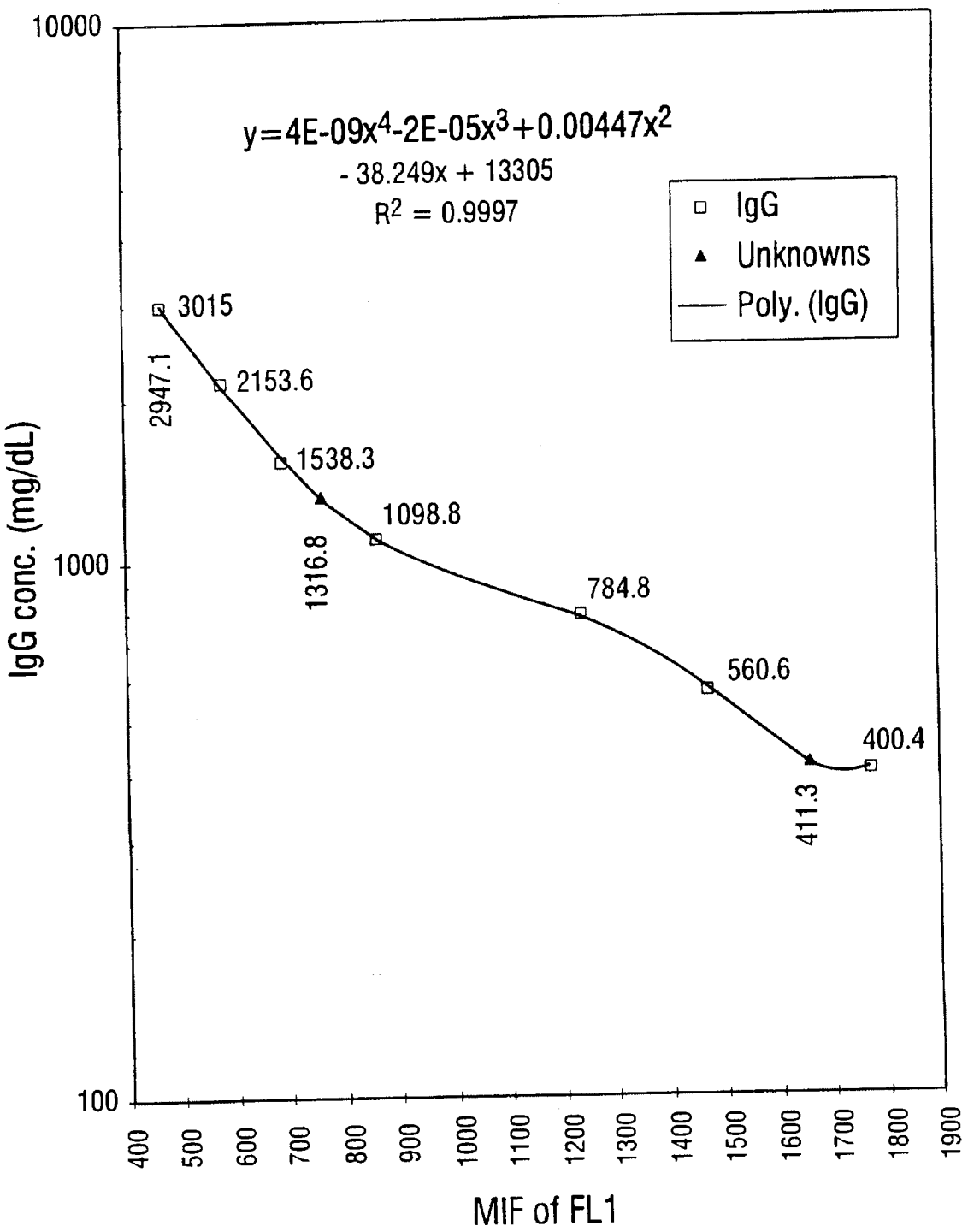

y=4E-09x4-2E-05x3+0.00447x2
- 38.249x + 13305
R2 = 0.9997
IgG
Unknowns
Poly. (IgG)
3015
2947.1
2153.6
1538.3
1316.8
1098.8
784.8
560.6
411.3
400.4
IgG conc. (mg/dL)
10000
1000
100
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
MIF of FL1

FIG. 17

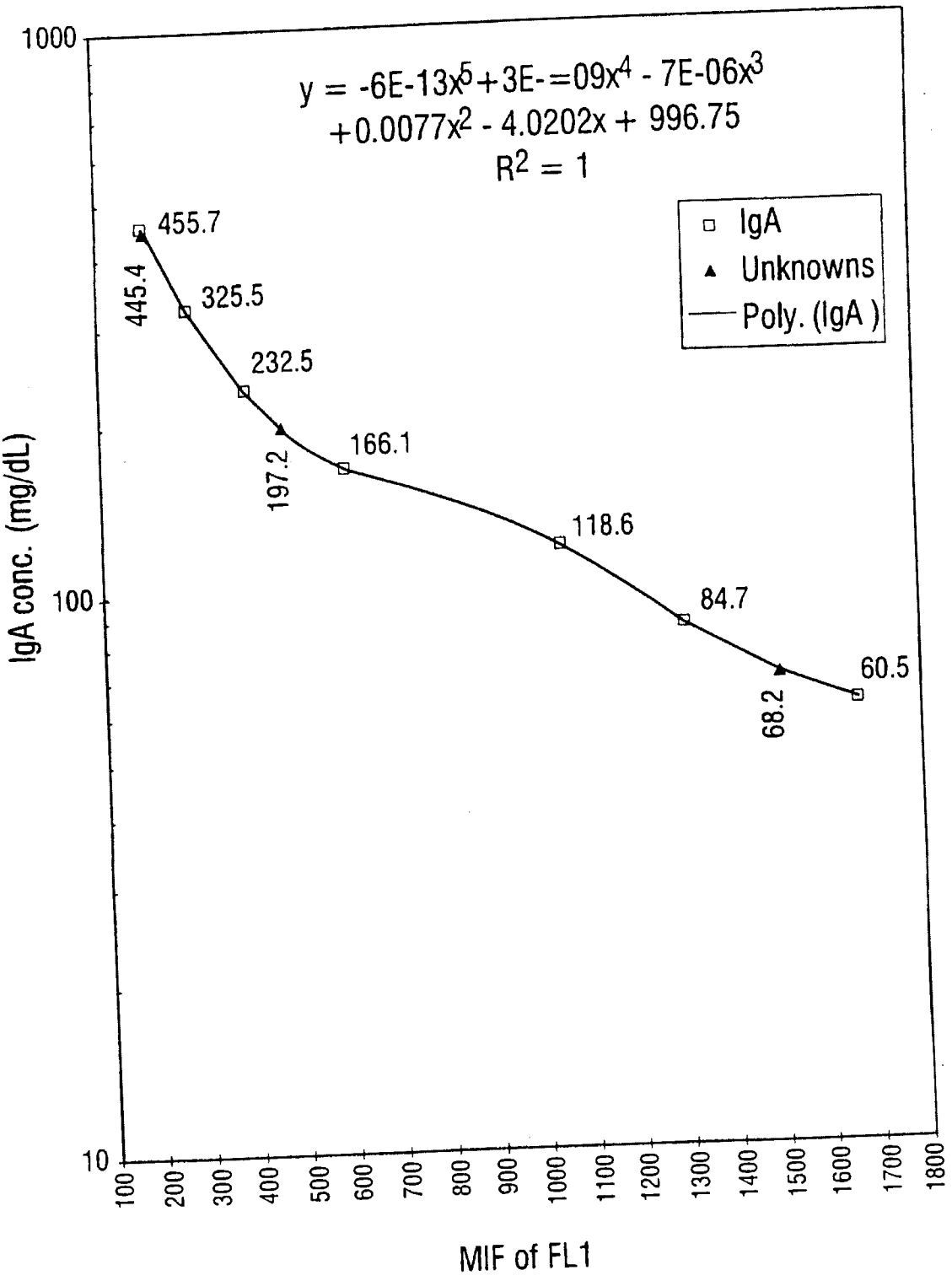

y = -6E-13x5+3E-=09x4 - 7E-06x3
+0.0077x2 - 4.0202x + 996.75
R2 = 1
IgA
Unknowns
Poly. (IgA)
455.7
445.4
325.5
232.5
197.2
166.1
118.6
84.7
68.2
60.5
IgA conc. (mg/dL)
1000
100
10
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
MIF of FL1

FIG. 18

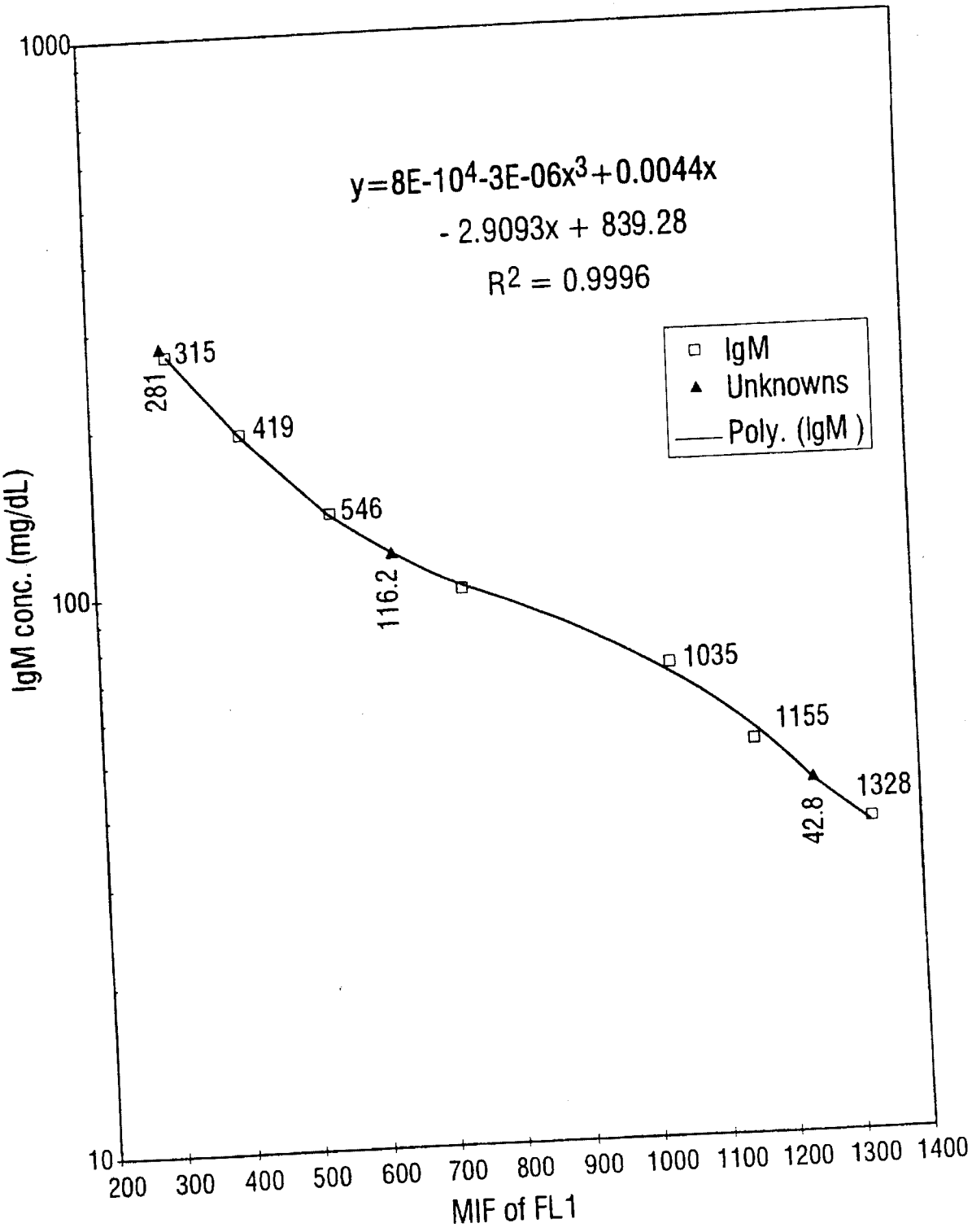

y=8E-10^4-3E-06x^3+0.0044x
- 2.9093x + 839.28
R2 = 0.9996
IgM
Unknowns
Poly. (IgM)
IgM conc. (mg/dL)
1000
100
10
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
MIF of FL1
315
281
419
546
116.2
1035
1155
1328
42.8

FIG. 19

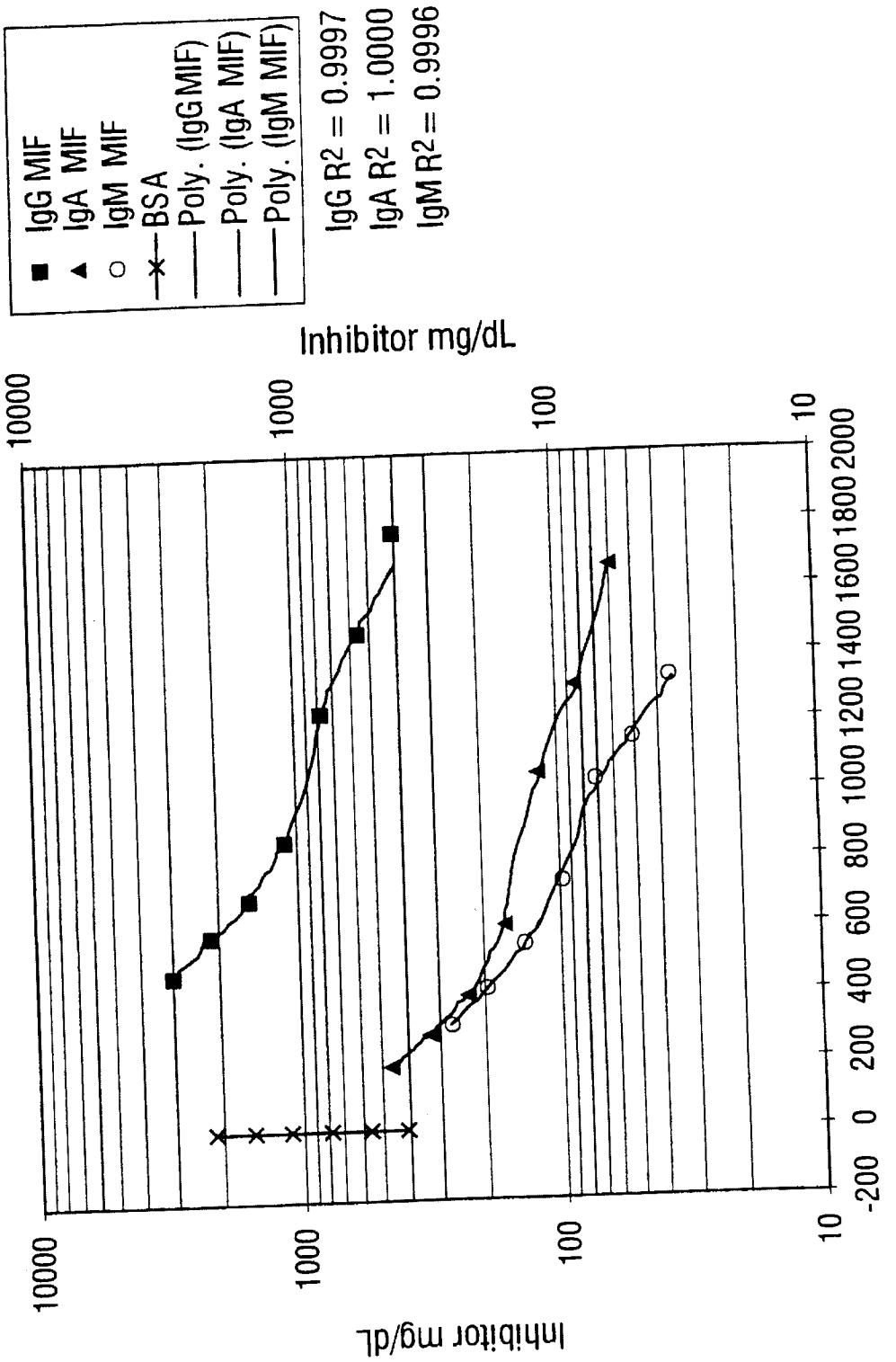

IgG MIF
IgA MIF
IgM MIF
BSA
Poly. (IgG MIF)
Poly. (IgA MIF)
Poly. (IgM MIF)
IgG R2 = 0.9997
IgA R2 = 1.0000
IgM R2 = 0.9996
Inhibitor mg/dL
10000
1000
100
10
-200 0 200 400 600 800 1000 1200 1400 1600 1800 2000
MIF of FL1
Inhibitor mg/dL

FIG. 20

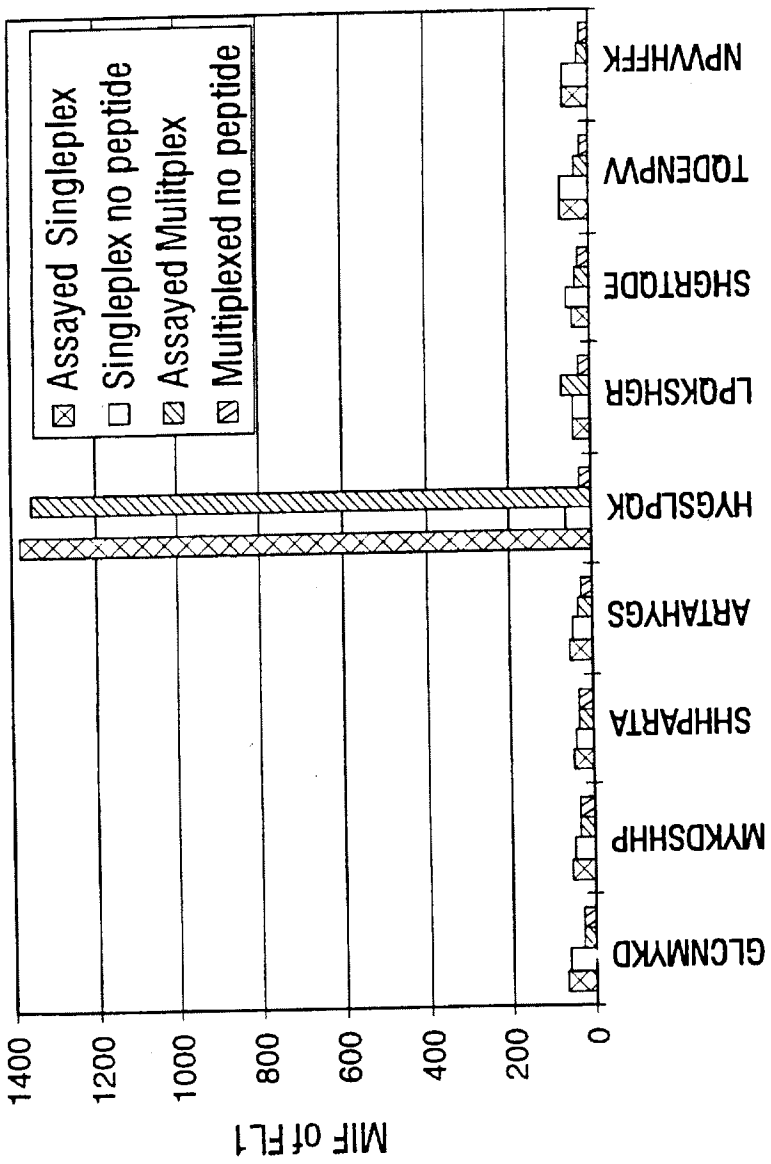

Assayed Singleplex
Singleplex no peptide
Assayed Mulitplex
Multiplexed no peptide
MIF of FL1
1400
1200
1000
800
600
400
200
0
GLCNMYKD
MYKDSHHP
SHHPARTA
ARTAHYGS
HYGSLPQK
LPQKSHGR
SHGRTQDE
TQDENPVV
NPVVHFFK
Peptides # 1-9

FIG. 21

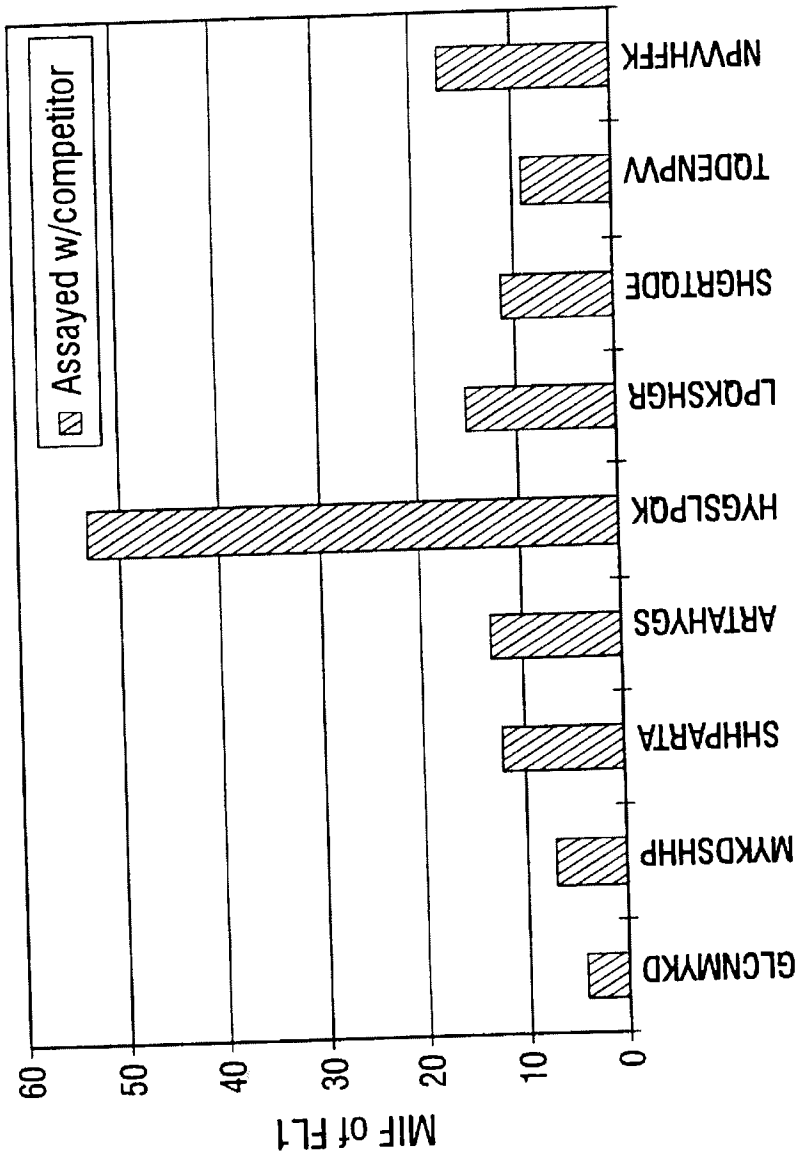

Assayed w/competitor
NPVVHFFK
TQDENPVV
SHGRTQDE
LPQKSHGR
HYGSLPQK
ARTAHYGS
SHHPARTA
MYKDSHHP
GLCNMYKD
0
10
20
30
40
50
60
MIF of FL1
Peptide

FIG. 22

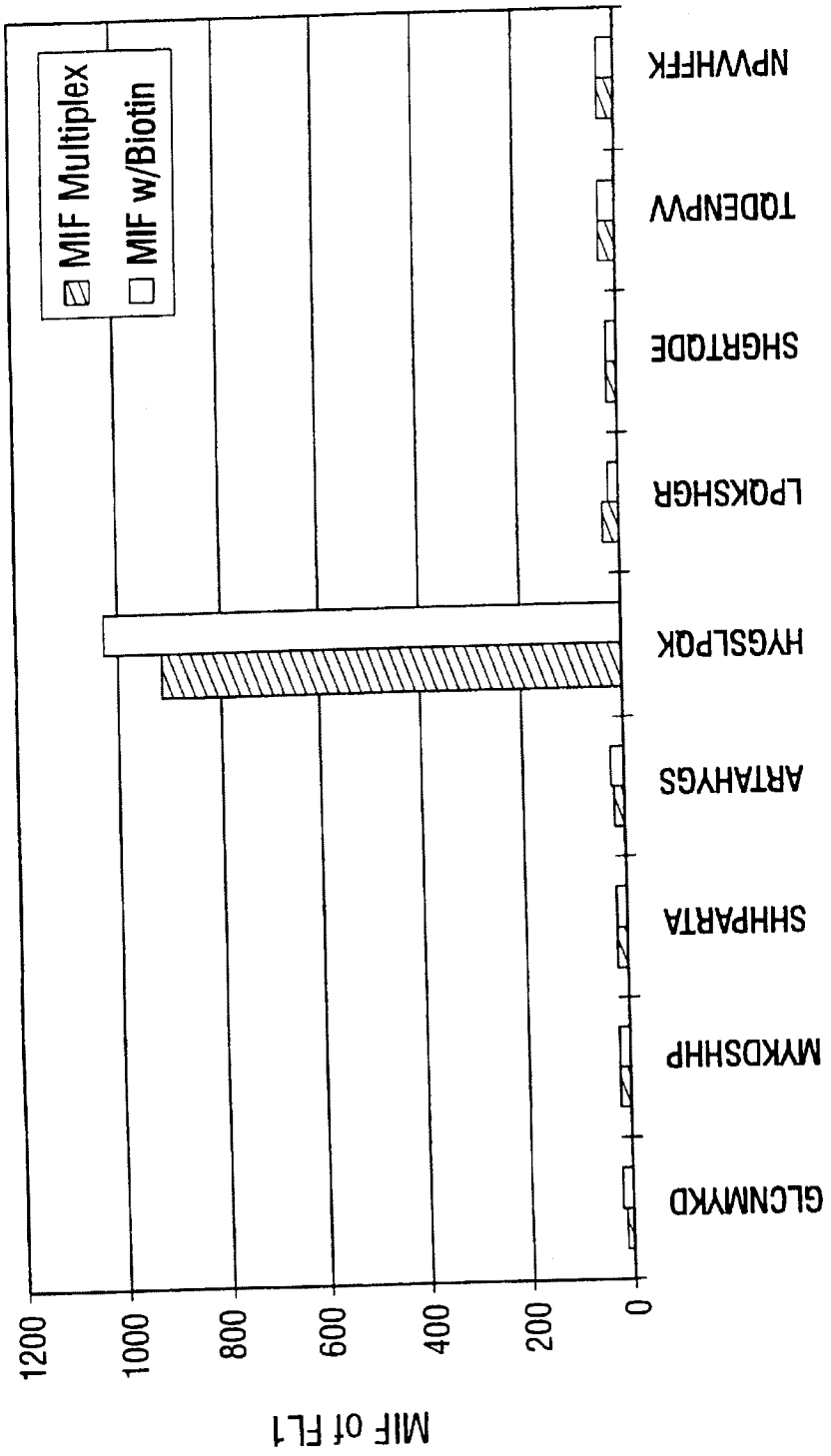

MIF Multiplex
MIF w/Biotin
MIF of FL1
1200
1000
800
600
400
200
0
GLCNMYKD
MYKDSHHP
SHHPARTA
ARTAHYGS
HYGSLPQK
LPQKSHGR
SHGRTQDE
TQDENPVV
NPVVHFFK
Peptide

FIG. 23

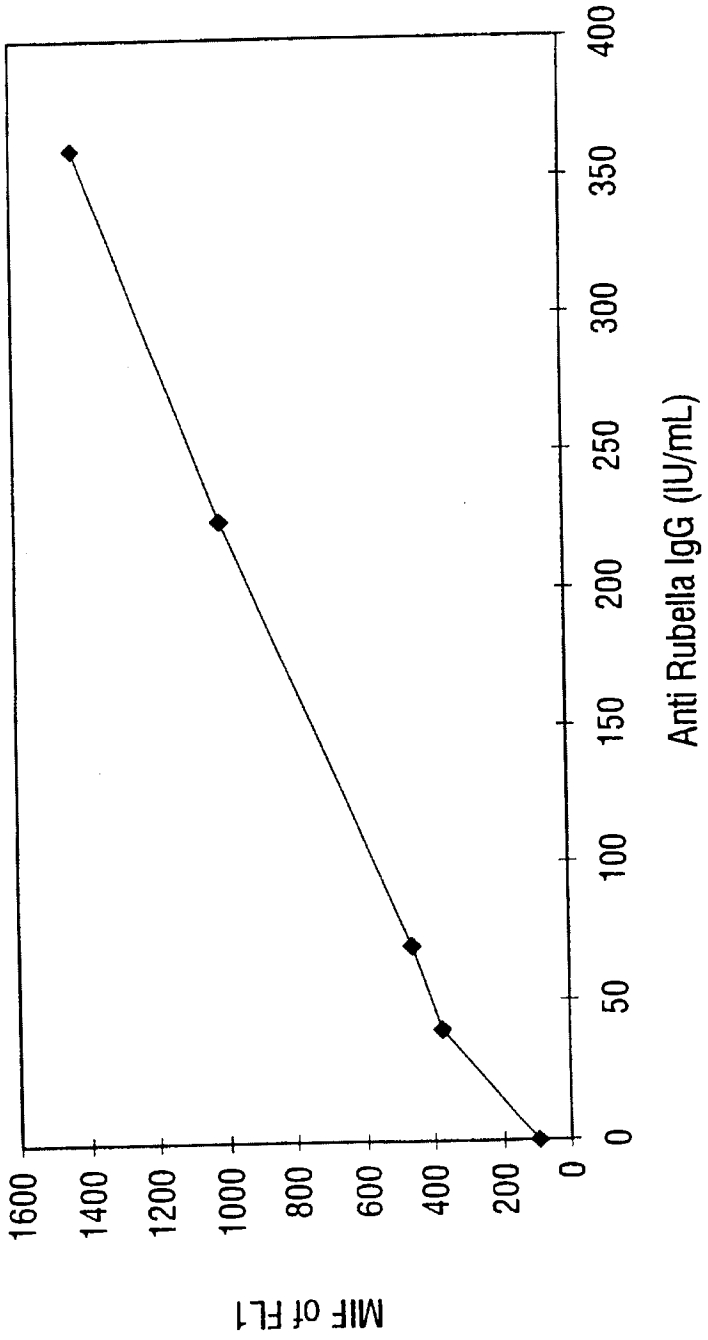

MIF of FL1
1600
1400
1200
1000
800
600
400
200
0
0
50
100
150
200
250
300
350
400
Anti Rubella IgG (IU/mL)

FIG. 24

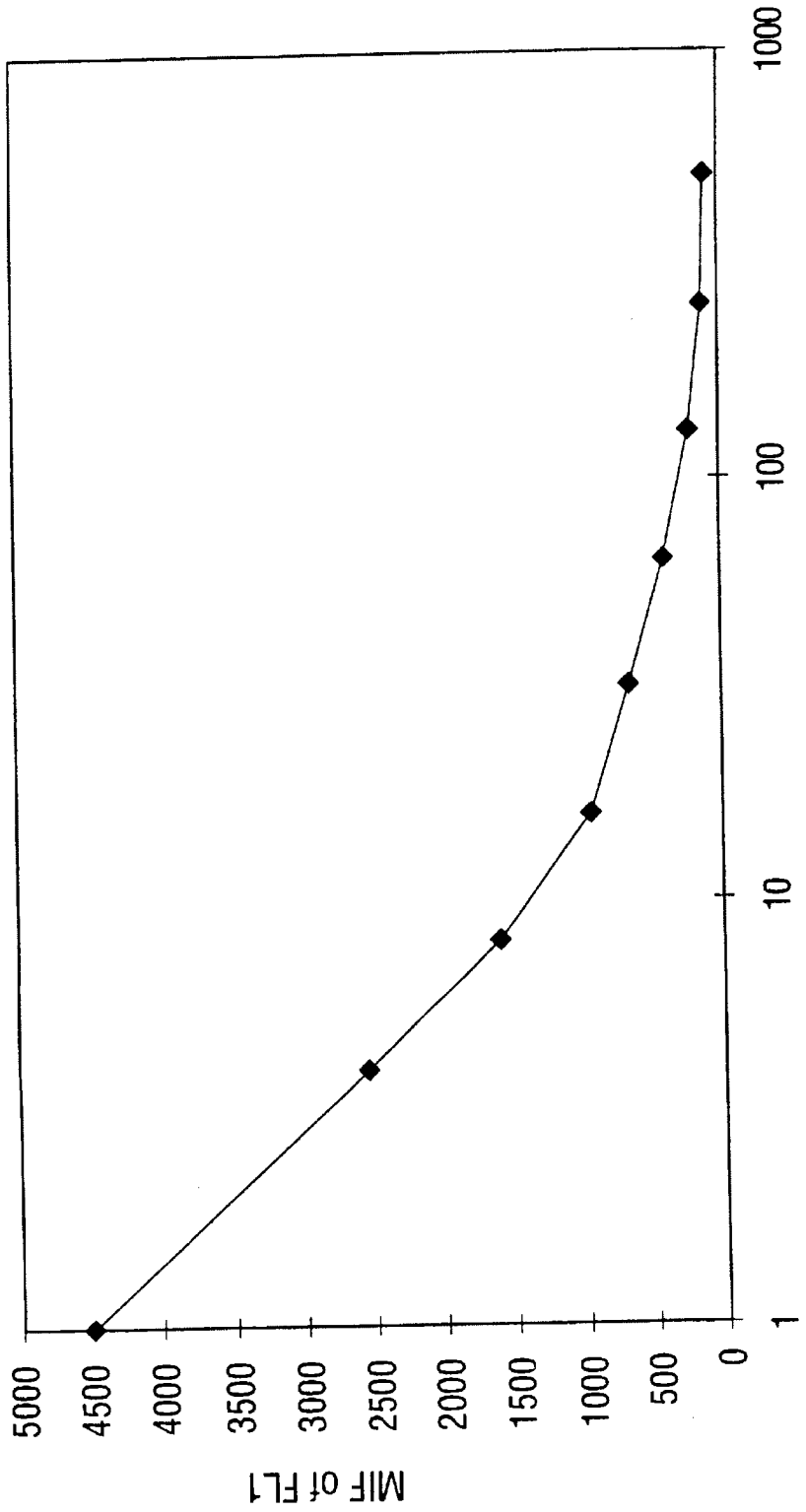

MIF of FL1
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
1
10
100
1000
Reciprocal of Calibrator 70 Dilution

FIG. 25

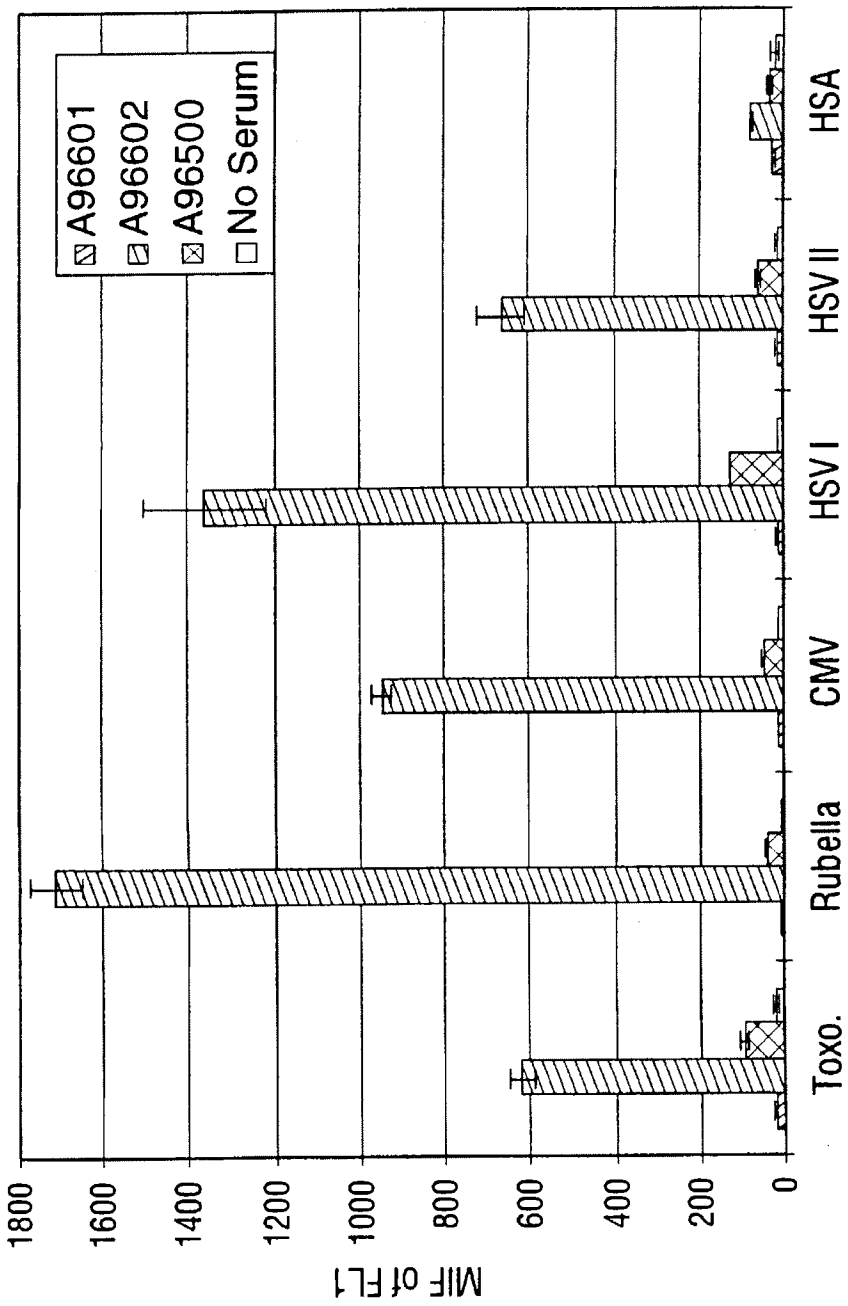

A96601
A96602
A96500
No Serum
MIF of FL1
1800
1600
1400
1200
1000
800
600
400
200
0
Toxo.
Rubella
CMV
HSV I
HSV II
HSA
ToRCH Antigen

FIG. 26A

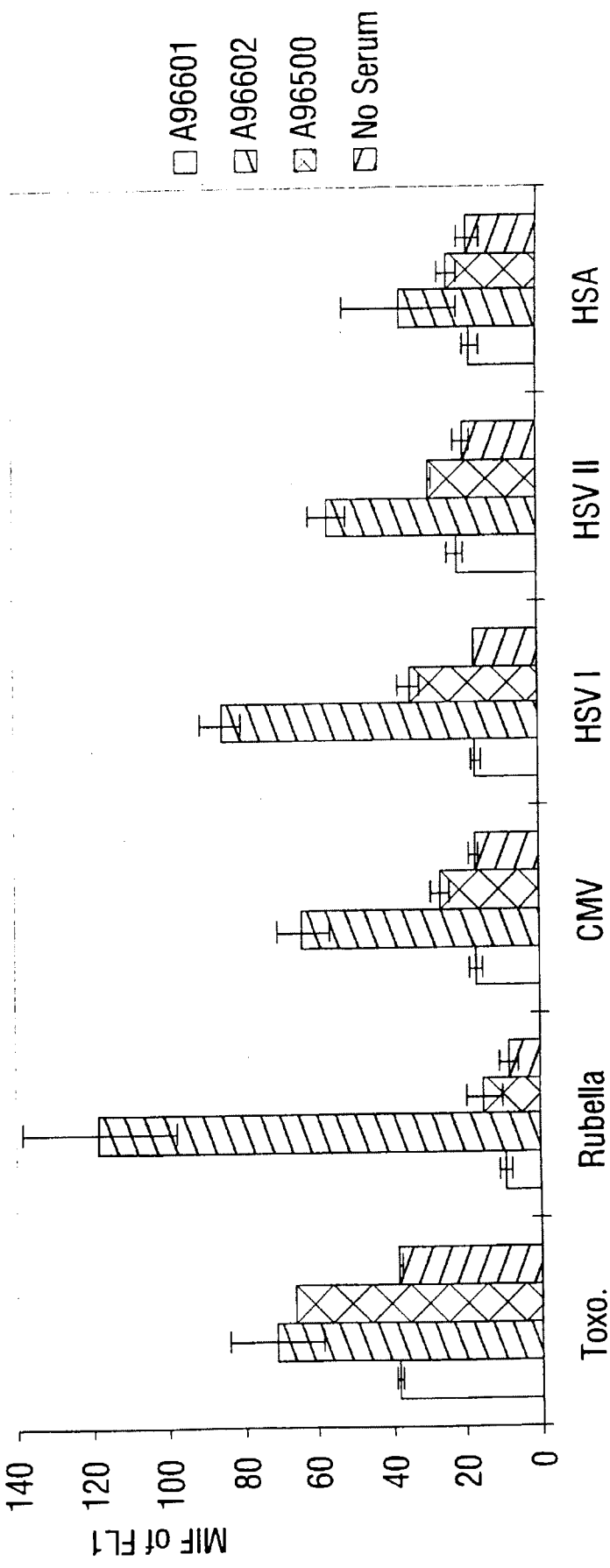

A96601
A96602
A96500
No Serum
MIF of FL1
140
120
100
80
60
40
20
0
Toxo.
Rubella
CMV
HSV I
HSV II
HSA
ToRCH Antigen

FIG. 26B

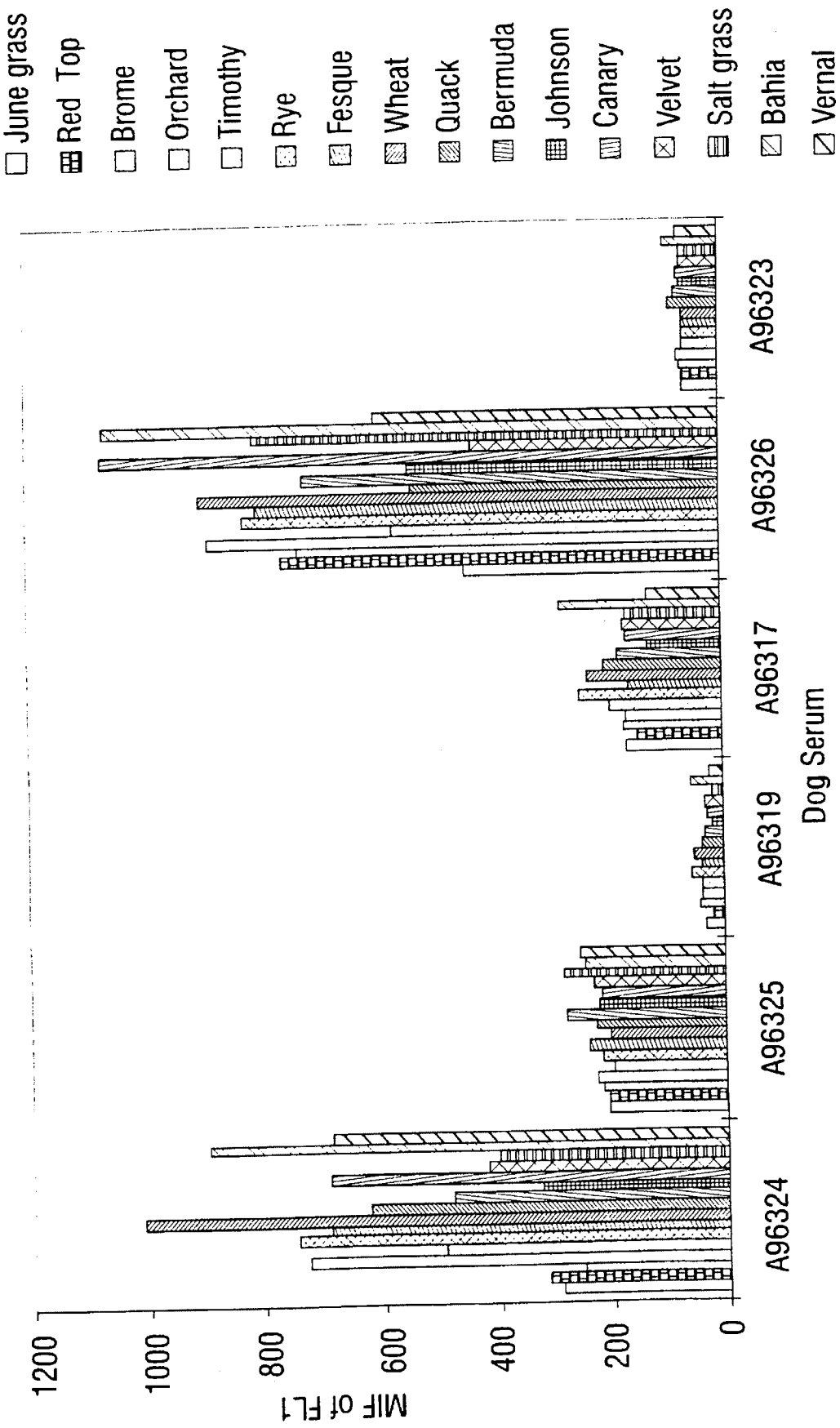

June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
MIF of FL1
1200
1000
800
600
400
200
0
A96324
A96325
A96319
A96317
A96326
A96323
Dog Serum

FIG. 27

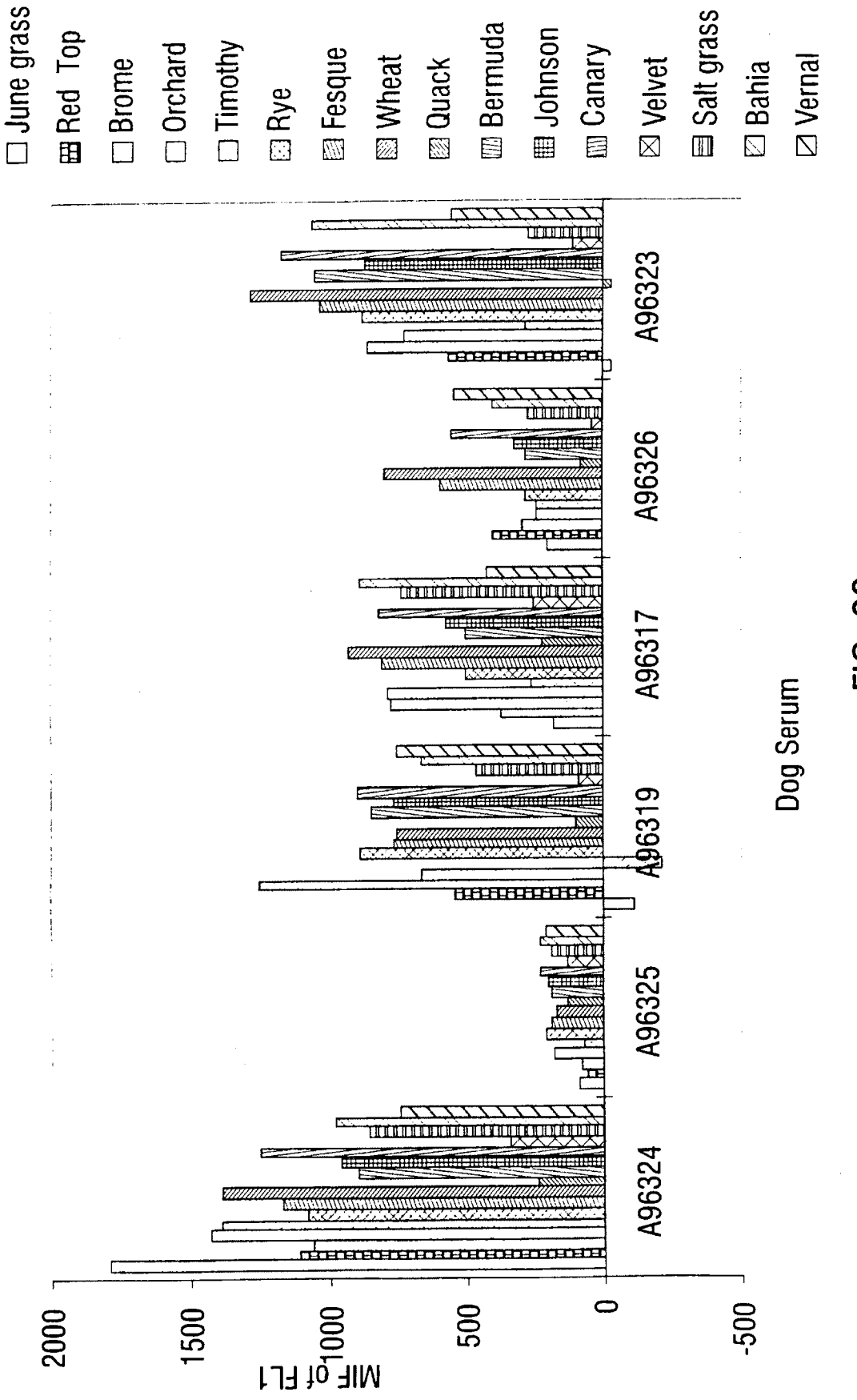

June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
2000
1500
1000
500
0
-500
MIF of FL1
A96324
A96325
A96319
A96317
A96326
A96323
Dog Serum

FIG. 28

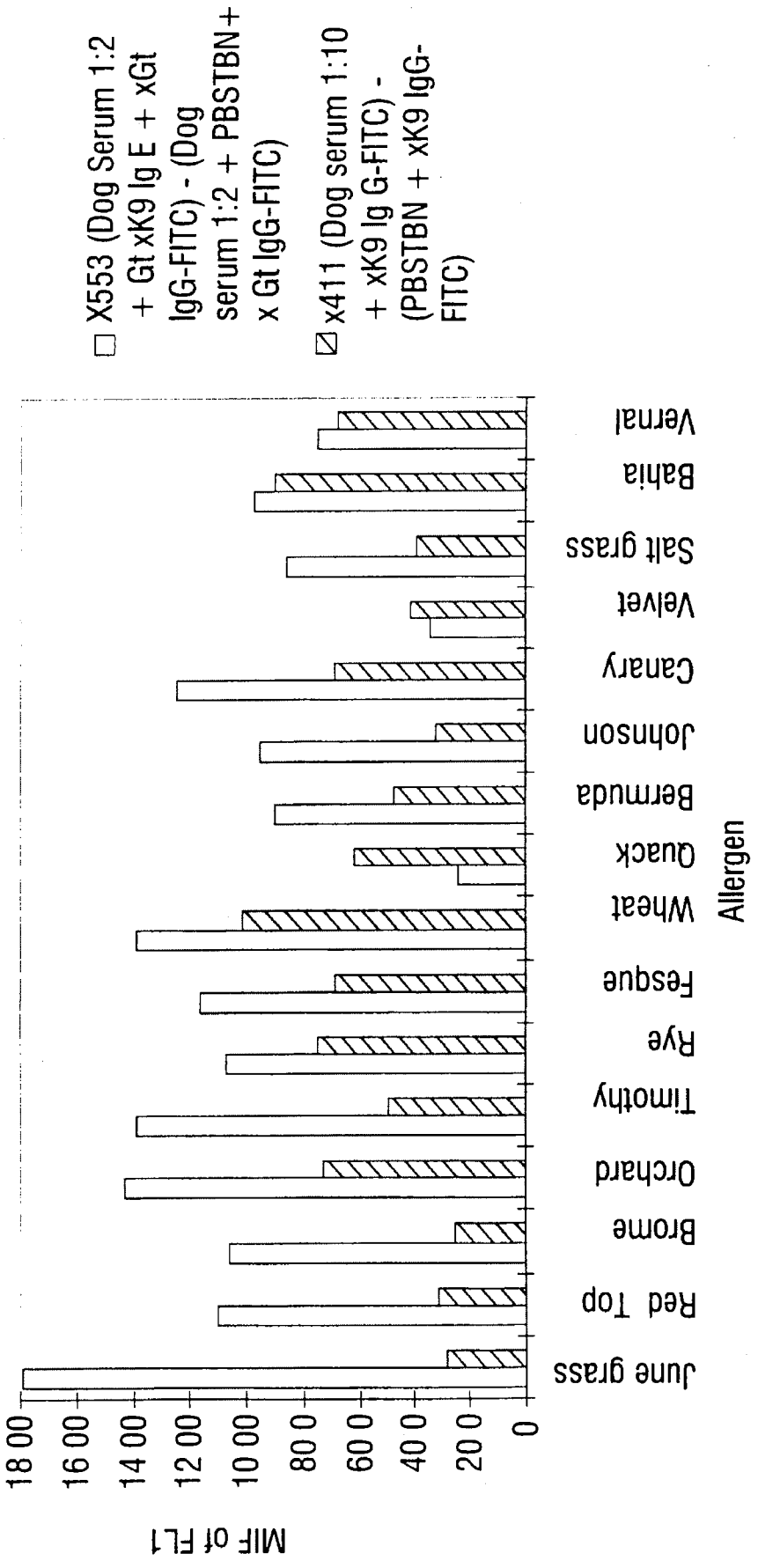

□ X553 (Dog Serum 1:2 + Gt xK9 Ig E + xGt IgG-FITC) - (Dog serum 1:2 + PBSTBN+ x Gt IgG-FITC)
☒ x411 (Dog serum 1:10 + xK9 Ig G-FITC) - (PBSTBN + xK9 IgG-FITC)
MIF of FL1
18 00
16 00
14 00
12 00
10 00
80 0
60 0
40 0
20 0
0
June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
Allergen

FIG. 29

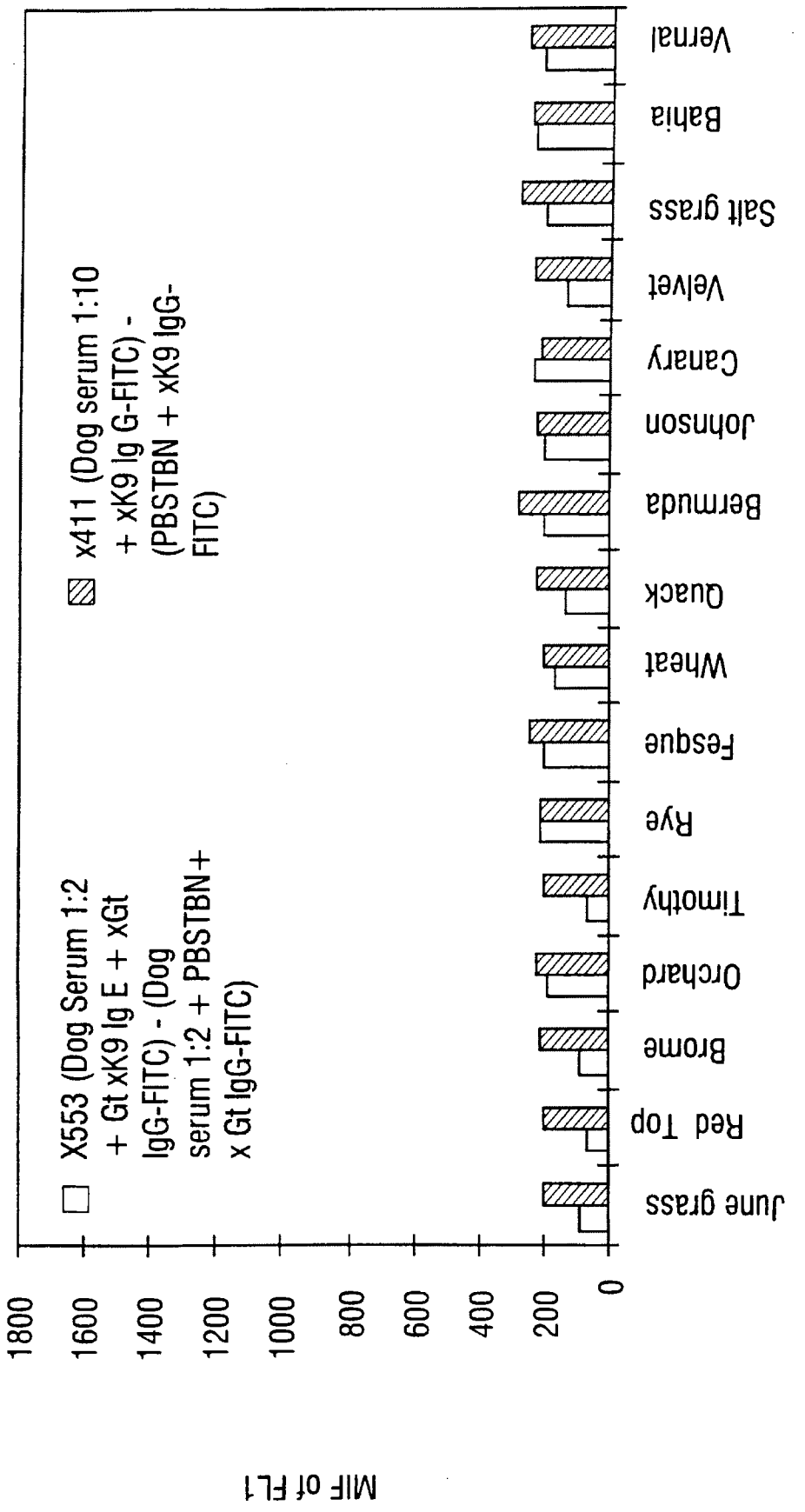

MIF of FL1
1800
1600
1400
1200
1000
800
600
400
200
0
X553 (Dog Serum 1:2 + Gt xK9 Ig E + xGt IgG-FITC) - (Dog serum 1:2 + PBSTBN+ x Gt IgG-FITC)
x411 (Dog serum 1:10 + xK9 Ig G-FITC) - (PBSTBN + xK9 IgG-FITC)
June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
Allergen

FIG. 30

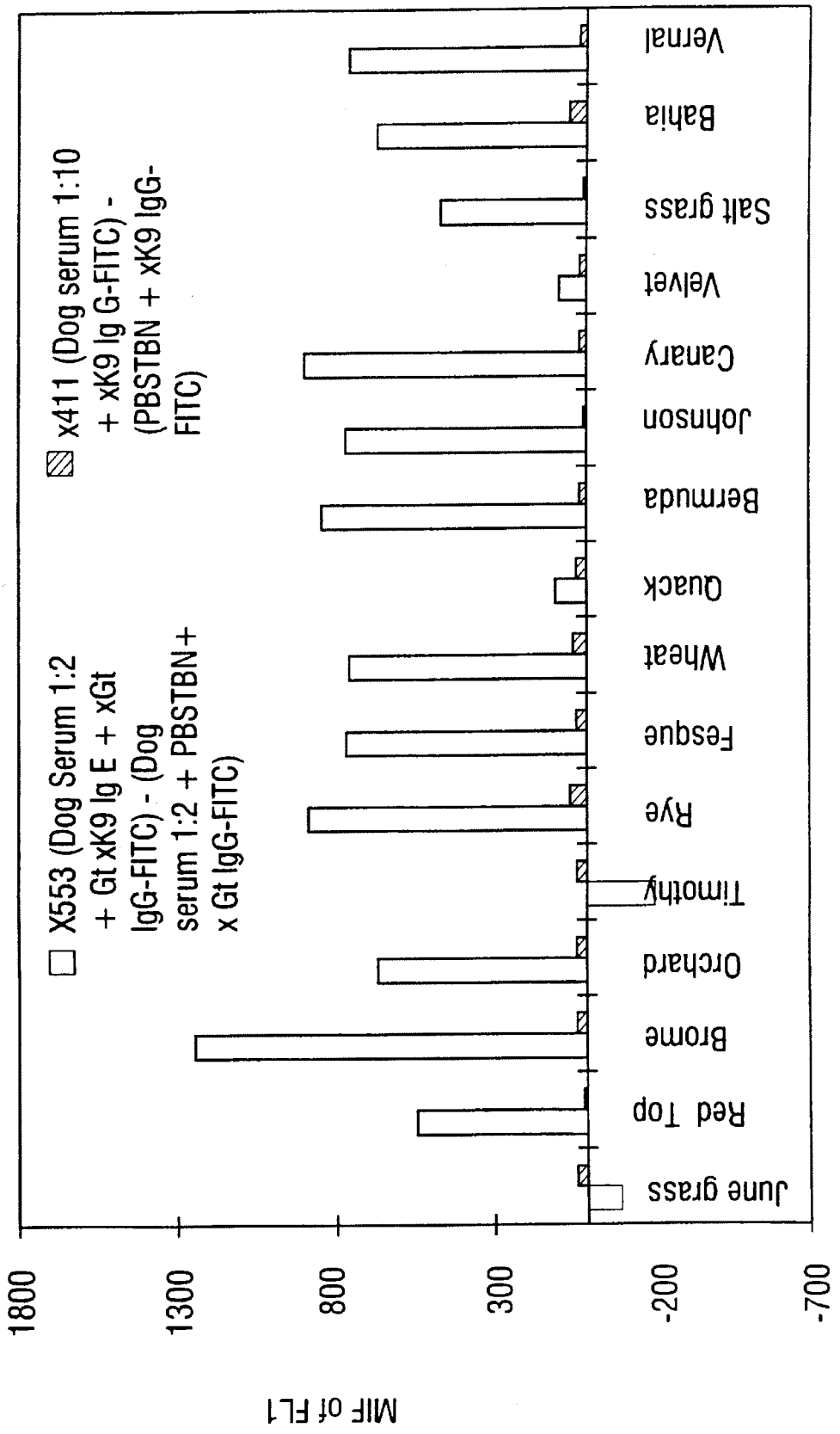

X553 (Dog Serum 1:2 + Gt xK9 Ig E + xGt IgG-FITC) - (Dog serum 1:2 + PBSTBN+ x Gt IgG-FITC)
x411 (Dog serum 1:10 + xK9 Ig G-FITC) - (PBSTBN + xK9 IgG-FITC)
MIF of FL1
1800
1300
800
300
-200
-700
June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
Allergen

FIG. 31

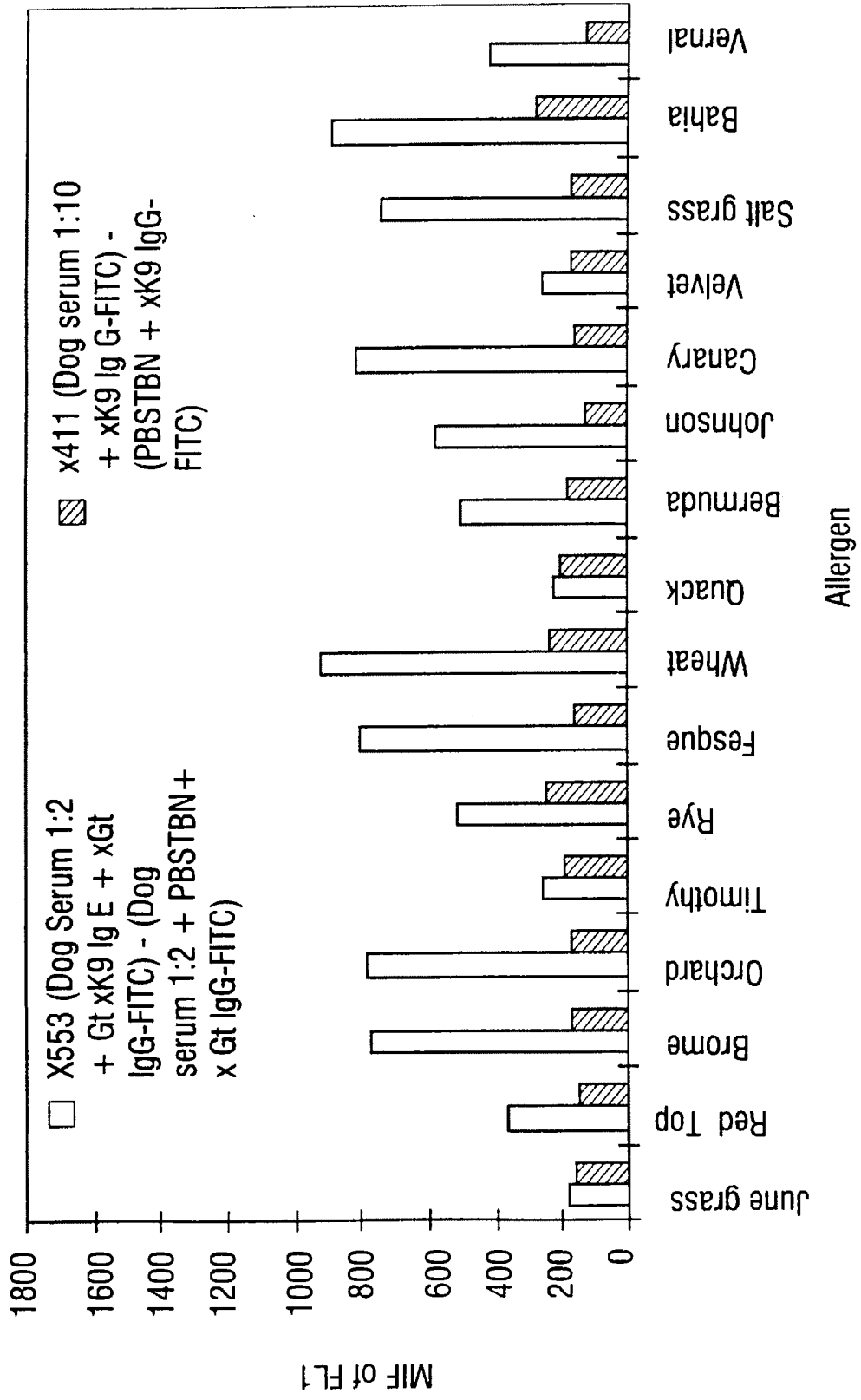

MIF of FL1
1800
1600
1400
1200
1000
800
600
400
200
0
X553 (Dog Serum 1:2 + Gt xK9 Ig E + xGt IgG-FITC) - (Dog serum 1:2 + PBSTBN + x Gt IgG-FITC)
x411 (Dog serum 1:10 + xK9 Ig G-FITC) - (PBSTBN + xK9 IgG-FITC)
June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
Allergen

FIG. 32

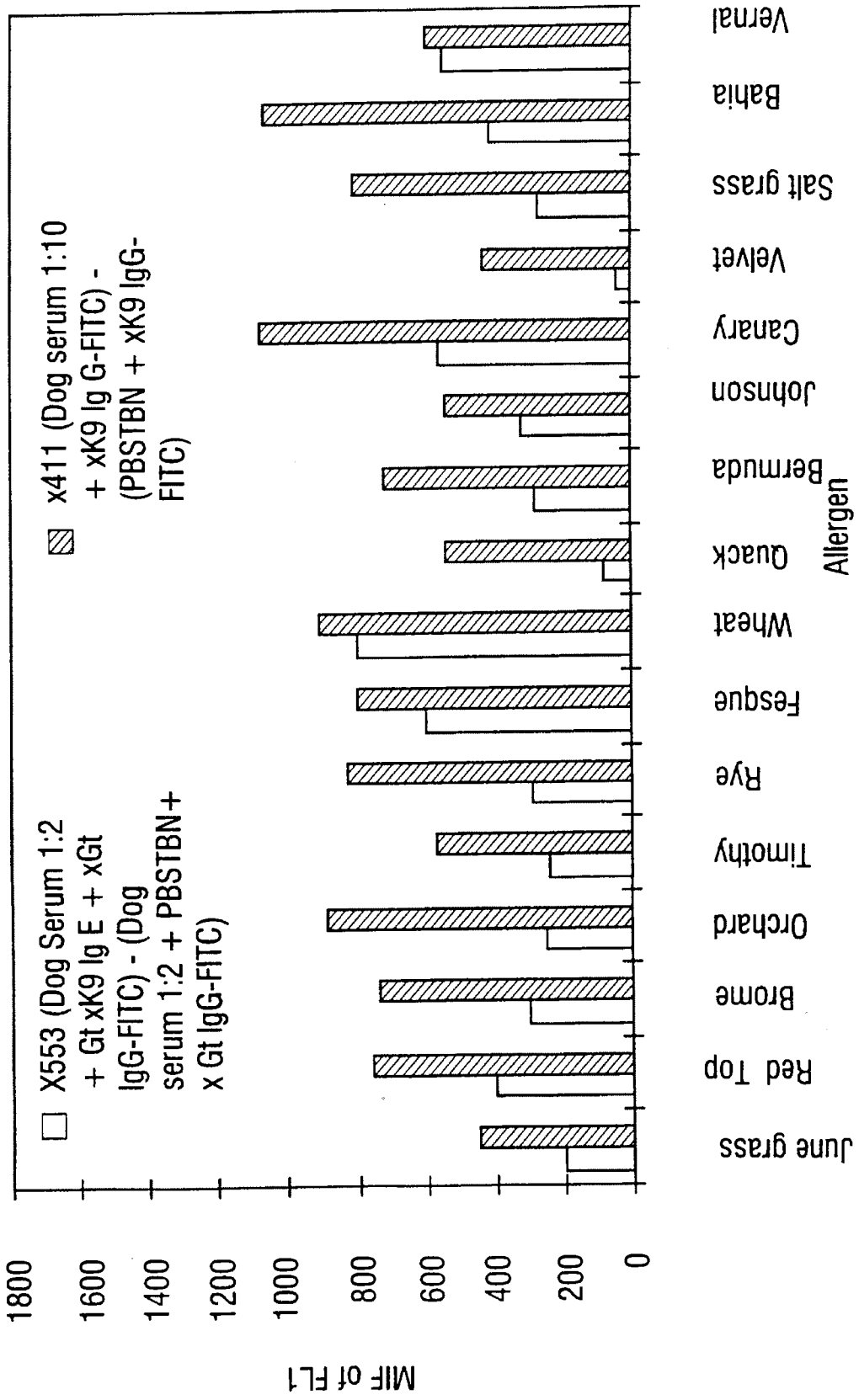

MIF of FL1
1800
1600
1400
1200
1000
800
600
400
200
0
X553 (Dog Serum 1:2 + Gt xK9 Ig E + xGt IgG-FITC) - (Dog serum 1:2 + PBSTBN+ x Gt IgG-FITC)
x411 (Dog serum 1:10 + xK9 Ig G-FITC) - (PBSTBN + xK9 IgG-FITC)
June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
Allergen

FIG. 33

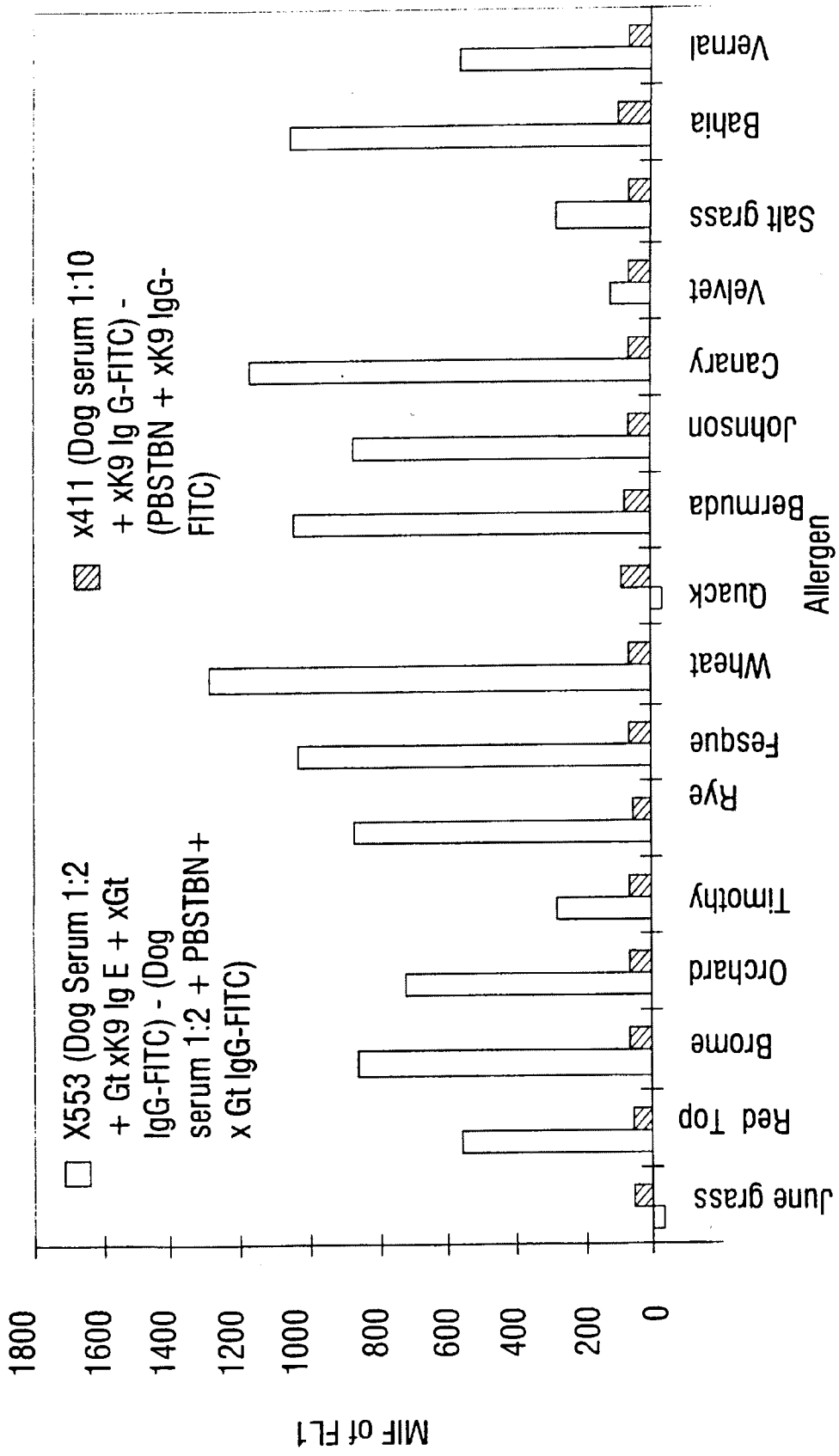

MIF of FL1
1800
1600
1400
1200
1000
800
600
400
200
0
X553 (Dog Serum 1:2 + Gt xK9 Ig E + xGt IgG-FITC) - (Dog serum 1:2 + PBSTBN + x Gt IgG-FITC)
x411 (Dog serum 1:10 + xK9 Ig G-FITC) - (PBSTBN + xK9 IgG-FITC)
June grass
Red Top
Brome
Orchard
Timothy
Rye
Fesque
Wheat
Quack
Bermuda
Johnson
Canary
Velvet
Salt grass
Bahia
Vernal
Allergen

FIG. 34

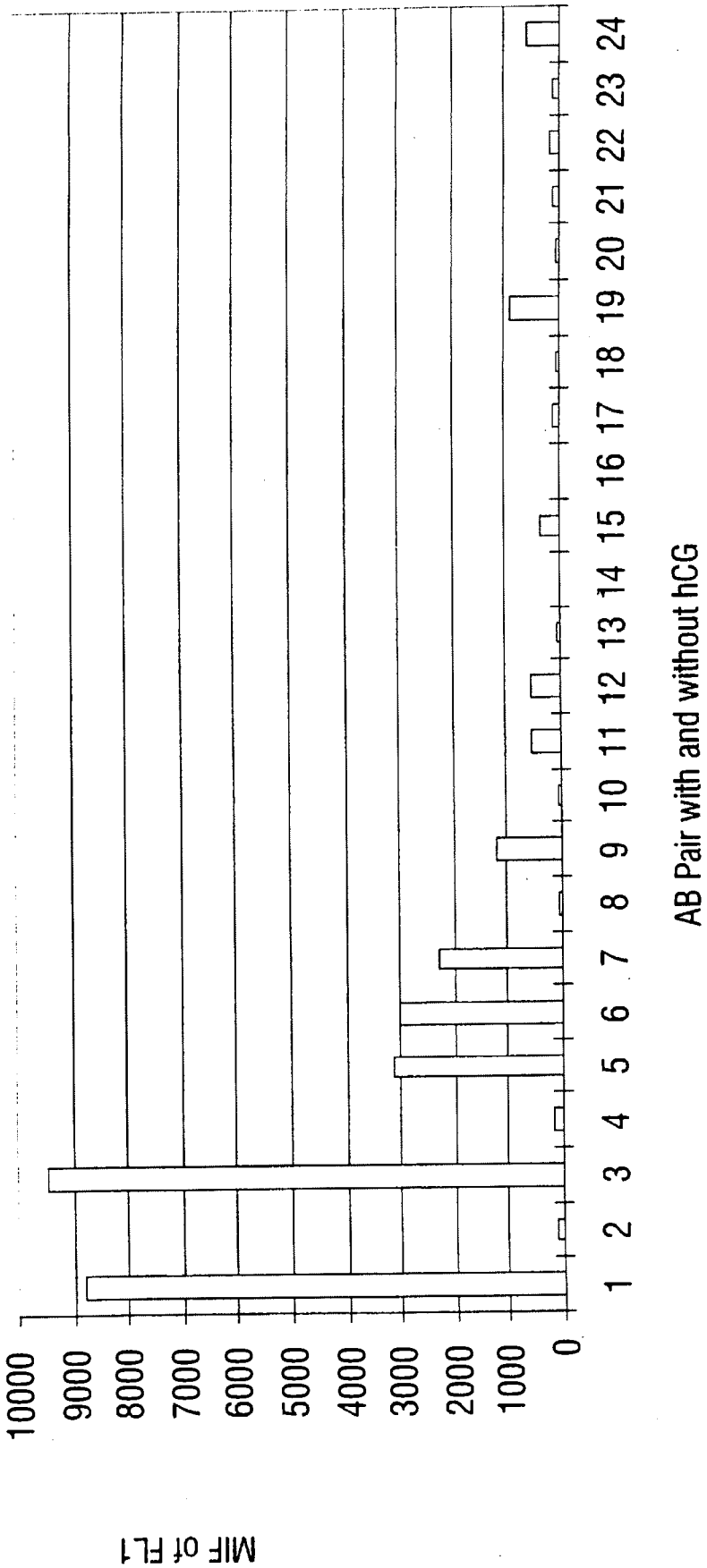

MIF of FL1
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
AB Pair with and without hCG

FIG. 35

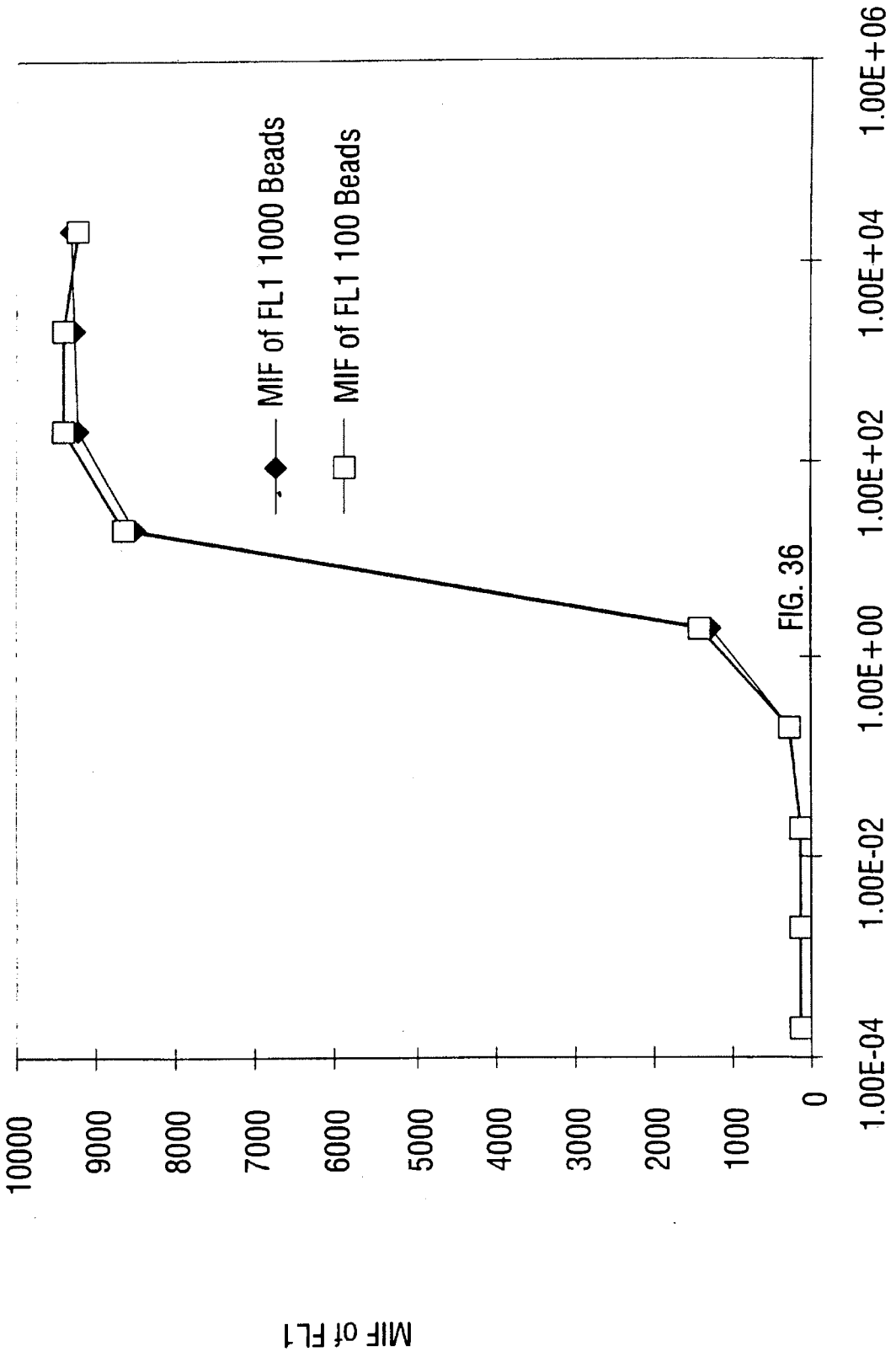

MIF of FL1 1000 Beads
MIF of FL1 100 Beads
MIF of FL1
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
1.00E-04
1.00E-02
1.00E+00
1.00E+02
1.00E+04
1.00E+06
hCG conc. (ng/mL)
FIG. 36

FIG. 36

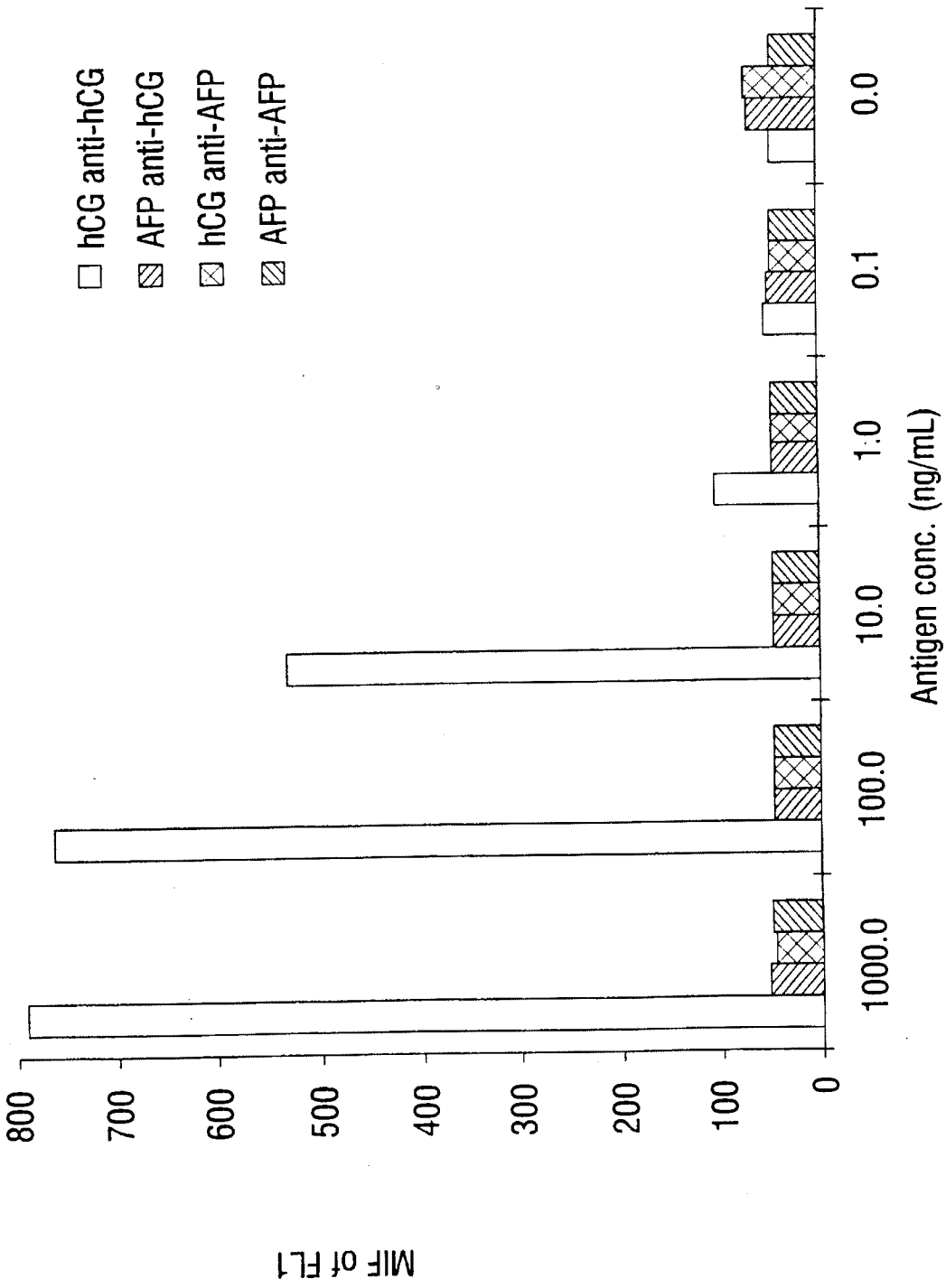

hCG anti-hCG
AFP anti-hCG
hCG anti-AFP
AFP anti-AFP
MIF of FL1
800
700
600
500
400
300
200
100
0
1000.0
100.0
10.0
1.0
0.1
0.0
Antigen conc. (ng/mL)

FIG. 37A

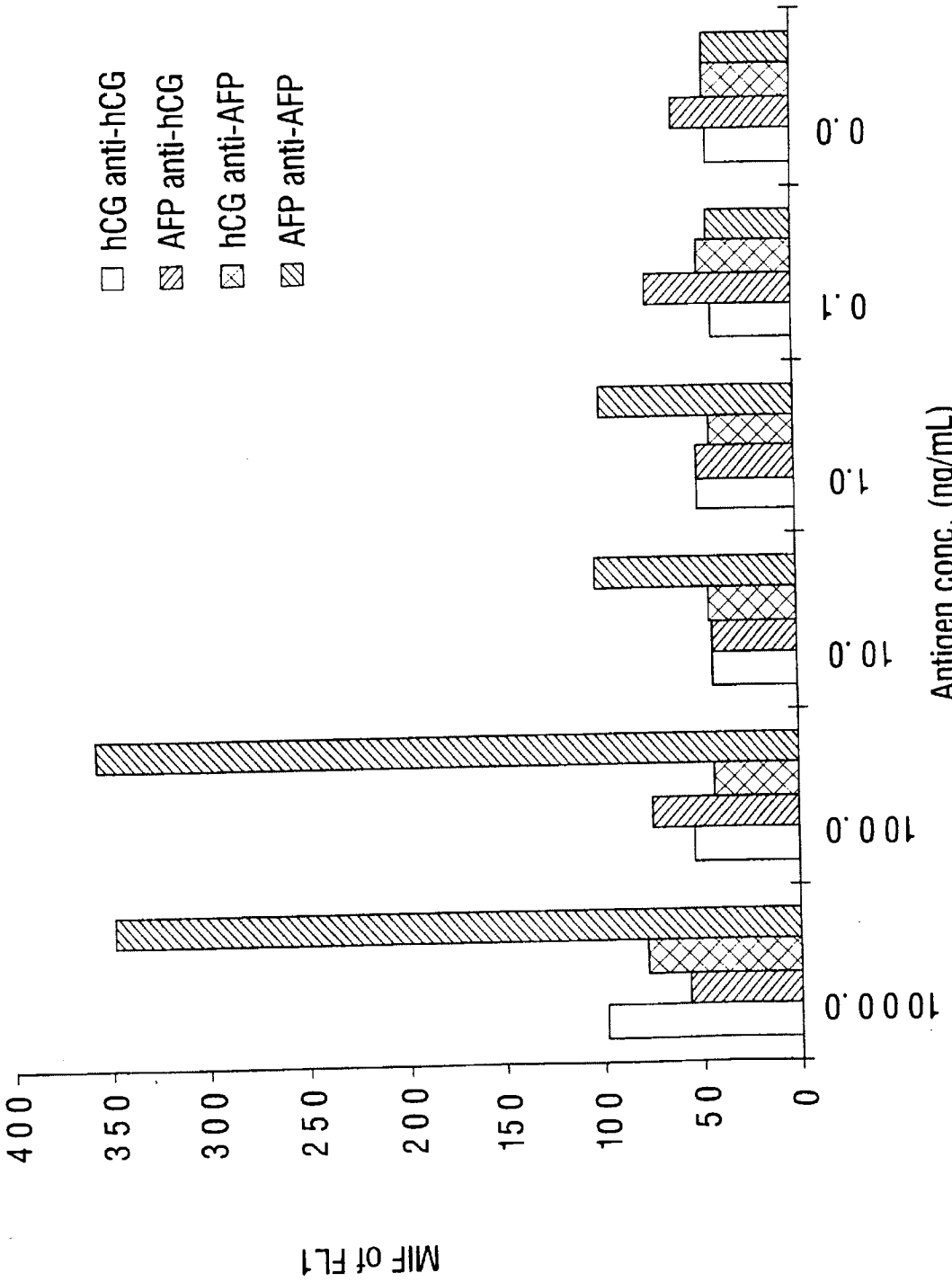

hCG anti-hCG
AFP anti-hCG
hCG anti-AFP
AFP anti-AFP
MIF of FL1
400
350
300
250
200
150
100
50
0
0.0
0.1
1.0
10.0
100.0
1000.0
Antigen conc. (ng/mL)

FIG. 37B

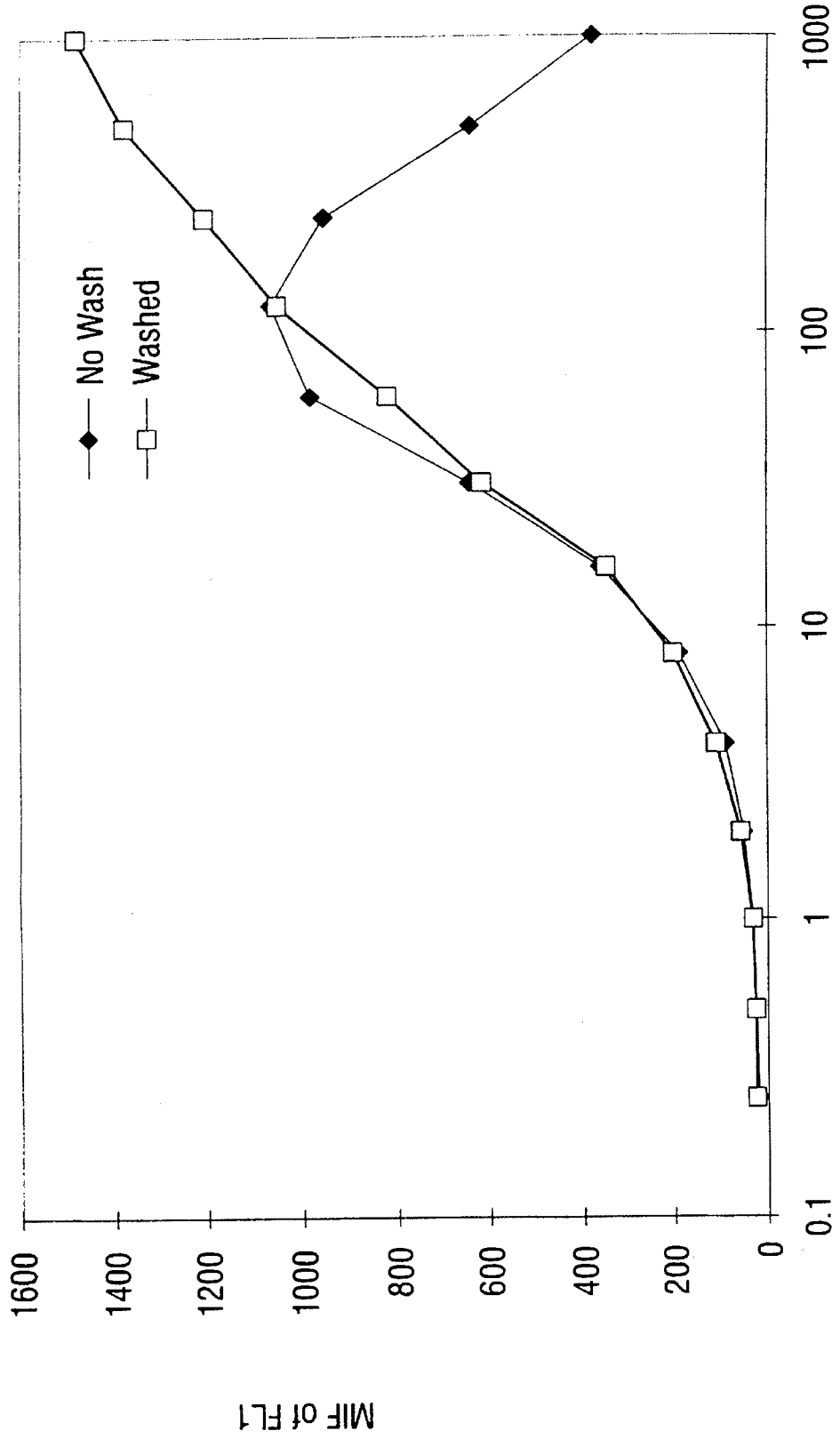

No Wash
Washed
MIF of FL1
1600
1400
1200
1000
800
600
400
200
0
0.1
1
10
100
1000
AFP Conc. (ng/mL)

FIG. 38A

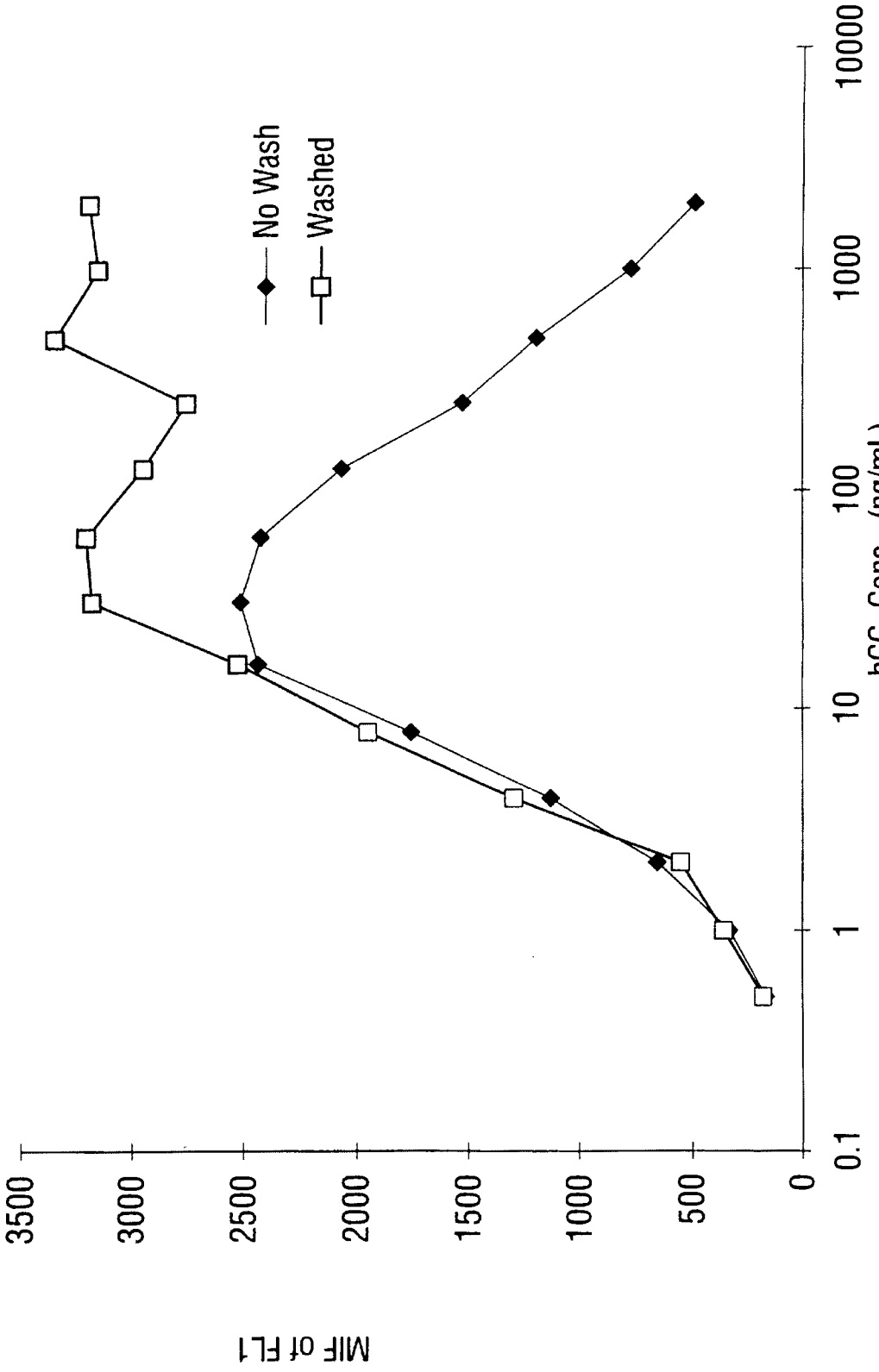

MIF of FL1
3500
3000
2500
2000
1500
1000
500
0
0.1
1
10
100
1000
10000
hCG Conc. (ng/mL)
No Wash
Washed

FIG. 38B

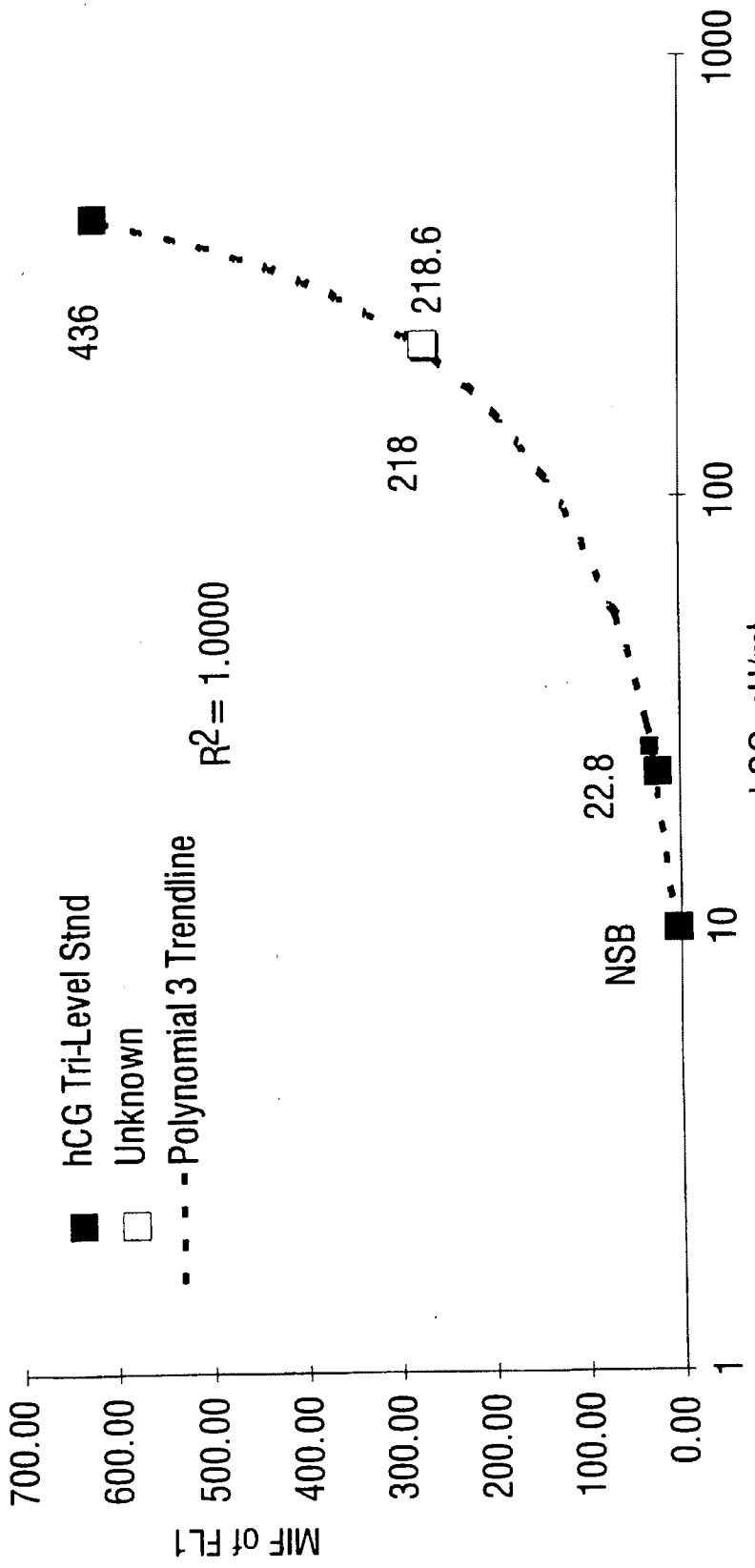

hCG Tri-Level Stnd
Unknown
Polynomial 3 Trendline
$R^2 = 1.0000$
436
218
218.6
22.8
NSB
MIF of FL1
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
1000
100
10
1
hCG mU/ml

FIG. 39A

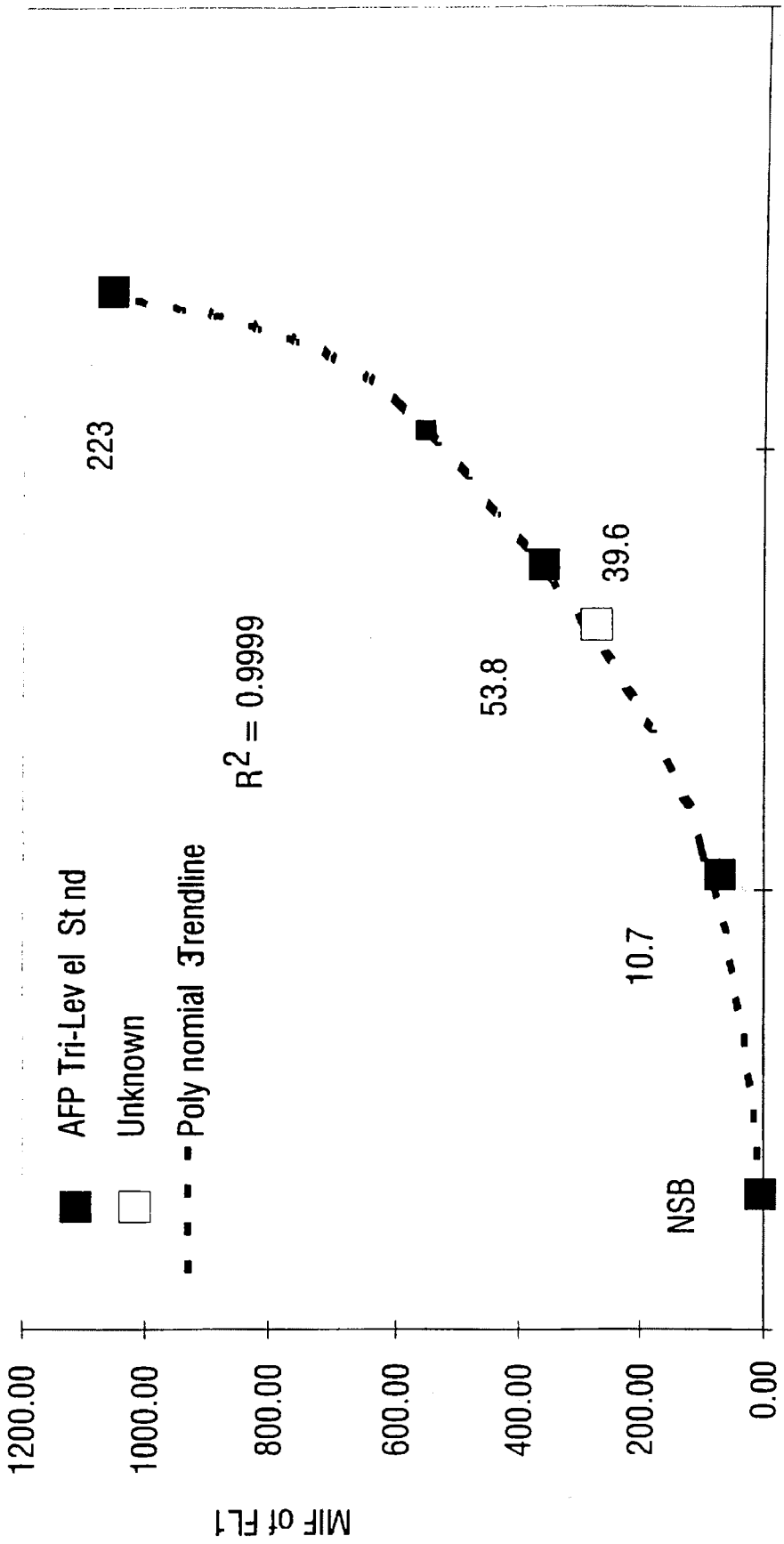

MIF of FL1
1200.00
1000.00
800.00
600.00
400.00
200.00
0.00
AFP Tri-Level Stnd
Unknown
Polynomial Trendline
R2 = 0.9999
223
53.8
39.6
10.7
NSB

FIG. 39B

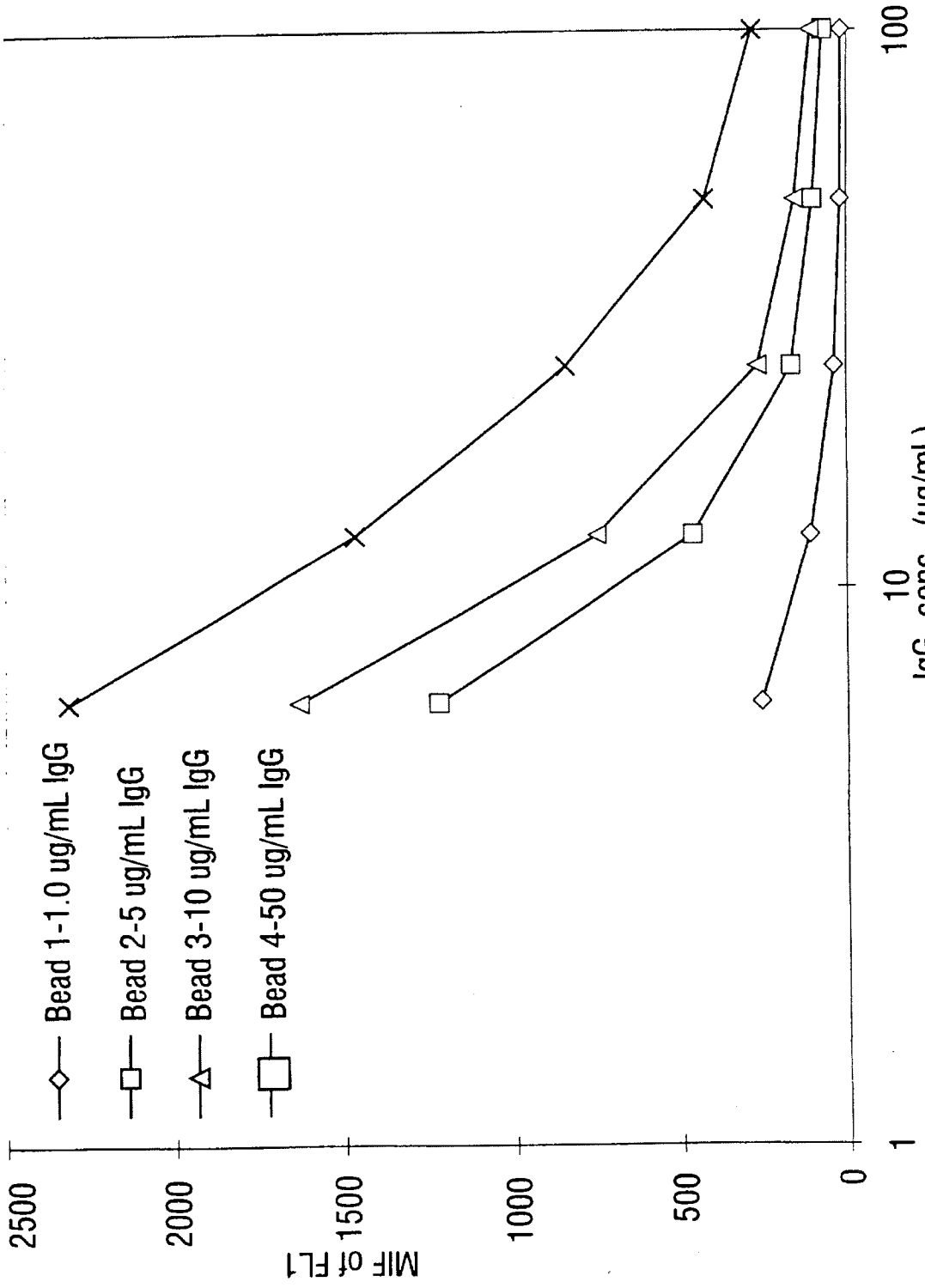

Bead 1-1.0 ug/mL IgG
Bead 2-5 ug/mL IgG
Bead 3-10 ug/mL IgG
Bead 4-50 ug/mL IgG
MIF of FL1
2500
2000
1500
1000
500
0
1
10
100
IgG- conc. (ug/mL)

FIG. 40

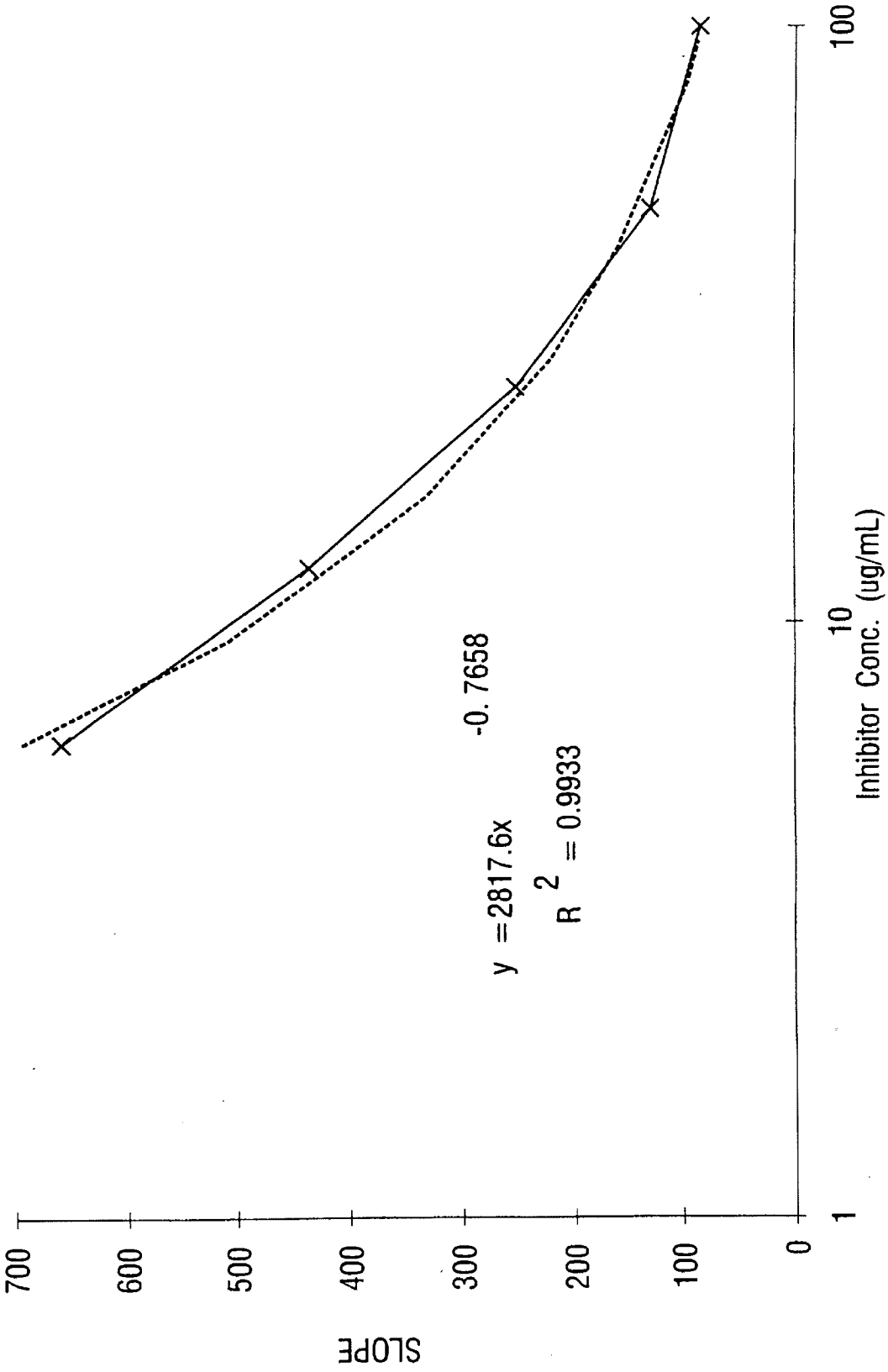

SLOPE
700
600
500
400
300
200
100
0
1
10
100
Inhibitor Conc. (ug/mL)
y = 2817.6x -0.7658
R 2 = 0.9933

FIG. 41

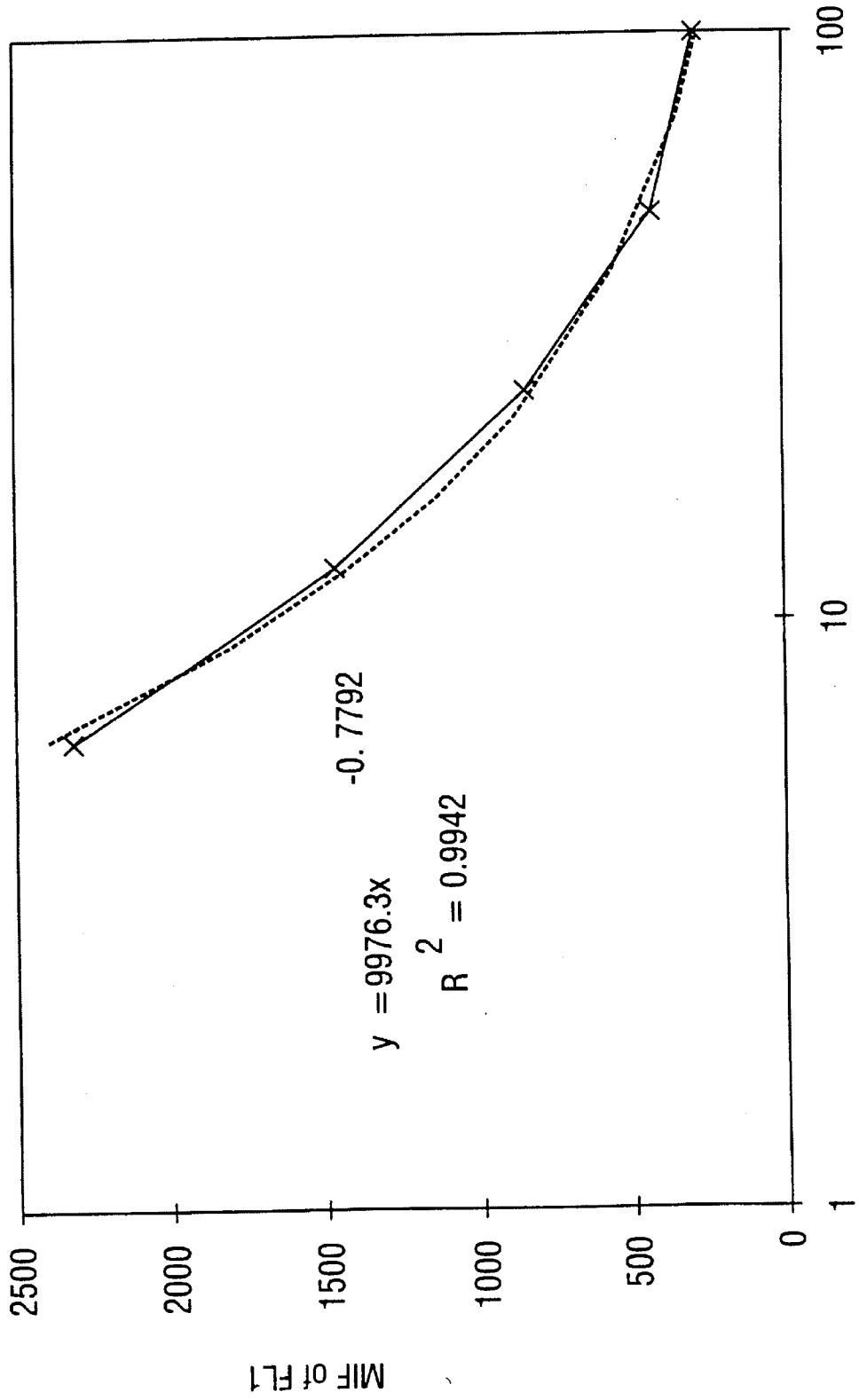

MIF of FL1
2500
2000
1500
1000
500
0
1
10
100
hIgG Conc. (ug/mL)
y =9976.3x -0. 7792
R 2 = 0.9942

FIG. 42

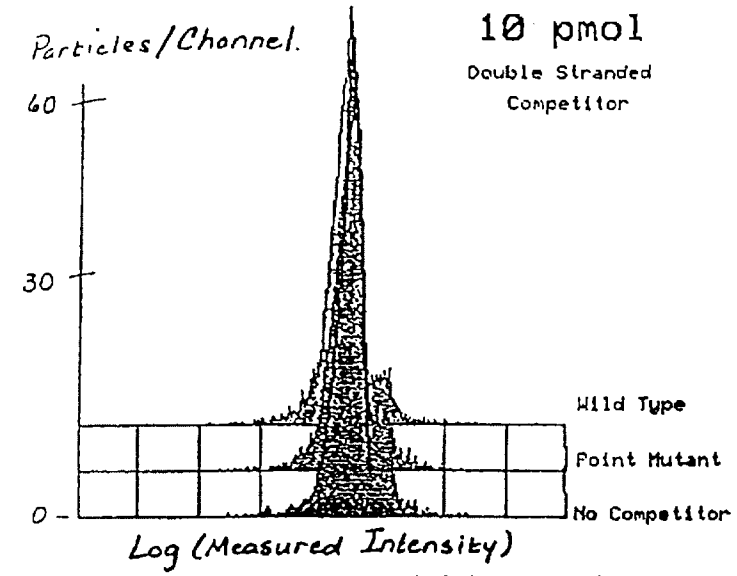


Fig. 43a

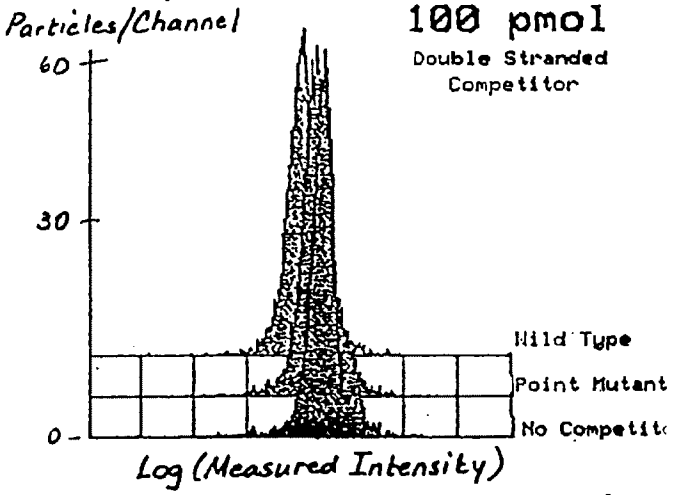


Fig. 43b

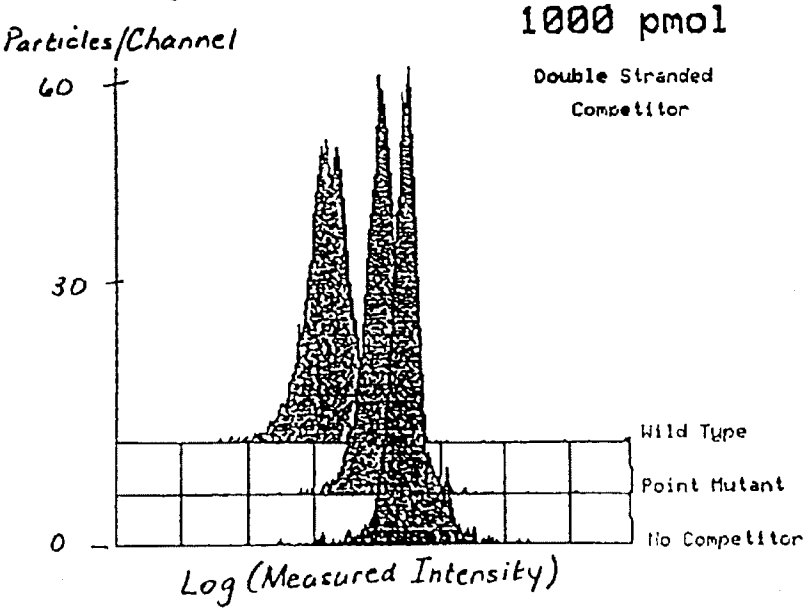


Fig. 43c

Fig. 44a

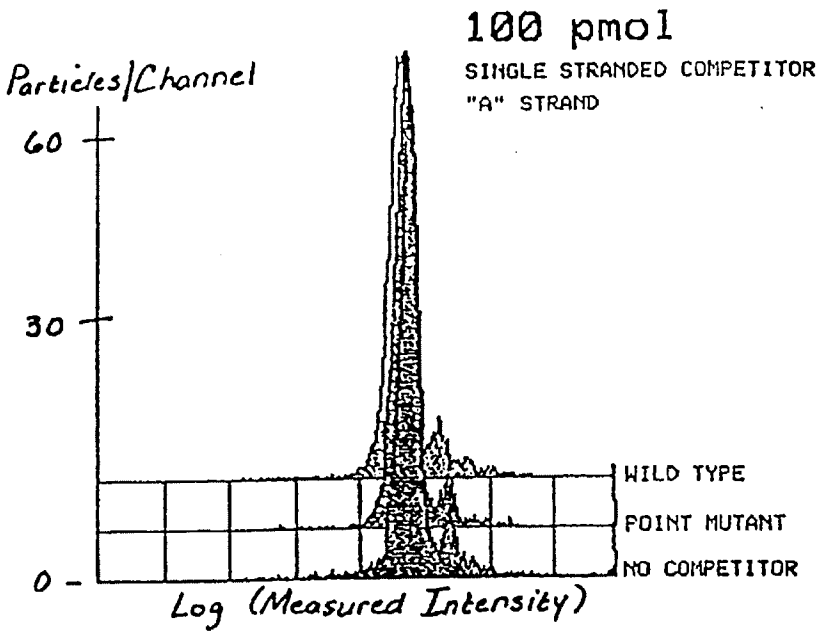


Fig. 44b

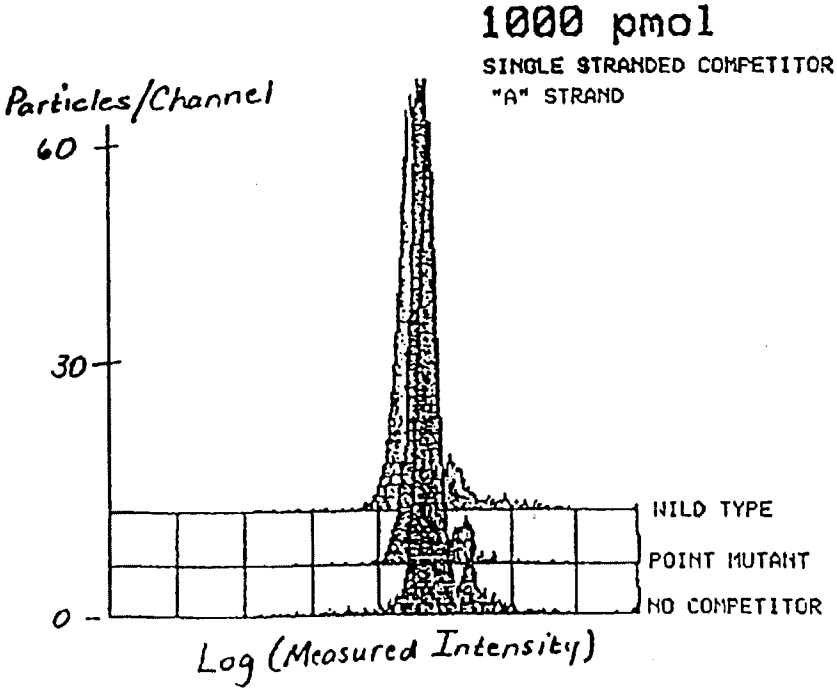


Fig. 45

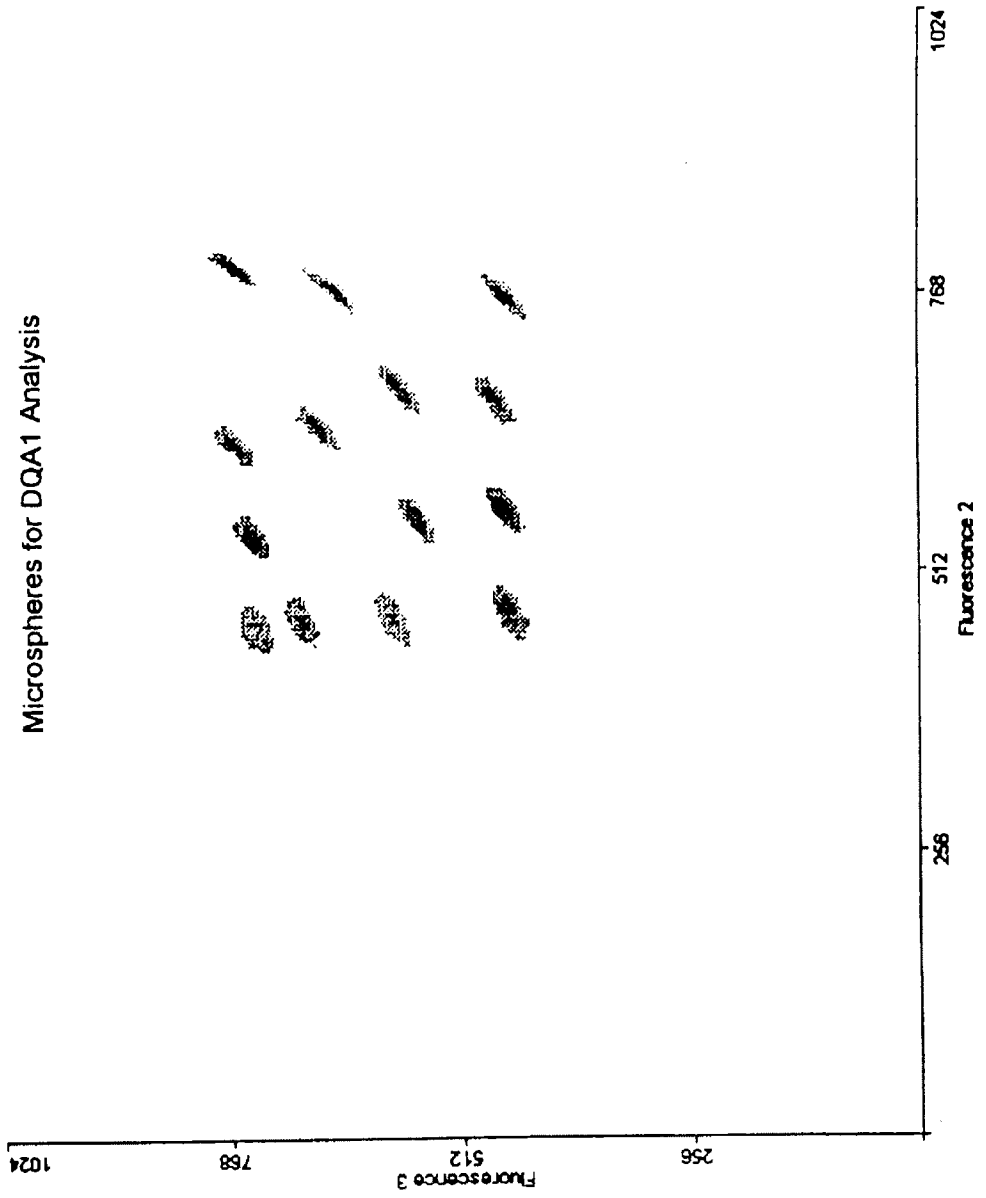

Microspheres for DQA1 Analysis
1024
768
512
256
Fluorescence 2
1024
768
512
256
Fluorescence 3

Fig. 46

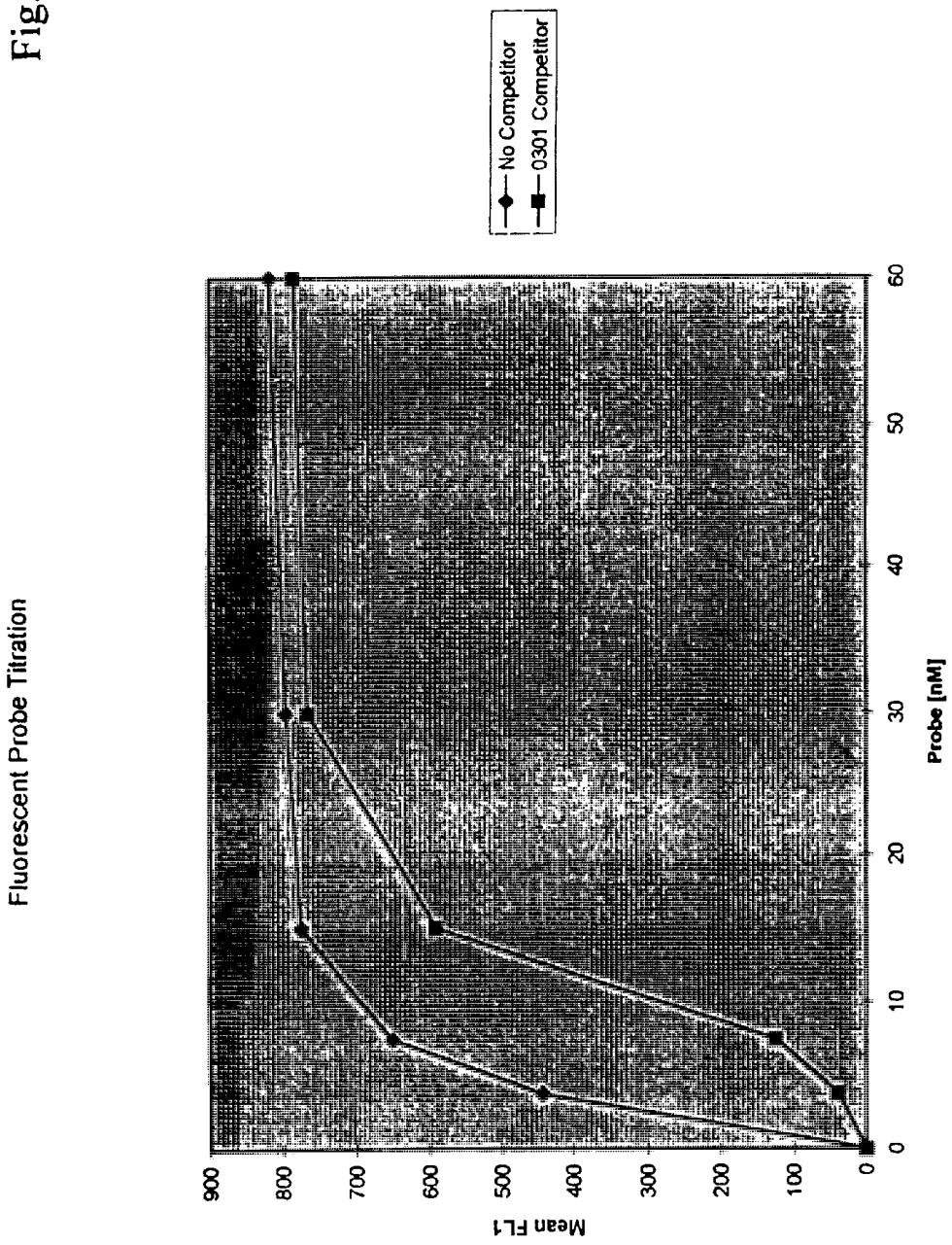

Fluorescent Probe Titration
No Competitor
0301 Competitor
Mean FL1
0
100
200
300
400
500
600
700
800
900
Probe [nM]
0
10
20
30
40
50
60

Fig. 47

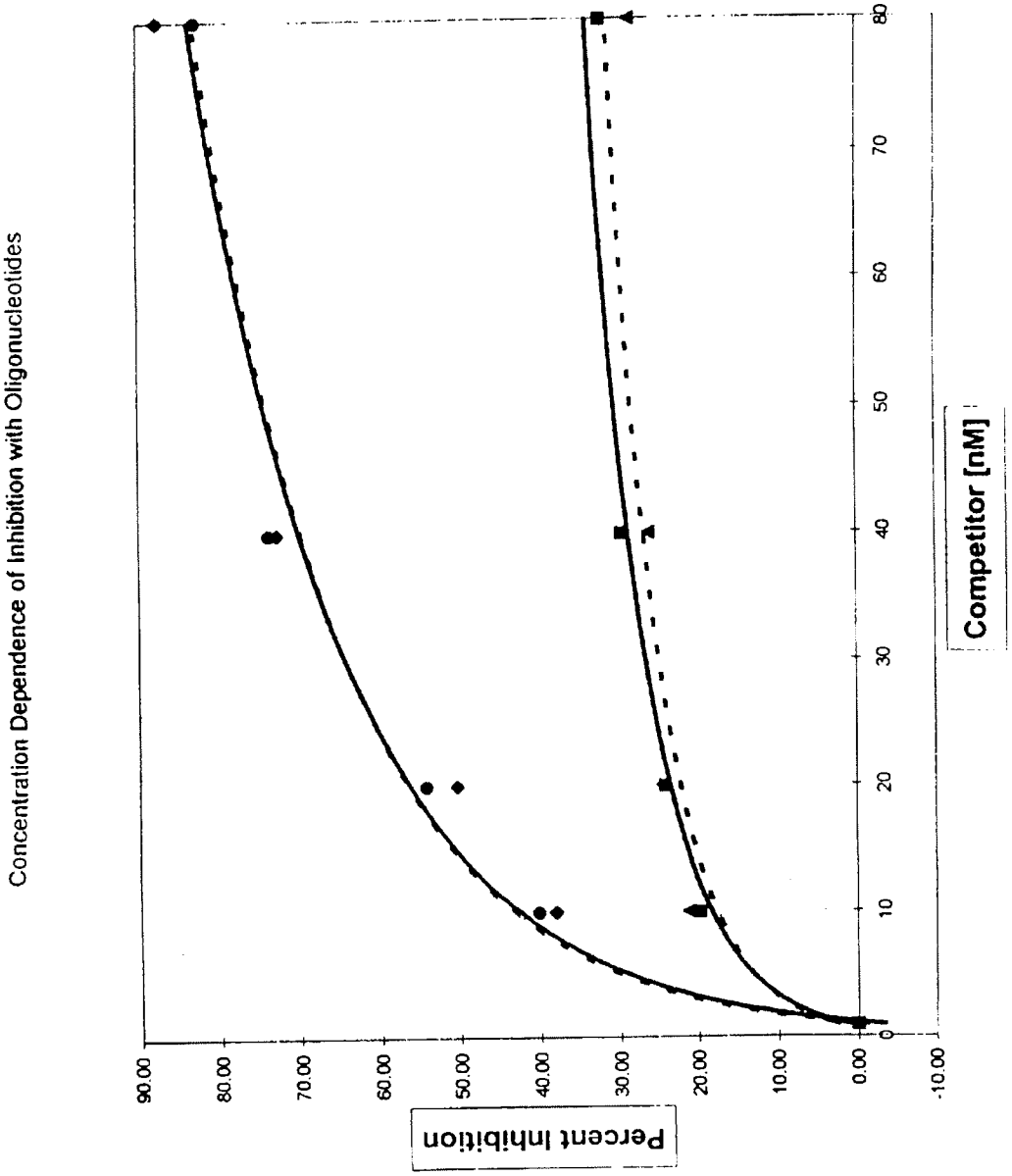
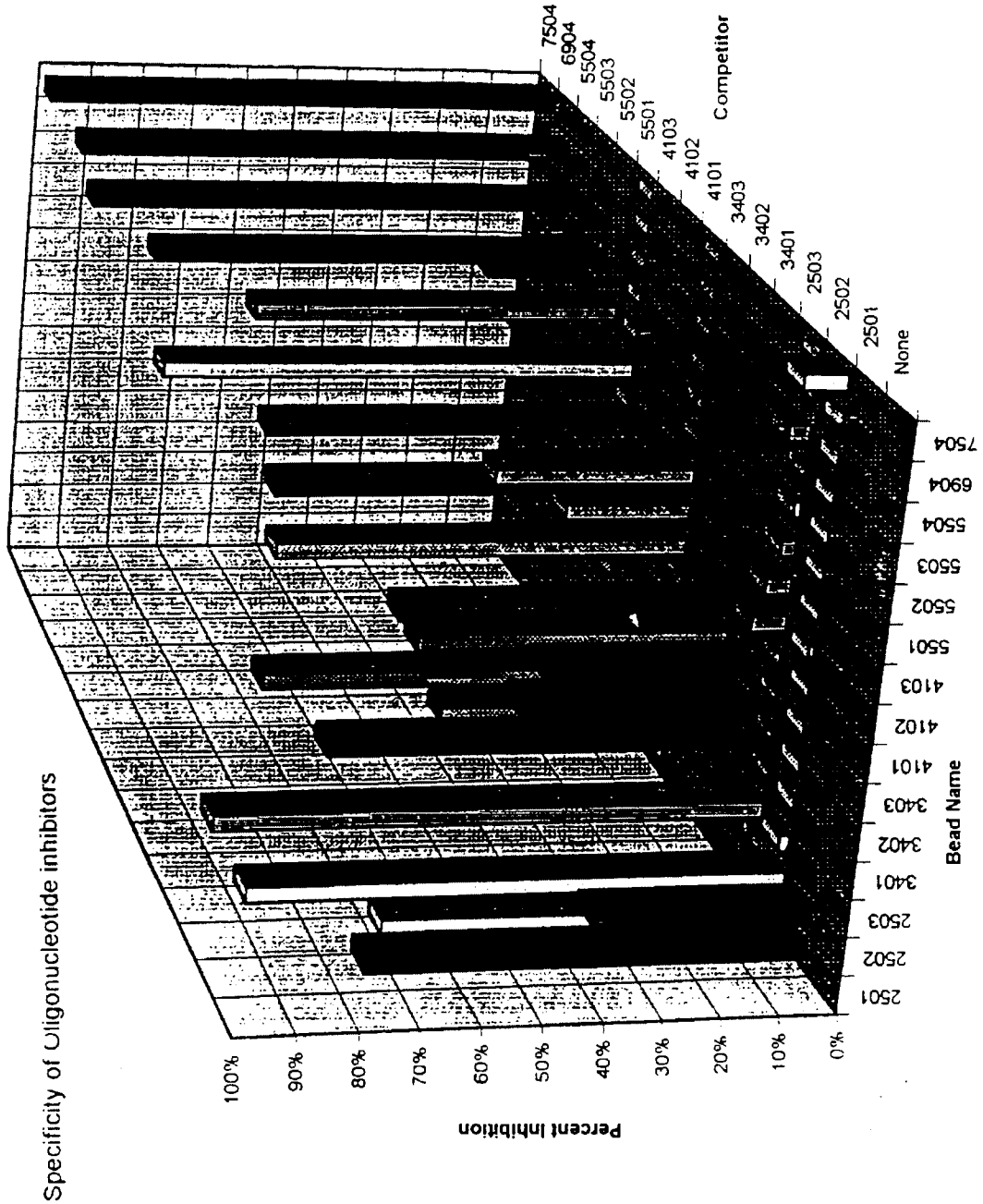

Concentration Dependence of Inhibition with Oligonucleotides
Percent Inhibition
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
-10.00
0
10
20
30
40
50
60
70
80
Competitor [nM]

Fig. 48


Specificity of Oligonucleotide inhibitors
Percent Inhibition
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Competitor
7504
6904
5504
5503
5502
5501
4103
4102
4101
3403
3402
3401
2503
2502
2501
None
Bead Name
2501
2502
2503
3401
3402
3403
4101
4102
4103
5501
5502
5503
5504
6904
7504

Typing of Mock DQA1 Alleles

Fig. 49

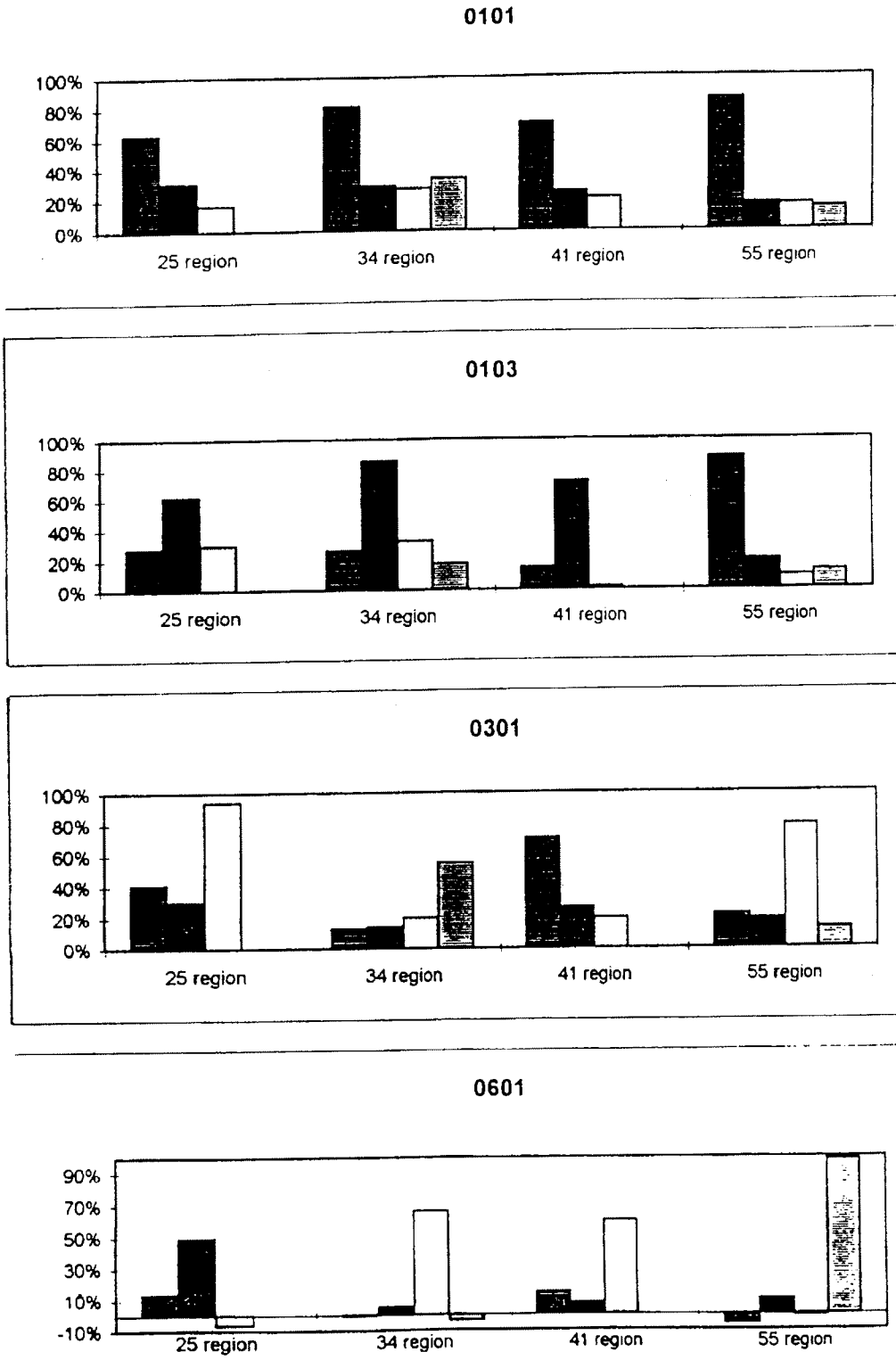


Fig. 50

Typing of DQA1 PCR Products

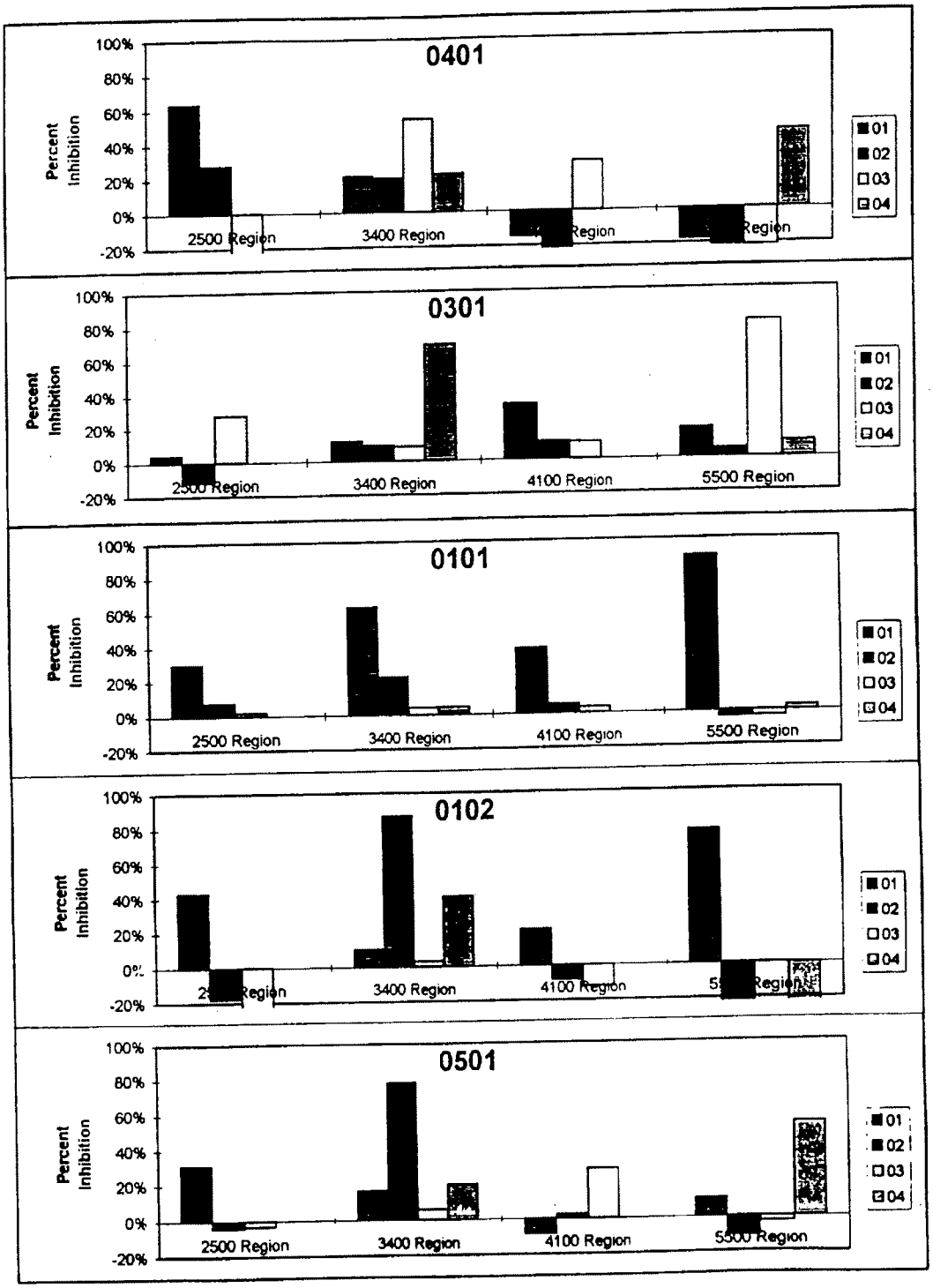


Fig. 51c

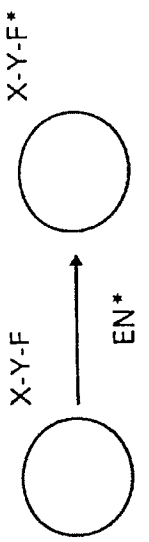


Fig. 51d

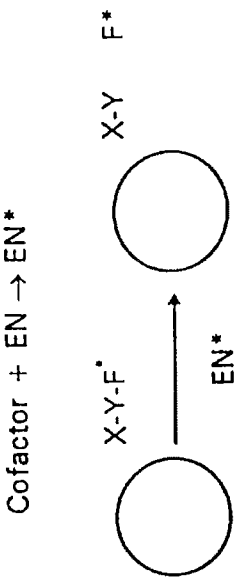


Fig. 51a

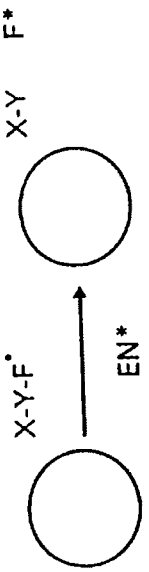


Fig. 51b

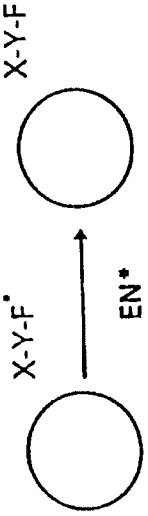


Fig. 51f

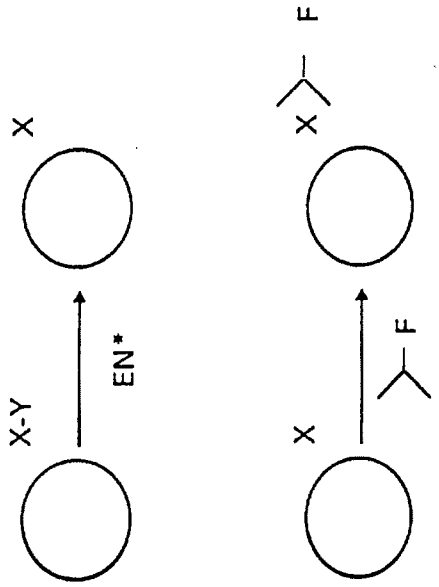
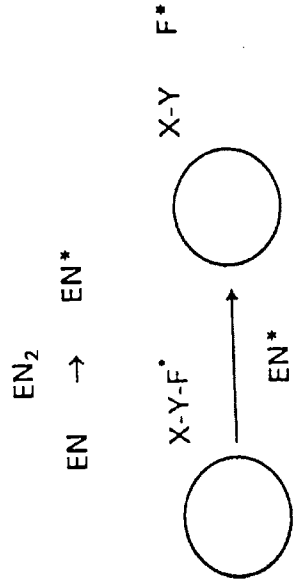


Fig. 51e



PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, G01N 27/26, C07H 21/00	A1	(11) International Publication Number: **WO 97/07245** (43) International Publication Date: 27 February 1997 (27.02.97)

(21) International Application Number: PCT/US96/13313

(22) International Filing Date: 14 August 1996 (14.08.96)

(30) Priority Data:
08/514,820 14 August 1995 (14.08.95) US

(71)(72) Applicant and Inventor: RABANI, Ely, Michael [US/US]; 4495 Vision Drive #1, San Diego, CA 92121-1942 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS AND DEVICES FOR PARALLEL MULTIPLEX POLYNUCLEOTIDE SEQUENCING

(57) Abstract

The present invention is directed to methods of multiplexing sequencing reactions such as the Sanger method when used with appropriately prepared complex samples. The method is advantageous because it allows the sequencing of many polynucleotides simultaneously. The method is primarily based on the Church and Gilbert multiplex sequencing reactions, in that the polynucleotides to be sequenced are constructed to contain terminal base-specific labeling. Furthermore, the method obviates the need to cross-link reaction products to membranes with chemical reagents. Method steps comprise, parallel separation of sequencing templates or sequencing reaction products from mixture prior to gel electrophoresis, and parallel detection of different samples or reaction products. To achieve separation, plural samples are applied to specific separation channels by parallel sample-specific binding means to said specific channels, thereby combining both multiplexation and parallelism. Detection is achieved by detector means that comprises tag-probe arrays. Suitable devices for performing the method are also described.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

"METHODS AND DEVICES FOR PARALLEL MULTIPLEX POLYNUCLEOTIDE SEQUENCING".

Field of the Invention:

The invention relates to the field of molecular biology, and more specifically the determination of polynucleotide sequence composition.

Related Art:

Due to the importance of the structure of genetic material to the understanding of biological mechanisms and genetic engineering, much effort has been focused on the development of techniques for the determination of the nucleotide base sequence of polynucleic acids, especially of DNA. At present there are only two well established basic methods for the determination of polynucleotide sequence. These are Maxam and Gilbert base specific chemical cleavage[1] and Sanger enzymatic chain termination[2]. Both of these methods serve to generate, from a sample of identical polynucleotide molecules, sub-populations of polynucleotide molecules with a particular one of the four naturally occurring nucleoside bases, according to treatment of the sample aliquot, at the 3' terminal position. With both of these methods, each such differently terminated sample aliquot comprises a so-called "nested set" of populations of molecules of different length, up to the length of the input sample molecules, but identically having the specified 3' terminal base composition. Thus, the length of each subpopulation within each such sample aliquot corresponds to a position of the base moiety, for which that sable aliquot has been treated, within the sample sequence. The nested

sets produced by both of these methods are conventionally separated according to length by electrophoresis through gel matrices, generally of polyacrylamide composition. Generally, the sample is labeled with radiolabels or dye moieties such that appropriate detection techniques may localize the position of the bands which are produced by electrophoresis of different length molecules in the gel. By comparison of the pattern generated by such an electrophoresis step, the base sequence of the original sample is inferred. These methods are well known within the field of molecular biology.[3] To improve the rate at which sequence information is accumulated, many different modifications of the original methods described above have been attempted or implemented. These include substituting different resolution media for polyacrylamide gels in the separation step[4], temporal or temporal and spatial separation of fragments involving their departure from the separatory gel matrix (e.g. onto moving membranes or passed detectors), automation of fluid handling, reaction and electrophoresis steps with either conventional slab gels or gels confined to capillary tubes. Despite the improvements in efficiency attained by such modifications, established sequencing technologies are a rate limiting step in much molecular biological research and the several genome projects, including the Human Genome Project. In part, such limitations stem from the characteristics of electrophoretic separation, specifically the time involved and the decreasing resolution power for successively longer fragment length, which decreases approximately logarithmically. It is thus not generally feasible to sequence stretches of longer than 500 bases in length by gel electrophoresis, and much effort and care is required to extend this length beyond 1,000 bases at present.

Multiplex Sequencing:

One innovation which attempts to negotiate the limitations associated with gel electrophoresis is multiplex sequencing.[5] This method has been taught by G.M. Church in U.S. Patent Number 4,942,124 and further by G.M. Church and S. Kieffer-Higgins in U.S. Patent Number 5,149,625. These workers have noted that multiple samples may be separated within the same gel region (lane). When these are uniquely labeled with a tag sequence or otherwise uniquely detectable tag, the length-fractionated sequencing reaction products are then

serially and uniquely detected according to tag identity. Thus, the pattern resulting from the electrophoretic fractionation of the sequencing products from each sample species is accessed by probing for the respective tag. Generally, this method is implemented by transferring or blotting the gel after electrophoresis onto a nitrocellulose or nylon membrane such that the size-separated polynucleotide molecules are deposited onto said membrane with preservation of spatial configuration of the resultant banding pattern. Usually a further step, such as UV illumination of said membrane, is performed to further strengthen the immobilization of said polynucleotide molecules. Probing is generally performed with a radiolabeled, dye labeled or enzyme linked oligonucleotide complementary to the respective tag sequence. After image capture of the banding pattern thus revealed, the oligo probe is denatured from its cognate tag sequence and washed away from the membrane. The membrane is then serially reprobed in the same way until many or all distinctly tagged samples have been detected. Data from each probed sample is used to reconstruct sequence information in the same manner as with Maxam and Gilbert or Sanger sequencing methods. Beyond the reduction of electrophoresis time and effort per sample by this method, distinction with unique tags permits several samples to be pooled together, which enables several samples to be processed in only the four reactions specific to each base, significantly reducing handling steps. A drawback of this method has been the difficulty of automating the serial probing of membranes, though efforts have been made in this regard.[6,7,8]

Electrophoresis methods have been described in a volume entitled *Gel Electrophoresis of Nucleic Acids: A Practical Approach*.[9] These and other methods are taught by J. Sambrook, et al.[10]

Manifold Sample Transfer:

So-called manifold sequencing, a method developed by A. Lagerkvist et al.,[11] reduces manual manipulations in the sequencing of many samples by using modified gel combs or similar structures for the transfer of several samples from sample wells of, for example, microtitre plates, wherein reactions such as PCR and enzymatic sequencing are performed, as well as to a slab-gel. Such a method permits handling of more samples than parallel pipettors commercially

available with greater simplicity and economy. This is accomplished through the immobilization of sample molecules to the solid phase of the comb teeth, similarly by a method comprising the step of dipping these into wells initially containing sample material. This multiplies the number of samples which may be separately handled in a given manipulation step by the number of comb-teeth used, and has the further advantage of reducing handling errors, an issue critical in clinical applications of sequencing methods.

Parallel Capillary Electrophoresis:

E.S. Yeung and J.A. Taylor have taught, in U.S. Patent Number 5,324,401, a method for the facile simultaneous use of plural capillary tubes for either capillary gel electrophoresis or capillary zonal electrophoresis. The parallelization of sample separation possible due to the macroscopically small lateral size of these capillaries, as well as the more rapid separations possible by the use of such capillaries and other advantages over conventional slab-gel electrophoresis, facilitate an increase in the number of separations that can be preformed in a given time with given separation effort. A drawback of this method is the lack of correspondingly rapid sample loading methods and means, which limits the ability to take advantage of the parallelism thus enabled.[12]

Nanochannel Array Glass:

R.J. Tonucci et al.[13] have demonstrated a method whereby two or more types of glass, at least one of which is acid-etchable and provided initially in the form of a rod, and one of which is inert, are heated, drawn and fused together repetitively, followed by an annealing step. The number of times such a set of steps is repeated, and the number of contiguous lengths fused together in each step, determine the number of channels ultimately formed, as well as the size of the resulting channels. Channels are formed in a final acid etch step with regularity across the entire medium. These workers note that the definition of the channels formed after etching results from interdiffusion of the different glass types during drawing and annealing, Though these authors do not note this, decreases in such definition entail increases in internal surface roughness of the capillaries thus formed. Note that such articles with channels of

larger than nanometer dimension may be termed a channel array glass most generally, and with channels of micron-range dimension may be termed microchannel array glass.

Microfabricated Separatory Devices:

In recent years there has been increasing interest and effort directed at the microfabrication of analytical instruments and separatory devices. A.T. Woolley and R.A. Mathies[14] have recently described microfabricated capillaries integrated on the surface of a solid substrate. These are formed by patterning and then etching plural trenches in a glass substrate according to photolithographic methods, and then sealing these to form a capillary array by fusion of said substrate with a juxtaposed glass plate. Such microfabricated capillary arrays are then filled, by injection, with a separation medium, for which these workers choose hydroxyethyl cellulose. Using devices with 30-120 micron capillary size, these workers demonstrated the separation of the DNA fragments resulting from HaeIII restriction digestion of phiX174 replicative form DNA in 120 seconds. Capillaries of similar internal dimension but fabricated by conventional means have been used by these workers to separate sequence reaction products[15,16] rapidly over sufficiently broad size ranges with single base resolution, indicating that sequence determination will be feasible with such microfabricated capillary arrays.

Replicated Relief Structures and Microcontact Printing Therewith:

This section reviews related art methods for fabricating structures and decorated articles useful as means in various aspects of the present invention.

Microfabrication with Replication:

Integrated optical devices and other microscale optical components have been fabricated by the replication of relief patterns into polymeric materials[17]. The optical properties of these materials and the structures resulting from pattern replication determine device function and characteristics. Replication methods used include injection molding and casting into polymeric, elastomeric or metallic

molds or mold inserts, and hot embossing with reliefs, for example, metal reliefs. Multiple cycles of replication may be performed to yield a large number of reliefs for the rapid mass production of devices, according to the fidelity of replication and acceptable device tolerances.

Over the last decade, methods have been developed to form three dimensional patterns on the micron and sub-micron scales, combining lithography with electroforming, micromolding and mold replication. An original microrelief is formed by the lithographic depth patterning of a relief material, which is then developed to form the desired predetermined structure (first relief). Electroforming is then performed to yield a metal negative mold insert conforming to the polymer resist structure, and then separated from the original resist. This first generation mold may then be as a mold for the casting or microinjection molding of a second polymeric relief nearly identical to the first relief. The molding or casting process may be repeated with the first generation mold, and the electroforming process may be repeated with second or subsequent generation polymeric relief structures produced with said molds by appropriate methods. Thus, polymeric objects and metallic objects with sub-micron patterns may be economically reproduced. This set of methods, known as LIGA, was developed by E.W. Becker et al.,.[18] These workers reported reproduction of lateral features smaller than 0.1 micron. These methods are particularly useful in the production of devices with microscale features and moving parts. Where the lithographic method employed is synchrotron X-Ray lithography, high aspect ratios may be achieved with resists of up to millimeter depth. More recently, M. Abraham et al.[19] have reported the fabrication of micro-optical systems with LIGA, and further extend these methods by using mold inserts thus produced for hot embossing.

Similarly, M.T. Gale et al.,[20] extend methods from the embossed diffractive foil and compact disk (CD) industries by the replication of an original (e.g. microfabricated) microrelief on the surface of a replication shim through electroforming, followed by replication of this first shim by surface passivation followed again by electroforming. These workers find that the shim replication process occurs with only slight (<2nm) increases in surface roughness per generation, thus enabling the reproduction of nanoscale relief

features, with corresponding cost advantages. Related injection molding methods are described by A. Neyer et al.,[21] and R. Klein and A. Neyer.[22]

Contact Printing of Resist by Lithographic Plate:

In U.S. Patent Number 5,380,620, T. Suzuki and F. Shinozaki teach a method whereby a lithographic plate comprising regions of ink binding and ink repelling regions is contacted with an ink or resist sheet and then contacted with a substrate to which said ink or resist which bound to said ink binding regions of said ink. This method forms lithographic plates by exposing materials similar or identical to those used as resists in conventional microfabrication, such that exposed regions have affinity or repel ink or other liquids, such that a pattern of differential retention of ink or said other liquids may be used to form a pattern of said ink or other liquids on said lithographic plates and then transfer this pattern to the substrate.

Photolithography Generated Relief Patterns, Mechanical Transfer Thereof, and Use as Lithographic Plate:

Methods for the production and transfer of relief patterns in polymeric resist materials or variants thereof are reviewed by B. Bednar, J. Kralicek and J. Zachoval.[23] A relief is produced by selectively exposing regions of a photoresist coated onto a first surface and mechanically transferring either the exposed or non-exposed regions to a polymeric foil second surface by juxtaposing said second surface to the image-exposed polymeric resist and relying on the differential adhesion properties resulting from exposure. Thus, a positive image is produced on one surface while the other surface retains the corresponding negative pattern. Such relief patterns may provide differential wetting or liquid-retention properties differing from those of the underlying surfaces, and thus be used as offset lithographic plates.

Microcontact Printing with Elastomeric Microreliefs:

G.M. Whitesides and co-workers have done considerable work developing methods for the use of replicated elastomeric microreliefs for microcontact printing.[24] It is presently important to note that

this work demonstrates the conformation of the raised portions of an elastomeric microrelief surface to the shape of another surface to which it is juxtaposed and impressed. These workers have similarly used such reliefs for micromolding, which further indicates that conformed contacts resulting from impression of such microreliefs onto a juxtaposed surface forms a liquid-tight seal.

Chemical Manipulation and Synthesis with Scanning Probe Microscopes:

W.T. Müller, P.G. Schultz et al.[25] describe a patterning process whereby a surface is adsorbed with molecules comprising a chemical functional group which may be transformed to a second chemical functionality by contact with a metal catalyst, an AFM is used to scan a tip coated with said metal catalyst over areas of said surface which are intended to be transformed, and said chemical functional groups transformed by the resulting, spatially limited catalytic process are then subjected to further chemical reactions which are selective for the catalytic reaction product second chemical functional groups. In particular, these workers situate azide (alkylazide) functionalities on a substrate surface, and contact a platinum coated AFM tip to these in the presence of dissolved H_2, to effect the catalytic hydrogenation of those of said nitriles contacted by said AFM tip under positional control effected by an AFM. The nitrile groups thus contacted are reduced to amines, which these workers detect by reaction to aldehyde modified fluorescently labeled beads, which are observed to bind only to regions which underwent said contact with said platinum. B.J. McIntyre, M. Salmeron and G.A. Somorjai[26] have similarly effected positional control of metallic surface catalysis using a platinum-rhodium STM tip in an atmospheric-pressure chemical reactor to catalyze the rehydrogenation of molecules comprising alkyne functionalities on (111) platinum surfaces.

Alternative Sequencing Methods:

Various novel methods for the determination of polynucleotide sequence have been proposed. These include substitution of mass spectroscopy[27] (including time-of-flight mass spectroscopy) separation for gel electrophoretic separation, exonucleolytic degradation with

nucleotide transport and single molecule detection of the effluent[28], scanning probe microscopic visualization of polynucleotides at high resolution[29,30] with the aim of base discrimination within a single molecule, and various sequencing by hybridization methodologies. Of these, only the last appears at present to have significant practical potential, but significant obstacles and challenges remain.

Sequencing by Hybridization:

Two general variants of sequencing by hybridization methodology have been proposed. In the first, a sample under investigation is hybridized to an array of oligonucleotides situated on a surface, fabricated, for example, by the spatially controlled light directed synthesis methods of S.P.A. Fodor et al.,[31], where the composition of an array element is related to the position of said element within said array of oligonucleotides. For example, an array may comprise all possible 10-mer oligos, in which case there are over 10^6 array elements. An unknown sample is denatured and permitted to anneal with said array under stringent conditions. Said array is examined for the presence of bound sample molecules, generally by optical means. Ideally, the sequence composition of a sufficiently short sample molecule may be reconstructed by using information about which array elements have bound sample molecules and about the compositions of each of these array elements, with algorithms which overlap each of the short sequences (corresponding to the 10-mer oligos in this example) to produce a linear sequence. Two difficulties arise with this method. First, because DNA includes tandem and dispersed repeat elements (including transposons), a particular 10-mer, which is one of the more than 10^6 possible 10-mers may occur many times in naturally occurring sequences much shorter than a megabase.[32] Thus, this method will lead to branch points which ambiguate the sequence data thus obtained. Longer oligo probe libraries or arrays will contain more elements, which thus occur less frequently, but require more surface area and synthesis steps, and will further diminish the ability to discriminate between perfect and mismatched hybridization. Second, even for shorter oligos, such as 7-mers, hybridization is not perfect for all sequences under any given conditions, as is clear to those familiar with other probe hybridization techniques such as northern blot gel analysis; optimal conditions for a particular probe

to form perfect hybrids and not form imperfect or mismatched hybrids vary in a sequence dependent manner, and it is not always simple to predict this optimum even where only a single probe is concerned. Thus, for any given array of all possible n-mers, only a subset of elements will display optimal binding discrimination characteristics under any set of particular conditions. Further, because of the formation of imperfect mismatches, discrimination between strongly and weekly binding probes must be considered rather than simply the success or failure of an array element to bind sample molecules.[33] Various attempts have been made to address these issues, with some success[34], but the technique is not yet competitive with established methods.

In the second variant of sequencing by hybridization methodology, the sample sequence (which is generally in the form of a phage or plasmid library) is immobilized, generally in the form of bacteriophage plaque or bacterial colony replicas, on a membrane, which is serially probed with a different labeled oligonucleotide of known sequence and length n. Here, because only one kind of hybrid is formed during each step, conditions may be closely optimized for homoduplex formation in each probing step. Plaques or colonies forming hybrids are noted. Repetition over the full combinatorial library of oligonucleotides of sufficient complexity relative to the complexity of the sample thus provides information about all of the n-mer sequences present in a particular plaque or colony. After each element in the full combinatorial library has been used to probe the sample sequence, linear sequence information may be reconstructed as in the first variant of sequencing by hybridization methodology. The two main drawbacks of this method are the requirement for synthesis of large probe libraries, which is much more laborious than the production of surface immobilized arrays by optically patterned deprotection[35], and that significant linear sequence information may only be reassembled only after many (thousands or more) probing steps.

The first variant of sequencing by hybridization is, however, proposed as a confirmatory sequencing method to identify any errors in sequence information obtained by other methods, and the same methodology is proposed as a means of classifying polynucleotides in unknown samples, for example, to detect the presence and type of

polynucleotides present in clinical isolates. In this latter application, even where a particular polynucleotide displays a hybridization pattern which permits only ambiguous sequence reconstruction, a characteristic pattern which may serve as a highly specific signature may nonetheless be observed.

Parallel Multiplex Polynucleotide Sequencing:

In U.S. Patent Application 08/445,094 filed on 19 May 1995, E.M. Rabani teaches a method for the sequencing of polynucleotides, especially DNA, combining the separation throughput advantages gained in multiplex methods with the detection methods made possible by oligonucleotide arrays. In the most general aspect of that invention, separation is performed so as to effect length dependent fractionation of complex mixtures of tag-primer primed sequencing reaction products, which said fractions are separately scored for the presence of tags, or for labeling moieties in communication with tags, on oligonucleotide arrays which comprise tag-probe sequences (i.e. such arrays may comprise as many distinct tag-probe oligonucleotides as there are distinct elements in such arrays, which may exceed 10^6 or 10^8 different oligonucleotide elements per cm^2, depending on the fabrication methods used to produce these.) In a preferred embodiment, said tag-primer primed sequencing reaction products are affinity labeled during that reaction with one of four distinct labeling moieties the identity of which corresponds to the identity of the terminal base moiety. After fraction collection, hybridization to an array comprising said probe-tags, and washing steps, four differently colored fluorescent beads each distinctly derivatized with one of four corresponding receptor types, which bind one of said distinct labeling moieties, are contacted to said array, with the result that array elements to which said tag-primer primed sequencing reaction products have bound are labeled, in this instance by a fluorescent bead. The color or identity of the bead bound to a particular array element or tag-probe reveals the identity of the terminal base moiety of the sequencing reaction product fragment thereupon, and hence the identity of the base at the position within the sequence of the tag-primer primed sequencing fragment probed by said tag probe (corresponding to said particular array element) corresponding to the position of said fraction within the series of

fractions collected (in a manner precluding the presence of successive base terminated species of the same tag-primer primed sequencing reaction product in the same fraction) and more particularly within the ordered set of labeling data obtained from similar scoring of all successive fractions. With this method and the several variations disclosed therein, multiplexation advantages in sample preparation, sequencing reaction and separation are combined with the rate advantages of parallel detection, such as is possible with said large (i.e. complex) probe-tag arrays and the technical facility of photolabel detection (e.g. fluorescence detection, especially of fluorescent beads, with appropriate CCD arrays with data capture by an associated digital computer and storage device.) These advantages come primarily at the cost of the requirement to examine a large number of fractions (greater than or equal in number to the length in bases of the linear sequence to be obtained for each or most tag-primer primed samples.) At the array hybridization step, fractions may be applied successively to a single array, which may be referred to as serial fraction array tag-probing, or multiple fractions may be applied (on a one to one basis, preserving information regarding ordering) to plural said tag-probe arrays, which may be referred to as parallel fraction array tag-probing. Additionally, data regarding labeling of array elements (in association with ordering according to fraction) may be collected from such tag-probe arrays sequentially by a single suitable photodetection means or in parallel by multiple such photodetection means. CCD based detection is, nonetheless inherently parallel, so this latter distinction refers to hierarchy of detection parallelism. Thus, despite the great advantages in overall sequencing rate attainable by such methods over conventional and other alternative methods, the sequence data collection rate with this method remains proportional to the complexity of the probe arrays used (i.e. the number of distinct probes per array), but the effort or number of automated steps required is primarily determined by the number of fractions thus examined. It would be desirable to extend this method using appropriate means so as to reduce or eliminate the need for or the number of separate fractions which must be separately handled (whether manually or by automated means) without thus reducing the rate at which data are collected or the economy of sequence

determination by such a method. It may for some purposes be further desirable to rely on probe arrays of generally smaller diversity (i.e. consisting of fewer array elements), since the effort required to fabricate these is primarily proportional to the length of the variable portions of the oligonucleotides therein, and further since more methods are capable of efficiently producing lower density arrays.

Analysis of Conventional Methods:

Logical Steps of Nested Fragment Set-Based Sequencing:

This discussion delineates the essence of conventional sequencing methods in a manner which will facilitate the understanding of the present invention both in its full generality and the large number of variations comprehended by the present invention. In both enzymatic and chemical methods, samples are in a first step prepared as suitable reactant molecules (e.g. as single stranded and primed molecules for enzymatic methods or as end-labeled fragments for chemical methods). In a second step, these samples are subjected to desired sets of reactions or combinations thereof to form fragments which are either separate nested sets each comprising only fragments with a single terminal base moiety or fragments differently labeled according to the identity of the terminal base moiety which said fragments differently labeled comprise. In a third step, these samples are each applied to a different separation channel. In a fourth step, samples are separated into the constituent molecular species which the nested set comprises according to length or mass separation (e.g. electrophoretic mobility or transport). In a fifth step the distinct said constituent molecular species in each separation channel are detected (either according to the final position reached in said separation channel at the time said fourth step was concluded by spatial detection method, the coordinate of which corresponds to the order of molecular species along the length of said sequencing channel, or according to the time or time delay at which a fragment is transported passed a detector situated at a fixed position along the length of a separation channel, concurrently with the separation step, in which case the separatory coordinate corresponds to the temporal ordering of transport dependent events;

in short, these are spatial and temporal separation coordinate resolution methods, respectively.) As more broadly considered, a sixth step consists in the assembly of sequence information obtained in the preceding steps into longer stretches of sequence data where the sequences of distinct samples presents both sufficient identity and unambiguous or unique match at the putative sites of overlap. In summary, these steps are:

1. sample preparation, which may include creation of libraries or plasmids, growth, purification and/or amplification steps;

2. base specific fragment generation (sequencing reaction), which may be effected by enzymatic or chemical methods;

3. application to separation channel;

4. separation, which may avail slab gel electrophoresis, capillary tube electrophoresis, microfabricated channel electrophoresis, or mass spectroscopic separation;

5. detection of separated species; and,

6. reconstruction of sequence information from separation pattern information.

Automated implementation of some or all of these steps have increased efficiency and reduced labor requirements and costs per base of sequence information, generally only modestly departing from the above basic scheme. Capillary array electrophoresis yields gains in parallelism at step 4 and thereafter. The original multiplex sequencing protocol of Church and Gilbert, and known art variations thereupon may gain advantages in reduction of sample handling and separation starting at step 1 above, gain these advantages at the cost of repetitions of serial reprobing steps in step 5. Multiplex PCR may reduce the number of samples which must distinctly be handled in step 1. Manifold transfers may reduce the number of manipulations as the results of step 1 are subjected to step 2, but still require the same number of vessels and separatory channels. The various encoding schemes (described below) which have been used to permit the sequencing products of two to four sequencing reactions yielding fragments terminating at a predetermined base moieties in predetermined proportion and/or with distinct, base-specific or reaction specific labels to be separated in the same separation channel, up to four, while not confounding data collection with

appropriate detection means, have provided two- to four-fold more efficient utilization of however many separation channels are thus used, primarily gaining advantage in step 4, but also in step 5. CCD arrays, photon cameras or other suitable electronic devices for image capture having increased resolution may improve the detection rates in protocols using optical labels including fluorescent labels. Improved image recognition algorithms and faster computers (including parallel computers) may increase rate of the final step of reconstruction of data from individual samples to ordered data about the original sample sequence.

Variations providing sequencing rate, economy advantages or reduction of manual steps:

Each of the aforementioned seven steps, as well as other associated steps and methods, admits either multiplexing or parallelization or both, with different advantages and tradeoffs according to the sequencing task intended. In conventional practice, slab gel sequencing admits the small-scale parallelism of multiple separation channels (i.e. lanes). As will be clear from the foregoing description, the multiplex sequencing methods of Church and Gilbert may avail multiplexing in sample preparation and reaction and separation, or reaction and separation, or only during separation; detection is performed serially, i.e. generally one sample per lane per detection step, where serial reprobing is made possible by the association of a unique tag sequence, generally as part of the priming molecule, though in the original method part of the sample sequence, by a labeled probe having a complementary sequence to said tag sequence, which such labeled probe will be referred to herein as a tag-probe. At the levels of separation detection, methods such as the encoding of terminator concentration or relative activity with corresponding effects on band intensity (e.g. a binary encoding scheme such as 2A+G and G+2C reactions of the same primer-template end or length labeled identically and run in adjacent lane such that 0,0=T; 2,0=A; 0,2=C; 1,1=G) permit two-fold advantages in the form of separation multiplexation (though here in the limited sense attaching to fragments derived from a single primer-template complex or the like) and corresponding advantages in fragment detection rate with sufficient intensity discrimination. More recently, the use of dye

molecule labels or reporters, which may be in communication with primers used (end-labeling), or with deoxynucleotides (length-labeling) permit the separation of the combined mixture of products of four separate sequencing reactions (each corresponding to a different chain terminator moiety and having been labeled with a dye moiety having a distinct absorption and/or emission spectrum) in a single lane, permitting further improvements in this limited sort of multiplexation of separation and detection. The extent of such multiplexation based on distinct labeling is limited by the number of suitable spectrally distinct dyes available. This category of base-moiety multiplexation limited to an individual primer-template sample has, however, been extended, through the use of distinct fluorescent or other optical labels in communication with each of the four chain terminator analog (generally, 2'-3'-dideoxy nucleotide triphosphates) such that the same synthesis event which terminates chain elongation also places a label corresponding to the identity of the incorporated chain terminator in communication with the resulting terminated elongation fragment. This further improvement results in a four-fold reduction of the number of reactions which must be performed to sequence a template sample. Even in this last case, such methods, by themselves, only provide four-fold advantages through the steps of sequencing reactions, separation steps and detection, providing only four-fold advantage at the system throughput level.

Object of the Invention:

It is an object of the present invention to gain the advantage of parallelism in the detection steps of multiplex sequencing and thus better exploit the parallelism of sample handling gained in sample preparation and size fractionation steps with multiplex sequencing.

It is a further object of the present invention to provide highly parallel base position detection without extensive use of high-complexity oligonucleotide arrays, for example, those comprising over 10^6 elements each, so as to provide the possibility of a comparable rate of detection with a larger number of oligonucleotide arrays each consisting of fewer elements.

It is a further object of the present invention to provide for the multiplexation of sample loading in parallel separation methods including parallel capillary electrophoresis methods.

It is a further object of the present invention to combine the parallelism advantages of parallel capillary electrophoresis with the sample processing advantages of multiplex separation.

It is a further object of the present invention to provide means to combine the advantages of parallel capillary electrophoresis and multiplex separation with the advantages of parallel detection. In particular, it is a further object of the present invention to minimize or eliminate the need to perform multiple manipulations to form separate liquid fractions for analysis on separate probe arrays while still gaining the parallelism advantages of detection comprising the use of probe array scoring of plural distinct species occurring within a size fraction.

It is a further object of the present invention to provide rapid and inexpensive methods and means by which complex polynucleotide samples including large portions of genomes may be analyzed with single base resolution.

It is a yet further object of the present invention to provide rapid and inexpensive methods and means for the high resolution analysis of many independent samples of genetic material, for example, one or more particular loci from many individuals within a population of organisms.

Methodologies and means fulfilling one or more of these objectives would increase the rate of progress of much biological research and applications including recombinant DNA technology, and have further relevance to both basic medical research and clinical medicine.

Summary of the Invention:

The essence of the present invention is to combine any of the various multiplexing methods possible at each necessary step of Sanger enzymatic sequencing protocols or Maxam and Gilbert chemical sequencing protocols for use with appropriately prepared complex samples, with the advantageous degrees of parallelism at appropriate steps of sequencing protocols. Prominent multiplex methods among these are multiplexing of samples in reactions directed at generating terminal base specific nested fragment sets or terminal base specific reporter-linked fragment sets, labeling and detection, as well as use of variations on aspects of the multiplex sequencing methods of Church and Gilbert obviating the crosslinking of sample sequencing

reaction products to membranes with chemical reagents. Prominent parallelization strategies enabled include parallel separation of sequencing templates or sequencing reaction products from mixtures prior to molecular mass fractionation (e.g. slab or capillary gel electrophoresis, capillary zonal electrophoresis, mass spectroscopy), parallel separation steps through the use of plural separation channels or means, and parallel detection of different samples or reaction products generated from different samples within the same mixture or mixture fraction. At the level of separation, single or plural sample mixtures or sample reaction-product mixtures may be prepared in multiplex, may each be separated in a manner directing them to specific separation channels or other separation means by parallel sample-specific binding means and thereafter separated by plural said specific separation channels or other separation means, combining both multiplexation and parallelism. At the level of detection, separation channels may comprise or suitably articulated with sample recovery means in turn comprising parallel sample separation means, may be detected by parallel detector means (e.g. tag-probe oligonucleotide arrays and photodetector arrays such as CCDs for visible labels) and alternatively or further discriminated by signal amplitude encoding schemes where appropriate with detection methods used. Suitable devices and apparata and variations thereupon facilitating such methods are further described.

Description of the Invention:

Definitions:

As used herein, the term parallelism will generally indicate concurrent performance of similar operations on different populations or sub-populations of molecules or samples across space. Parallelism thus entails the multiplication of devices or articles with the methods of the present invention to gain a corresponding multiplication in throughput.

In contrast, as used herein, the term multiplex will generally indicate the concurrent manipulation of different samples (or molecular products yielded by reactions performed collectively upon the ensemble of said different samples) within a continuous region of space, area or fluid volume. Multiplexing thus involves the

reduction of handling steps, and the ratio of devices or articles used to accomplish a multiplexed step per sample, by treating different samples identically, which is possible where each of said different samples comprise unique features facilitating subsequent separation and/or discrimination according to the origins of molecules (or of molecular products yielded by reactions performed collectively upon the ensemble of said different samples) of from each of said different samples. This is stressed because within much of the related art the term "multiplex" has been applied to arrangements or methods which are more accurately described as parallel. Note that the present usage is consistent with that employed by Church and Gilbert.

Thus, it is important to note that in some regards, parallelism may be exploited as a method of de-multiplexing differently derived fractions of sample mixtures or the like as well as demultiplexing the information obtained during detection steps.

These distinctions will facilitate a clearer understanding of the essence of the present invention as it is set forth below and as it may be variously embodied by those of ordinary skill in the relevant fields of related art (which include molecular biology, recombinant DNA technology, microfabrication, integrated optics replication, combinatorial chemistry, systems operations theory, electronics and electronic sensing, and, computer architecture engineering.)

Further note that the term separation channel is used generically to include such diverse separations means as a mass spectrometer, a slab gel lane, a capillary gel, a capillary or the like. By contrast, the terms formed channel, formed relief channel and articulated channel, all of which may serve as separation channels and may be of varied form and composition, do not apply to structures formed from conventional capillary tubes. Lateral closure of such channels need not be permanent, and thus separated species may be recovered by laterally accessing said thus separated species by reversing said lateral closure of such channels, i.e. separating one or more portions of materials making up said channels so as to expose a substantial proportion of the internal volume of said channels. Conventional capillary tubes, with which conventional capillary gel electrophoresis and capillary zonal electrophoresis are accomplished,

in contrast have only terminal openings and generally poorly accessible internal surfaces.

Brief Description of the Figures:

Figure 1 depicts demultiplexation array means (10). In Figure 1a, said demultiplexation means are shown in a sealed channel formed by juxtaposition of trench (20), while in Figure 1b, said demultiplexation means are shown each in one among many plural sealed channels formed by juxtaposition of trenches (30) which thus form a separatory channel array. Figure 1c shows an isometric cross-section of one of said separatory channels depicted in Figure 1b.

Figure 2 depicts a bin array (100) comprising a grounded electrode contact (110) across the bottom surface of said bin array (shown with schematic symbol) and shows a portion of a transparent channel array (140) of appropriate pitch offset from said bin array.

Figure 3 depicts a superimposed view of and array (two sets of groves shown) comprising fraction transfer grooves and plural surface immobilized tag-probes (330) of a striped-patterned tag-probe array, and electrodes (350a and 350b) for lateral electroelution, positioned facingly over the trenched surface for forming a separatory channel array, with one such trench shown (320). Note that features 400 and 420 are extensions of the surface comprising said fraction transfer grooves which serve to define two distinct openings to said separation channel trenches when the fraction transfer grove array surface is facingly juxtaposed to the separatory channel trench array surface, and further (420) to prevent cross-transfer of fractions.

General Description of the Invention:

As recited above, established methods for the determination of sequence composition of DNA or other polynucleotide samples generally comprise six categories of steps:

1. sample preparation;
2. base specific fragment generation (sequencing reaction);
3. application to separation channel;
4. separation;
5. detection of separated species;

6. reconstruction of sequence information from separation pattern information.

Heretofore, with only a few exceptions, overall methods for the determination of sequence composition of DNA or other polynucleotide samples have availed multiplex handling or parallel manipulations at only one or a few of the above categories of steps. This is attributable to a general lack of previously identified means for sample demultiplexation, provided in the present invention, or to a lack of parallel detection of plural species (of similar length in nucleotides) separated simultaneously in the same separation channel, such as those recently described by E.M. Rabani in the above cited patent application. Demultiplexing and parallel detection of distinct species of similar length simultaneously separated in the same separation channel individually, and especially in combination, provide the crucial links necessary to exploit combinations of parallelism and multiplexing, in many possible combinations, in order to multiplicatively increase sequence determination rate and improve sequence determination economies, expressed as cost per finished base.

Thus, for example, samples may be prepared (and as necessary, tagged) in multiplex (including multiplex amplification), subjected to sequencing reactions and/or labeling in multiplex, partially or completely demultiplexed with predetermined transfer to parallel separation channels, separated in multiplex in each of said parallel separation channels or merely in parallel with only one sequence per channel, and detected in parallel (including, as applicable, parallel detection of or by parallel detection of samples separated in multiplex, for each separation channel, favorably performed simultaneously and hence in parallel for all separation channels.)

Manifold Loading of Parallel Separation Means:

This aspect of the present invention combines manifold loading of samples with the use of capillary arrays or other parallel separatory channels or separatory channel arrays. Manifolds may be one dimensional combs as previously described by A. Lagerkvist et al.,[36] or may be two dimensional regular arrays of teeth (e.g. rods) projecting from a two dimensional surface, fabricated by suitable

means such as molding, including LIGA techniques or variations thereupon, or other suitable means.

In a second variation, manifold separation may further provide separation of probe targeted species from complex mixtures. Oligo-probes are situated on manifold teeth and used to fish for probed sequences from a complex pool, e.g. several libraries or complex samples of different origin are individually contacted with one comb tooth, such that a comb may carry several separated sample fragments in a manner ordered according to which tooth was contacted with which sample, which is particularly useful for application to parallel separation channels. This second variation is actually a specific case of demultiplexation, and manifolds with plural distinct oligos situated on said teeth may be seen as a specific case of demultiplexation arrays when used in this way, with multiplexed samples. This variation however suffers from shortcomings associated with not integrating probe-tags with separatory channel arrays, particularly that each of said teeth must be contacted with each sample from which it is to capture sample material, as opposed to the unitary contacting provided by the demultiplexation arrays of the present invention.

Random Loading of Parallel Channels With Parallel Amplification:

The present aspect of the present invention involves the use of one- or two- dimensional arrays of sample bins, wells or chambers (hereafter referred to as bins). A single sample volume, comprising a number of (individual) molecules equal to the number of wells used, is distributed among these bins. Alternatively, a sample volume, containing an appropriate concentration of (distinct) sample molecules, and larger than the total volume of all bins of said array, is applied to the array of bins in a manner such that a predetermined portion of said sample volume sufficiently fills each of said bins of said array, such that on average each bin contains one sample molecule; the excess of said sample volume is drawn off of the surface of said array. These aspects of the present invention may be termed Poisson statistical random sample deposition, or simply random isolation. Samples are then (if they had not previously been, or are not otherwise suitable for amplification with standard

primers) ligated to PCR primers, and then PCR or other amplification is performed in each chamber, where thermal cycling of all chambers is effected simultaneously, followed by reactions to generate nested set sequencing products with appropriate labels. Other means of effecting parallel amplification are, of course, similarly comprehended by the present aspect of the present invention. A sufficient number of PCR amplification cycles are performed to yield a quantity of DNA that will ensure that sufficient signal will result from the labels in each sequencing reaction fragment for reliable detection. Thus, samples comprising distinct sample molecules which each comprise substantially different insert sequences or are each derived from substantially different regions of a genome, are generated in an ensemble of bins or chambers (e.g. in a microtitre plate or higher density chamber array); according to Poisson statistics, approximately one third of said chambers contain sequences amplified from only one molecule. The remaining chambers will contain material amplified for either zero or two or more sample molecules. Said array of bins is designed so as to conveniently be juxtaposed to a separatory channel array, or alternatively to be used with manifold transfer means for transfer to a separatory channel array; i.e. said array of bins is arranged in either a mirror image of the configuration of entry ports to said separatory channel array with identical array pitch (for transfer by juxtaposition), or arranged identically to the configuration and pitch of said separatory channel array. In the case of transfer by juxtaposition, physical transfer may take place by capillary action and fluid flow, effected by suction at the exit ports of said separatory array where said separatory array comprises entry ports which may be immersed in said bins, or physical transfer of sample molecules may take place under the action of an applied electrical field. In the case of transfer with manifold, denaturation may free unlinked strands from the teeth of said manifold, as in the method of Lagerkvist *et al.*,.

It should be noted that the approximately two-thirds inefficiency of random bin filling, in absence of other measures to fill unfilled bins with single samples, may be greatly offset in various embodiments by the ability to load samples simply and in parallel into a very large number of parallel separation channels.

Note that the method and means of the present aspect of the present invention may provide further advantages in cases where a sample is in the form of a complex heterogeneous mixture (e.g. a genomic library or a cDNA library comprising thousands to billions of different chimeric clones) having an estimated number of distinct clones per unit volume. Clones may be isolated by such methods availing statistical separation by first diluting an appropriate volume of said complex heterogeneous mixture such that the desired number of molecules are present in a appropriate volume suitable for application to a bin array according to one of the two variations described above. On average, different clones will appear in a bin array in proportion to their relative copy number or representation within said complex heterogeneous mixture. For example, shotgun libraries may be thus sampled and separated into bins each containing one individual sample molecule, amplified, and directly applied to parallel separation channels or means. Further, transformation steps or phage packing steps may be eliminated in certain cases where a library need not be maintained, by randomly fragmenting genomic material, subjecting to treatment with phosphatase, and ligating into vectors comprising appropriate amplification primer sequences, sequencing primer sequences, and optionally, tag sequences, and then separating into bin arrays for amplification. This eliminates host-induced biases arising in library growth. Thus, genomic material may be shot-gun sequenced directly, without even any transformation steps, providing important rate and cost advantages in certain applications. cDNA libraries may similarly be sequenced without transformation or phage packaging and growth steps. Multiplex PCR products of genomic material may similarly be sequenced directly, for example yielding sequence information regarding several regions of each sample genome.

Various assays (e.g. detection of pyrophosphate, a product of polynucleotide polymerization arising in enzymatic sequencing reactions, and most favorably colorimetric pyrophosphate assays) may be performed on all bins subsequent to sequencing reactions and prior to loading to separation channels to determine whether any polymerization took place within each bin, and further assess the degree of reaction which will depend in part on the number of individual molecules randomly distributed into each of said bins.

Thus, empty bins may be readily detected, and with sufficiently quantitative photodetection methods, bins containing material amplified from two or more individual polynucleotide molecules may similarly be detected. Such photodetection of polymerization in said bin arrays is favorably accomplished with photon cameras or CCDs or the like interfaced with one or more digital computer. With an appropriate arrangement and electrical or electronic design of individually controllable electrodes and electrode controlling means, or of individually controllable parallel fluid transfer means (e.g. comprising controllable valves), parallel transfer of samples from each bin of said bin array to the corresponding separatory channel of parallel separation means may be individually controlled. The individual transfer controlling means of such an arrangement may be interfaced to one or more digital computer or the like, which may use data gathered as described above or equivalent data pertaining to the occupancy of each bin to control sample transfer of the contents of each bin of said bin array to said parallel separatory channels or separatory channel array such that only the contents of bins containing exactly one or alternatively fewer than two distinct samples are transferred to said separation means, while data concerning which separation channels have not thus received samples are further maintained. Such a procedure is repeated with additional bin arrays, such that separatory channels not previously receiving samples will receive samples from any of said additional bin array if and only if the juxtaposed or otherwise corresponding bin contains either exactly one or alternatively fewer than two distinct samples. With each repetition of such a transfer cycle with different bin arrays, the probability that any given separatory channel has not received an individual sample decreases by a factor of approximately one-third. Thus, after n transfer cycles, the probability that any separatory channel of said parallel separation means or separatory channel array has received sample material amplified from exactly one molecules is given by $1-(2/3)^n$, for $n \geq 0$, which indicates that after 4 transfer cycles over 80% of said separatory channels will have received sample material amplified from exactly one molecules, which percentage increases to over 91% after 6 transfer cycles. Because application of samples to bin arrays, amplification of material in bin arrays, and transfer of contents of bins to separatory channels

are not involved steps with appropriate apparatus, such repetitions of bin-array transfer cycles may not represent significant costs of time or effort which maximizing the quantity of useful information obtained per unit time from a parallel separatory channel array of a given degree of parallelism.

Loading of Multiplex Samples onto Parallel Separation Channels With Demultiplexation:

The present aspect of the present invention uses one or more immobilized probes or tag-probes (which are generally oligonucleotides, but may, for example be sequence specific binding proteins, etc.) to selectively cause only polynucleotide sample molecules comprising complementary sequences to bind to a selected surface region or spatial region, to facilitate the transfer of only bound polynucleotide molecules to a particular separation channel. The ensemble of distinct said immobilized probes or tag-probes, each of which is chosen and/or synthesized to be substantially distinct from any other probes or probe-tags of the ensemble, are referred to as a demultiplexation array, which term also refers to the surface and object on which they are situated when used most generally. Generally, said probes or tag-probes are provided as a linear array and are situated with a pitch equal to that of the entry ports of a separatory channel array such each element of said demultiplexation array is in, or may be placed in, convenient proximity to exactly one of said entry ports, and such that upon release of said sample molecules from said probes or tag-probes, said sample molecules are either situated within or are readily transported in a targeted manner to the corresponding said entry port. For the case of oligonucleotide based demultiplexation arrays, release of bound sample molecules may, for example, be effected by increasing temperature sufficiently, or exposure to a sufficient concentration of denaturing reagents (e.g. 7M urea).

A complex sample (termed a multiplexed sample) is contacted as a solution with said demultiplexing array such that the sample molecules therein may bind to complementary probes and such that most or all of the probe elements of said demultiplexation array will bind sample molecules of substantially identical origin. For example a flat surface with a linear trench may be juxtaposed with said trench

oriented lengthwise juxtaposedly along said demultiplexation array such that said trench passes over all array elements and forms a channel. A sample mixture containing sequencing reaction product fragments, for example, which are derived from plural distinct sequence regions (i.e. the products of a multiplex sequencing reaction in this example) is flowed through the channel thus formed such that it passes over each of the elements of said demultiplexation array, and non-denaturing conditions of sufficient stringency are maintained. Then a rinse solution is passed through said channel thus formed to carry away and unbound molecules, as non-denaturing conditions of sufficient stringency are maintained. Said array elements of said demultiplexation arrays are designed and produced so as to comprise a sufficient number of probe molecules to capture a sufficient number of sequencing reaction product fragments for reliable detection of each species corresponding to termination at a particular base with the detection method used. Said sample mixture is prepared so as to either not have any two identical tag sequences in communication with different sample sequences to eliminate probing ambiguity, or is produced such that probing ambiguity will be improbable. Alternatively, said probe sequences or other binding means are chosen so as to target only one sequence in said multiplex sample based either on knowledge about the sample or statistical predictions concerning probe sequence occurrence. Said flat surface with a linear trench is then removed from said demultiplexation array, which is then positioned juxtaposedly to the array of entry ports of a separatory channel array, if not part of the same structure. If said demultiplexation array is situated on the same surface of an article comprising separatory channels as the input ports of said channels, then a surface conforming to said demultiplexation array and input port array surface but comprising trenches for juxtaposedly forming channels permitting the transport of molecules from each of said demultiplexation array elements to the corresponding separatory channel input port is then appropriately juxtaposed, and either temperature is briefly increased or denaturants are introduced into said trenches, in either case as the electrophoresis field is applied to initiate separation of the thus demultiplexed sequencing reaction products in parallel.

Distinct tag-probe oligos or internal sequence probes or other such probes of different composition are situated at the entry port to each of two or more, and preferably several, and most preferably thousands or more, separation channels which may be capillary tubes or may be a channel array glass (or microchannel or nanochannel array glass) or other types of separation channels.

In the case of these channel array glass types, probe-tag arrays, which may favorably be of triangular or hexagonal geometry, are either fabricated on the surface of said channel array glass, which surface comprises said entry ports or openings, or are most favorably situated on a surface which is juxtaposed to the surface comprising openings of said channel array glass in a loading step.

Complex mixtures of different sample molecules are first bound to said probe tag arrays or demultiplexation arrays, and unbound species are removed by one or more wash steps. In the case of a probe-tag array which is situated on an article or surface other than that of said channel glass array, sample molecules are bound to said probe tag array which is then washed before juxtaposition or the contacting of said probe-tag array with said channel array glass. An electric field directed along the axis of said separatory channels and oriented with an increasingly positive electrical potential from said probe-array to the distal opening (exit port) of said separatory channels is established, and conditions (e.g. temperature) are adjusted to disrupt the binding of said sample molecules with said probe-tag array such that sample molecules are freed and immediately undergo field directed transport into and along the said separatory channels situated closest to the respective probe-tag array element. Thus, sample molecules enter particular separatory channels according to the array element to which said sample molecules first bound, and according to which separatory channel said array element was juxtaposed. Thus, means, including automated instrumentation, embodying this aspect of the present invention should preferably juxtapose such arrays in a repeatable manner to said surface such that the sample molecules bound to a known array element will be separated by a particular, thus predetermined, separatory channel.

Here, the problem of loading a particular sample (or set of samples in the case of multiplex separations) into one particular separatory channel out of many is solved by situating a binding

activity in proximity to the entry port of a separatory channel. A variation comprises the transfer of sample molecules from a probe-tag array to a surface of a manifold, with binding of sample molecules to said surface of a manifold, after which said manifold is used to carry sample molecules into arrays of vessels, if any, to perform reactions, if any, and comprises transfer of sample molecules or reaction products therefrom to separatory channel arrays, channel glass arrays or capillary arrays.

Using such arrays, two sample fraction detection schemes are possible. The first involves photodetection of molecular species, which may involve species labeled with fluorescent dyes. In this case said channel array glass must be transparent at least to the excitation frequencies used, and where emitted photons traveling in any direction other than that of the channel axis direction, further transparent to the wavelength of emission of the dyes used. Fractionated species may be detected by scanning a laser beam of appropriate frequency through said arrays, and detecting fluorescence of fluorescent labels as a function of beam position and linear position along said beam, as proposed by S. Esener et al.[37] for three dimensional optical memories with one-beam addressing. Such a one-beam approach may comprise the scanning of said one beam across planes comprising the axis of each of many channels or across a plurality of channels at a fixed position along the length of all such channels as separation proceeds, and in either case recording the dye emission signals thus associated with particular positions or times with each channel of said array. Planar beams, also used by Esener et al. for two dimensional readout, may similarly be used with two-dimensional photodetection means (e.g. CCD array camera) to either collect spatial separation data along the lengths of those channels whose axes substantially occur in the plane of said planar beams, or to collect signals regarding the passage of species through said plane (i.e. according to a temporal separation coordinate), here oriented perpendicularly to the axis of each channel Note that for such far-field photodetection based data collection schemes, the useful minimal dimensions of glass channel arrays will be determined by the resolution limit, i.e. related to the wavelengths of the light used.

In the second detection scheme used with channel array glass parallel separation means, molecular species are detected as they exit separation channels. Here, the output port comprising surface of said channel array glass may be briefly contacted with a receiving surface, such as an ion exchange paper surface or nylon membrane, onto which sample molecules are transported by the action of applied electrical field (or fluid flow in the case of ion-exchange border matrix capillary modifications described herein) and are bound to said receiving surface (e.g. by crosslinking with UV radiation in the case of nylon membranes.) Said contacting is performed such that sample molecule species bind to the portion of said sample receiving surface immediately juxtaposed to the channel from which each such molecule exited (i.e. such that sample molecules of an individual molecular species do not diffuse to regions juxtaposed to channels from which they did not respectively exit). This may be facilitated by performing said contacting with said receiving surface under high pressure normal to the juxtaposed surfaces, and may further be facilitated by the use of elastomeric surfaces or underlayers to cause said receiving surface to conform to said channel array glass surface comprising exit ports. Such brief contacting steps are repeated with fresh receiving surfaces as separation proceeds, preserving information as to the order in which receiving surfaces or receiving surface regions were contacted. Said receiving surfaces are then examined for the spatial location or pattern of labels, from which information about the sample sequences separated in each respective channel is reconstructed. Said pattern may favorably be captured by imaging visual labels (e.g. fluorescent labels, affinity labels bound by receptor-comprising beads, etc.,) with photodetector arrays such as CCD cameras, and analyzed by digital computer for information content, which is then reordered and further analyzed to glean sequence information.

It should be noted that various treatments of such channel glass arrays may further facilitate their use as means for parallel electrophoretic separations. First, though R.J. Tonucci et al.[38] do not note this, decreases in channel definition resulting from interdiffusion of the two or more glass types during drawing and annealing steps entail increases in internal surface roughness of the capillaries thus formed, when such interdiffusion is not so severe as

to completely interfere with channel formation. Surface roughness of internal channel surfaces will affect the separation properties of such channels in possibly favorable ways, which effect will depend partly on the internal volume to internal surface area ratio describing such channels.

A second category of treatments depend on the chemical modification or derivatization of said internal surfaces. As reviewed by A. Ulman,[39] the self-assembling monolayer systems based on organosilane compounds and derivatives thereof (e.g. $RSiX_3$, R_2SiX_2 or R_3SiX where R groups may be identical or distinct), which form covalent linkages to surface hydroxyl groups (e.g. formed on surfaces with exposed SiO_2 including native silicon oxides) may provide control over surface properties according to the particular compounds chosen. Other chemistries may similarly be used to derivatize such surfaces. This fact has been exploited in capillary zonal electrophoresis[40]. Derivatives conventionally include polyacrylamide derivatives, though these have generally been used for the reduction of protein adsorption to such surfaces. The embodiments of the present invention availing capillary tubes and channel array glass may further avail similar treatments. In particular, such treatments may provide for a channel comprising a surface localized separation matrix (border matrix) which may occupy a substantial portion of the inner diameter of such channels. Such a border matrix may comprise a predetermined polymer chain length or length distribution or composition distribution.

In particular, for the channels of channel array glass, or of capillary tubes, or of microfabricated separation channels, ionizable chemical functional groups situated on such border matrices may be used to modulate the degree of interaction of polynucleotides with such matrices, in analogy to ion-exchange resins. For example, pH of a running buffer (which occupies such channels and may flow through these independently of any border matrix) may be chosen to effect the desired degree of border matrix ionization, up to the maximum determined by the surface density of ionizable groups. Other conditions, such as temperature, affecting the ionization of chemical groups may similarly be exploited. Polynucleotide fragments comprise negatively charged phosphate linkages in direct proportion to length; for separation of polynucleotides, said ionizable groups are chosen

to yield cations upon ionization, in order to bind said negatively charged phosphate linkages. Degree of border matrix ionization will affect the probability that each negatively charged phosphate group associates with an ionized group of said border matrix. The different number of anionic groups corresponding to the different polynucleotide molecules will result in a difference degree of association of different length polynucleotides to such surfaces. Conditions and border matrix composition (and hence ionizability and ion exchange properties) are favorably chosen such that each polynucleotide fragment in the size range to be separated has an appreciable characteristic mobility, fractional association time or holdup, arising from differences in the mean number of associations with said border matrix per unit length. Optimal conditions and border matrix compositions will vary according to border matrix composition and size fractionation range, and the many parameters which may be varied to effect a desired result provide numerous possible combinations for many desirable fractionation properties. The mechanisms involved in these types of size fractionation may, for many cases, be envisaged as the association at a few points of polynucleotide backbones with a border matrix, where a characteristic proportion of said associations spontaneously dissociate per unit time, and said polynucleotide molecules thus *locally* freed (at one or more previously bound points along the length of said polynucleotide molecules) translate, to the extent possible under the constraint of other points of association persisting at that moment, in the direction of an applied electrical field or fluid flow. Thus, over time, a polynucleotide (or more generally, any polyions or polyanions subjected to such separation) advances in the direction of said applied field or fluid flow, at a rate which is negatively correlated with length (or more specifically, for polyions or polyanions in general, with number of cationic and/or anionic groups.) In the limit of less than on average one association per unit time, mobility will depend on the ratio of time which any such ion comprising species subjected to such separation spends in association with a border matrix ion to the time which it spends unassociated, and further depend on the mobility of said ion comprising species when unbound and in either the border matrix and overlying fluid phase (and the fraction of time spent in each of these, and mobilities in

each of these.) When transport in such a separation is directed by an electrical field, such methods may be termed ion-exchange resolved electrophoresis (or ion-exchange zonal electrophoresis), and when transport is directed by fluid flow such methods may be termed fluid injection directed ion exchange fractionation. Such methods bear some similarity to thin-layer chromatography, but are crucially distinguished by the presence of a free-fluid or mobile phase constrained to a limited proximity to said border matrix by channel dimensions.

Note that for any of these surface modification methods, particularly for surface modification aimed at providing surfaces suitable for ion-exchange based resolution methods, binary masking methods used with physical masks may be used to pattern surface density gradients or other distribution patterns of, for example, molecules comprising ionic or ionizable groups for separation methods involving ion exchange, where a surface used for forming a formed channel is exposed. Such a surface is a flat surface comprising linear trenches. Another such surface is a substantially flat surface which may be juxtaposed to said flat surface comprising linear trenches. Further, a large number of surface modification compositions and combinations thereof may be tested by such binary masking means for any desired sample type. This provides a method for the rapid optimization of surface modifying resolution media.

Combinations Including Partial Demultiplexation, Multiplex Separation or Parallel Detection:

Partial Demultiplexation and Parallel Multiplex Separation:

This aspect of the present invention combines partial demultiplexation, parallel multiplex separation and conventional serial reprobing of sequences separated in multiplex and transferred to membranes or the like. Parallel formed separatory channels capable of being opened length-wise may be used to enable transfer of separated species to membranes for probing. Alternatively, transfer to a moving membrane from the outlet port of capillaries or separatory channels simultaneous with separation may accomplish transfer of the fractionated sample conserving a separation

coordinate (i.e. the time of exit coordinate is thus transformed to a positional membrane transfer coordinate on said membrane, related by the velocity of said membrane relative to said capillaries or separatory channel). The principal advantages of this aspect of the present invention are a considerably reduced requirement for probe arrays (requiring only the demultiplexation array), and the combination of both highly parallel separations and the rapidity of separation in small channels, with separation in multiplex. In this instance, each serial reprobing cycle yields data up to a maximum equal to read length times the number of separation channels used. For example, for a read length of 500 bases with 1000 separatory channels and a serial reprobing time (probing, washing, probe detection and probe removal) of one hour, 12 million bases of sequence information may be collected per membrane per day. Of course, reprobing of multiple membranes (i.e. patterns transferred from multiple sequencing runs) will increase this rate correspondingly. Using the results of E.S. Yeung et al. of 18 minutes per fractionation, 80 membranes may be generated in one day. This will be practical for a separatory channel pitch of 100 microns and in the case of longtitudinal transfer, length less than 50 cm. It should be noted that such a method will require reduced use of a parallel separatory channel array corresponding to the number of serial reprobings performed. Automated membrane transfer and serial reprobing equipment such as has been described may be adapted to embody this aspect of the present invention.[41,42,43] Because of the small quantity of each sequence species in each fraction or band, a preferred probe oligo labeling method in this aspect of the present invention is involves labeling, (e.g. with biotin) and detection with dye labeled or fluorescent labeled affinity modified beads (e.g. streptavidin modified fluorescent latex beads). In this case, intensity encoding with four monochromatic levels in one lane (referring to a region of the membrane) or channel is the preferred means for reducing the number of channels required. Data capture is again performed with a CCD and digital computer.

Partial Demultiplexation, Parallel Multiplex Separation and Parallel Detection:

Because both demultiplexation and parallel detection of species separated in multiplex both rely on the binding of tag sequences, it will be explicitly noted that for these variations of the present invention, samples must be constructed or provided such that the identity of a clone or sequence is specified by at least two distinct tag sequences, though said at least two distinct tag sequences may be contiguous or even overlapping, and may be naturally occurring, generated randomly or designed deliberately. The first of said tag sequences is a demultiplexation tag, and the second of said tag sequences is a detection tag for probing in a parallel detection step.

This aspect of the present invention gains maximum multiplexation of sample preparation and of sequencing reaction multiplexation, exploits parallel separation and further exploits parallel detection of multiplexed species in each separatory channel, simultaneously (i.e. a further level of parallelism) for different separatory channels. The maximum number of sequence reads which may thus be obtained is equal to the degree of parallelism of partial demultiplexation (which equals the number of separatory channels used) times the number of parallel detection elements for each fraction. Variations on and methods and means for this aspect of the present invention are described below.

Parallel Multiplex Separation and Parallel Detection:

These aspects of the multiplex loading and parallel separation methods and means of the present invention are combined with the parallel multiplex methods described by E.M. Rabani, comprising the simultaneous detection of plural different species within the same separation fraction. Such methods and means permit the sequence determination rate possible with a given separation rate (r) and a given separatory channel parallelism (s) to be multiplied by the parallelism of detection (d), yielding a maximal rate of (r)(s)(d) determined bases per unit time. This is favorably accomplished by lateral transfer, which is to say transfer of all fractions in one step or substantially simultaneously to detection arrays in a manner such that each fraction or each portion of a fraction will be

transferred to a distinct detection array or distinct detection array region.

Lateral Transfer of Fractionated Multiplex Species to a Second Surface or Detection Array:

In many instances it may be desirably to avoid the need for continuously moving parts and other complexities associated with the transfer of eluted fraction from exit ports of laterally sealed separatory channels to a moving membrane. The drawbacks associated with such methods may be avoided by using separatory channels which may be laterally opened or have many lateral ports, such that transfer may occur laterally as in the case of transfer from slab gels to laterally juxtaposed or opposingly placed membranes in, for example, Southern blots.

In the case of the present invention, lateral transfer from one or more separatory channel array facilitates the use of parallel detection arrays, preferably comprising stripes of plural distinct tag-probe oligonucleotides or other probing compositions, with an array repetition of a similar pitch to that of said separatory channel array, such that a said stripe of each of said distinct tag-probe oligonucleotides or other probing compositions occurs for each separatory channel of said separatory channel array. Said detection array comprising stripes are aligned parallel to said separatory channels during transfer, such that the contents of any fraction in each of said separatory channels will be exposed, upon transfer to at least one, and preferably exactly one, repetition of said stripes of said detection array. Alternatively, said array may comprise stripes of fine pitch which will be oriented perpendicularly to said separatory channels, provided that the pitch with which a full repetition of said array is less than half the length occupied by a species of a particular length in said separatory channel, such that the molecules in each fraction (many of which may be identical to each other but different from most molecules in the multiplex fraction) will be juxtaposed to all detection array elements.

In this aspect of the present invention, separatory channels are preferably formed from molded or cast microreliefs or reliefs (first reliefs), preferably of elastomeric composition, comprising narrow parallel trenches, which define the ultimate separatory channels, and

with surfaces composed or treated so as to have good affinity to any resolution media used but no undesirable affinity to sample molecules (in cases of separation by a gel or low viscosity medium, as opposed to separation on border matrices, in which case the surfaces of said reliefs are treated accordingly.) Said reliefs are juxtaposed to either a substantially flat surface or a second relief surface comprising arrays (fraction transfer arrays) of fraction transfer grooves and electrodes for lateral electroelution, and preferably also integrated tag-probe arrays; such juxtaposition is followed by the application of sufficient normal pressure to compress regions of greatest extremity on said relief or reliefs (or in the case of non-elastomeric relief surface, of an elastomeric film or an elastomeric second relief surface comprising arrays of fraction transfer groves, which is used as the opposed surface) so as to cause complete lateral sealing of said trenches into separatory channels by compression of the elastomeric material.

Fluid separatory media such as low viscosity media including polyethylene-oxide or polyacrylamide may then be injected into the resulting channels. Alternatively, in cases where said channels are formed by the juxtaposition of said first relief comprising trenches with a flat surface, a liquid resolution media (from which bubbles have been removed, e.g. by degassing) may be spread uniformly on one of the two surfaces, preferably said flat surface, which is in this case preferably flexible, which is applied with a uniform unrolling motion or a squeegeeing motion such that only one meniscus forms and is evenly translated so as to prevent the formation of bubbles. By such means, the need to inject resolution media into small channels may be avoided.

Transfer may be favorably accomplished as follows. Subsequent to separation of plural different multiplex samples on said formed separatory channel array, a surface of said formed separatory channel array comprising lengthwise portions of the internal surface area of each of said separatory channels is removed, and later replaced with a transfer surface, which is thus juxtaposed to the surface comprising the remainder of said formed separatory channel array. Thus, lateral opening of said channels is performed by removing, for example, said flat surface from said relief surface, the aspect ratio or recession of the trenches therein favoring retention of any

resolution media. This removal step is performed, preferably with a lateral sliding motion to prevent adhesion of resolution media, such that resolution media and sample material substantially remains in the remainder of said formed separatory channel array, principally or exclusively within the trenches defining formed channels.

After removal of said surface of said formed separatory channel array comprising lengthwise portions of the internal surface area of each of said separatory channels, said relief surface comprising trenches Said transfer surface is substantially free of separation media before juxtaposition to said surface comprising the remainder of said formed separatory channel array.

Alternatively, in cases where said formed separatory channel array comprised said second relief surface comprising arrays of fraction transfer grooves and electrodes for lateral electroelution, and preferably also integrated tag-probe arrays, separatory channels may be performed by juxtaposition, as in the other case, but further with relative positioning such that said trenches of said first relief are immediately opposite a flat portion of said second relief surface. Said second relief is preferably coated with a degassed liquid layer to facilitate electroelution before said juxtaposition is performed, and the resulting meniscus is caused to translate evenly to avoid the formation of bubbles. After fractionation, said second relief surface is translated by a lateral sliding motion along the surface of said first relief such that two openings of each of said fraction transfer grooves (which preferably run perpendicularly to said formed separatory channel trenches) are opposite each trench on said first relief which defines said separatory channel. After said lateral sliding causes the proper alignment of separation and fraction separation channels, a voltage is applied to the electrodes on of said second relief so as to electroelute fractions from said separatory channels and cause molecules in said fractions to pass into a restricted volume contacting said detection array. After said electroelution, said second relief may be translated so as to minimize the opportunity for denaturants to seep from said separation medium in said separatory channels into said restricted volume. Conditions are then changed to establish the desired binding stringency and thus permit selective binding of tag sequences with the tag-probes of said detection array.

After lateral transfer of multiplex fraction to said parallel detection arrays (which occurs in a stratified manner) appropriate means are used for signal amplification, if any, and detection. For example, such detection methods and means may include contacting said parallel detection array to receptor or antibody labeled colored or fluorescent beads targeted to particular affinity labels used to label sequencing reaction products, e.g. according to terminator moiety. The positions in said parallel detection array and the associated fractionation coordinate are detected with suitable photodetection means, for example comprising a CCD array or photon camera and appropriate filters, and recorded by a digital computer interfaced to said photodetection means. Information concerning the sequence of sample material in communication with a particular second tag is reconstructed from the occurrence and identity of labels bound to the corresponding tag-probe of said tag-probe array in the ordered set of fractions, in a manner otherwise similar to sequence information reconstruction with other capillary gel electrophoresis sequencing methods. Said photodetection means and said detection array are preferably mounted on a system or translation table which moves one or both of these relative to each other such that data may conveniently be collected across the entire array by means of such relative motion or scanning. After data collection, said detection array is subjected to denaturing conditions and repetitive rinses to remove labels and sample molecules and thus prepare said detection array for reuse, care as always being taken not to contaminate said detection array.

Note that detection of samples on surfaces is facilitated by the signal amplification methods thus possible, including the aforementioned affinity label detection and signal amplification comprising the use of receptor labeled fluorescent beads, which methods are not favorably used with laterally sealed capillaries or gels.

Partial Demultiplexation, Parallel Multiplex Separation and Parallel Detection:

The same methods and means permitting the application of multiplex samples above to an array of separatory channels may be combined with the above described embodiment of the present invention, such that

partial demultiplexation by a demultiplexation array sequences sample material, preferably subsequent to sequencing reactions, targeting molecules comprising particular first tag sequences to corresponding particular individual separatory channels of said array of separatory channels, such that identity of said first tag (more specifically the tag or tag sequence utilized for demultiplexing, as distinguished from a second tag sequence used for parallel detection) determines the channel to which each molecule will be transferred. By these methods and the corresponding combined means, a large multiplex sample or sets of multiplexed samples may be applied to parallel separatory channels or separatory channel arrays in a single step, in preparation for the parallel multiplex separation and parallel detection for each channel described above.

Determination of Tag Sequences Not Present in a Complex Sample:

For many large scale sequencing applications of the present invention, it may be desirable to determine the sequences of a large number of candidate tag (and probe-tag) sequences suitable for use with that sample, where each such tag sequence occurs infrequently or not at all within a sample. One method for accomplishing this is as follows. A pool of oligonucleotides of a particular nucleotide length n is randomly synthesized from all for nucleotide monomer reactants by a suitable oligonucleotide synthesis chemistry (e.g. phosphoramidite). Length n must be chosen such that number of different sequences in said pool, 4^n, exceeds the size N (number of bases) of the sample by some factor greater than, for example, two. A sample of size N can comprise no more than N distinct n-mer sequences (N>>n). Thus, said pool comprises a large number of n-mers not present in said sample. The non-randomness and repetitivity of evolved DNA sequences increases this number.

We may consider the example of a genome of $N=10^9$ bases in size. Such a genome is of lower complexity than a pool comprising all possible 15-mer oligonucleotides. We may thus choose a pool comprising half of all possible 16-mers, by reducing the number of nucleotides (i.e. to <4, e.g. to two) at one position of said n-mers. Such a pool is produced in quantity sufficient to ensure that on average at least, for example, ten thousand copies of each possible

sequence occur. At least half of this quantity is left on the solid phase support used in synthesis but is nonetheless subjected to removal of protecting groups. A quantity equivalent to 100 copies of each sequence in the pool is fractionated against genomic DNA. This is equivalent to about 2ng or 0.4 pmol. This quantity of such an oligonucleotide pool is mixed with a tenfold excess of genomic DNA (which may have previously been filtered to eliminate small fragments from the sample). The resulting mixture is subjected to denaturing conditions and then hybridized under reduced stringency hybridization conditions. Reduced stringency conditions are favorable because of the variation optimal hybridization conditions for highest stringency with (oligo) sequence; reduced stringency conditions will eliminate sequences which, although they do not occur in said genome, nonetheless may bind to genomic sequences under the set of conditions used, while minimizing the likelihood that any oligos with sequences occurring in said genome will fail to bind. Unbound oligonucleotides are recovered after hybridization, for example by ultrafiltration (care being taken to avoid excessive fragmentation by shearing, etc.,). The unbound fraction may be subjected to repeated rounds of depletion by repetition of the foregoing mixing, denaturing, hybridization and filtration steps, against the same quantity of genomic DNA used initially, to greatly enrich for sequences not occurring in said genomic DNA and greatly reduce the presence of genomic sequences in said unbound oligonucleotide fractions. Different conditions may be used in different depletion rounds.

Following depletion against genomic DNA, the fraction unbound after the final depletion cycle may optionally then be subjected to selection for binding of the undepleted pool, under more stringent conditions. This is again performed by mixing the undepleted, solid phase bound pool in, for example, fifty-fold excess, denaturing and hybridizing under more stringent conditions. Washing or filtering is then performed to eliminate the unbound fraction. The bound fraction may then be recovered by denaturation and washing, and will appear in the eluted fraction. Such a selection will enrich for sequences which bind their complementary sequences under more stringent conditions. Such enrichment will reduce the number of oligo sequences not bound to genomic DNA but exhibiting unacceptable non-specificity of binding against the lager complexity random pool, due

to the increased stringency used. This treatment does not, however, formally preclude the possibility of recovery of oligos which bound to less than fully complementary targets because no discrimination of such events is possible, and these are only reduced with increasing stringency. Further, binding of oligos in this fraction to more than one tag-probe cannot be excluded, but will be improbable under sufficiently stringent conditions. Longer random tag sequences (relative to a fixed sample genome complexity) prepared as above and used under correspondingly higher stringency binding conditions in sequencing protocols and other related protocols may address this issue by rendering the likelihood of one probe-tag binding multiple tags or one tag binding to other tags correspondingly less likely in finite sample sizes. Further, to the extent that such events occur, these will only reduce overall efficiency of parallel detection or demultiplexing across a large sample diversity (e.g. samples comprising many different distinctly tagged sequence inserts); these inefficiencies are readily offset by the parallelism and multiplexing possible with the methods of the present invention.

The fraction thus recovered comprises sequences favorably used as tag-probes, sequences complementary to these may favorably be used as tags for library clone tagging. Sequence determination of the oligos obtained by the above methods may be performed as follows. Two 5' phosphorylated oligos, complementary to each other over most of their length but having, for example, a random three-base 3' overhang at one end (but blunt at the other) are annealed to an excess of depleted tag-probes, and then ligated to these, such that a ligation event occurs at the 3' end of oligos from the pool of said depleted tag-probes. The resulting products are then subjected to polymerization (primer extension) to fill in the overhangs thus formed comprising 13 of the 16 bases of said depleted tag-probes. Finally, the resulting mixture is subjected to blunt-end ligation (e.g. with T-4 DNA Ligase) to form large multimers, which are then ligated into a sequencing vector, or a vector comprising priming sites suitable for PCR amplification of the insert material. These vectors are then transformed into hosts, or the inserted material is amplified by PCR after dilution and distribution into wells (e.g. of a microtitre dish suitable for PCR cycling) at a concentration yielding on average one amplifiable molecule per well. Vectors used

favorably comprise a second tag sequence immediately downstream from a standard sequencing primer sequence and immediately upstream from the inserted material. Said second tag sequence comprising a limited number of variable bases at predetermined positions, e.g. 6 bases, not necessarily all contiguous, such that probe arrays complementary to all of the possible second tags may be formed and used for one or more of the demultiplexing or parallel detection methods of the present invention. A third tag sequence of predetermined complexity may similarly be included, for example, adjacent to said second tag sequence, and may comprise a similar complexity but a distinct consensus sequence, and be used in a step other than where said second tags were used for demultiplexing or parallel detection, said third tags used for parallel detection or demultiplexing. Thus, a set of tags from the depleted pool may be sequenced, and the methods of the present invention may be used to rapidly determine the sequences thus recovered, where said second and said third tags permit, for example, demultiplexing (here, partial demultiplexation) for separation of a complex pool in parallel separatory channels by binding to arrays comprising probes for said second tags, and parallel detection is accomplished with arrays of probes complementary to said third tags. Sequence information gained about sequences not occurring with such a genome is favorably then used to generate probe tag arrays for used with said genome. Further, the cloned or amplified depleted tag material may be constructed such that individual tags are liberated by digestion with a predetermined restriction enzyme, and these may then be used in the generation of a pool of tagged derivatives of a library vector construct. Alternatively, sequence information regarding said depleted tag material may be used to score consensus sequences and m-mers (m<n) which may be used as elements in the synthesis of pools of oligos comprising sequences of highly unlikely occurrence in said genome.

For tagging downstream from a primer, in which case a standardized primer of identical sequence may be used with all clones from such a library such tag sequences are preferably placed within a library vector immediately downstream from a sequencing primer region, and immediately upstream from a cloning site. Thus, such tag sequences will be copied (or otherwise included) during the generation of

sequencing reaction products, and all but the shortest sequencing products will comprise the tag sequence.

Constructs related to the vectors described above are disclosed by G.M. Church and S. Kieffer-Higgins in U.S. Patent Number 5,149,625. Such constructs may similarly be used with certain various aspects of the present invention. Where tag-probe arrays are to be used for demultiplexing or parallel detection, said probe-tag array oligonucleotide ensemble composition must comprise oligonucleotides having sequence composition which will generally, and preferably uniformly, each probe a distinct tag of such vectors under sufficiently stringent binding conditions. Thus, one aspect of the present invention is the demultiplexation based separation of libraries or other pools of different sequences constructed from the vectors of G.M. Church and S. Kieffer-Higgins.

Copolymer Array and Probe-Tag Array Patterned Synthesis:

Synthesis Patterned by Lithographic-Plate:

Lithographic plates such as those taught by T. Suzuki and F. Shinozaki in U.S. Patent Number 5,380,620, or those reviewed by B. Bednar, J. Kralicek and J. Zachoval,[44] which accomplish the spatially predetermined transfer of liquid patterns, where said liquid is conventionally an ink, colorant or etch-resist, may be applied to the fabrication of copolymer arrays. Copolymer types may include peptides and oligo- and poly-nucleotides, as well as other copolymers which may be synthesized by deprotection step-controlled chemistry on solid phase supports. This method is in analogy to that of S.P.A. Fodor et al., but substitutes spatial control of reagent placement according to the wettability of an juxtaposed surface contacted with an array substrate for spatial control of removal of photolabile protecting groups. The principle advantages of this present method is that monomers (comonomers, e.g. phosphoramidite deoxynucleotides of each of the four nucleobase compositions found in naturally occurring DNA) need not be chosen according to photodeprotection requirements, which are eliminated, and the use of light or actinic radiation exposure equipment is also eliminated. Thus, standard peptide synthesis reagents or polynucleotide (e.g. DNA) synthesis reagents and comonomers may be used. Three variations are possible.

In a first variation, lithographic plates are used to contact regions of said array, corresponding to those regions desired to undergo deprotection of the copolymers or initiators thereupon, with deprotection reagents included in said liquids ("inks") on said plates. The entire surface of said array is then exposed to the reactant (comonomer) to be added only to the copolymers or initiators deprotected by the preceding treatment. By such means, the addition of comonomers to only those copolymers or initiators situated in a spatially predetermined set of regions (according to the patter and relative positioning of the lithographic plate used in the immediately preceding step) is effected. Such steps are repeated, using different lithographic plates aligned relative to the position of precedingly used lithographic plates, in an overall procedure otherwise resembling the light controlled method of S.P.A. Fodor et al., to produce the desired pattern of copolymers with predetermined sequence composition and situated in a predetermined manner which corresponds to said sequence composition on said array substrate surface. This first variation may be termed lithographic plate patterned deprotection. A second, contrasting, variation may be termed lithographic plate patterned comonomer contacting and addition. Here, the totality of one or more such array is contacted with deprotection reagents to remove any protecting groups from any preceding steps, and then contacted with a lithographic plate "inked" with a liquid comprising a single comonomer reactant type. Comonomers are thus added only to those copolymers or initiators in those array regions contacted with said liquid, i.e. contacted with said reactant. With this method, a wash step is favorably performed using reagents that will inactivate unadded comonomer reactants, such that the reactivity of these is neutralized before they are swept into possible contact with regions comprising copolymers or initiators with which said reactants are not desired to react. The advantage of this variation lies in the reduction of steps from the identity of patterning and reactant contacting steps. As in any patterned copolymer synthesis method, the final pattern is determined by the lithographic plate pattern used, the order in which these are used, and the alignment and precision of alignment with which they are repeatably contacted with said array substrate. The third variation combines the patterning methods of the previous variations,

and may be termed lithographically patterned deprotection with redundantly patterned comonomer contacting and addition. Here, lithographic-plate control is exerted, as in the first variation above, to effect the deprotection of initiators or copolymers situated in spatially predetermined array regions in a first cycle step, and in a second cycle step either the same lithographic plate or another plate with an identical pattern is used, as in the second variation above, to contact reactants with only the desired predetermined regions substantially identical in location and extent to those deprotected in said first cycle step, after which a wash step is performed as in the second variation above. Such a cycle is repeated with different lithographic plates to produce the desired final copolymer array pattern.

The large areas which are readily patterned by such lithographic means make these methods particularly advantageous in the high volume production of copolymer arrays, particularly arrays of extended dimensions. It must be emphasized that the patterning method disclosed here relies primarily upon the differential wetting of surface regions of said lithographic plates rather than any other principle of liquid pattern formation and transfer.

These methods may be availed to produce oligonucleotide arrays useful for sequencing by hybridization, for polynucleotide characterization, for the determination of overlap of clones in a DNA library, for mutation or genetic disease detection, or for to produce the probe-tag arrays used in the parallel multiplex sequencing method previously described by E.M. Rabani cited above. These methods, as well as any other functionally adequate or comparable methods, including the photodeprotection methods of S.P.A. Fodor et al., or those taught by J.L Winkler, S.P.A. Fodor et al., in U.S. Patent Number 5,384,261 comprising steps of flowing reagents through groves on the surface of a channel block juxtaposed to a substrate, or those taught by R.E. Southern and co-workers in U.S. Patent Number 5,324,663, may be employed to produce the linear probe-tag arrays disclosed herein as useful for parallel multiplex sequencing with direct spatial transfer including parallel multiplex sequencing with parallel separation and direct spatial transfer.

Note that the lithographic plates used in this aspect of the present invention may have patterns comprising isolated inked or

isolated wettable regions, as retention of applied solutions (rather than transport through channels to desired surface areas, as in some prior art methods) does not require transport through contiguous regions. Note further that lithographic plates of this kind are favorably produced by the microfabrication methods described and referenced above, and thus generally comprise films of polymeric resist-type coatings which are usually of 50 nm or greater depth from the surface, being deposited by such methods as spin-coating.

In summary, this method comprises the steps of: (a) producing lithographic plates that specify the desired spatial patterns of chemical reactions, in particular, the desired spatial patterns of addition of reagents or reactants; (b) applying said reactants or reagents to one of said lithographic plates for use in a reaction cycle according to the final pattern desired, with the result that solutions are only retained in selected predetermined regions according to the pattern of the respective lithographic plate; (c) juxtaposing the lithographic plate of step (b) to the target surface on which copolymer initiators or intermediate products are situated such that reagents or reactants are contacted and thus permitted to react with only the desired regions specified by the pattern of said lithographic plate, where said juxtaposing constitutes opposing said lithographic plate to said target surface in correct alignment, register and relative position (e.g. relative to some reference point which facilitates the correct registry of material applied by different lithographic plates); (d) performing any further necessary chemical reactions on said target surface which do not require spatial limitation; and (e) repetition of steps (b) through (d) as many times, and with lithographic plates of correct pattern, as necessary for the desired copolymer array pattern to result. As will be obvious to those skilled in the relevant arts, such a method will be suitable for any copolymer which may be produced by a synthesis procedure wherein some controlling step is performed with a solvent or reagent for which there can be found a material composition to yield a surface which repels said solvent or reagent and for which there can be found a materiel composition to yield a surface which retains said solvent or reagent. Copolymers of particular interest which may be produced by this method are DNA, RNA, peptide-nucleic

acids, peptides and the various peptide analogues or mimics, and obvious variations upon all of these.

Copolymer and Modified Oligonucleotide Synthesis by Patterned Activation with Metal Catalyst Structures:

Probe-tag arrays, arrays comprising oligonucleotides which need not be probe-tags, and most generally copolymer arrays may be produced by patterned synthesis comprising one or more steps of spatially controlled catalysis, particularly spatially controlled deprotection or activation. One aspect of this method, comprising the catalytic reduction of nitrile groups in the presence of dissolved H_2 to yield amine groups will be presented here for illustration. It will be noted, however, that the essence of this method concerns the exertion of spatially predetermined step control in the synthesis of copolymers situated on surfaces, and may be adapted to other chemistries and corresponding reactions catalyzed by metal surfaces.

For the synthesis of backbone modified oligonucleotides, nucleotide synthesis reactant monomers comprising 5' nitrile groups, and not comprising phosphoester protecting groups which include nitriles (such as is the case for cyanoethyl groups used with phosphoramidite oligonucleotide synthesis chemistries) or any unsubstituted amides are used. The active acceptor group in preferred reactant monomers is a 3' -($OPOCH_3Cl$) activated phosphate. This group is chosen to be reactive with amines (by attack of amine nitrogen lone pair at the partially positive phosphate, with liberation of the chlorine as a stable anionic leaving group), and to not comprise any nitrile or unsubstituted amide functionalities. Initiators used, which are preferably of substantially linear structure, comprise a surface binding group (e.g. an organosilane for glass substrate surfaces), a linker of sufficient but not excessive length (e.g. having linkers 10-1000 atoms or bonds in length), and a terminal nitrile group. Longer initiator linkers will impose limitations on minimum array element size and spacing. Like azide groups, nitrile groups may be reduced by platinum in the presence of dissolved H_2.[45]

Patterned platinum catalyst structures or microstructures, which may, for example, be patterned and fabricated by LIGA methods,

comprising substantially flat surfaces from which a pattern is recessed or omitted (in the form of voids), provides for the spatially predetermined patterned activation of nitrile groups situated on a surface. The patterned surface of said patterned platinum catalyst structure is contacted, in the presence of dissolved H_2, with said substrate surface on which said nitriles are situated; said nitriles experiencing sufficient contact with the surface of said platinum are thus transformed to amine groups. Said contacting is performed in a manner and with means such that a particular platinum catalyst structure may be repeatably contacted with the same surface with high spatial accuracy, or with high accuracy stepping (as in semiconductor microlithography). The entirety of said substrate surface is then contacted with nucleotide reactant monomers such as those described above comprising the base moiety (in protected form) desired to be added to the amines formed in the preceding step; nitriles are substantially inert to methoxyphosphorochloride ester addition, as evidenced by their use as backbone protecting groups in other oligonucleotide synthesis chemistries employing PCl_3 derivatives. A catalytic patterning step followed by a monomer addition step comprise a patterned addition cycle. Other steps necessary are identical to those used in standard phosphoramidite chemistries; these include capping and oxidation steps. Patterned addition cycles are repeated with different patterned platinum catalyst structures or microstructures, or the same structures in different registry or juxtaposed orientation, and different reactant monomers such as those described above comprising different base moieties each of which said moieties is desired to be added to the amines formed in the preceding catalytic patterned activation step. The oligonucleotide synthesized by nitrile/amine step control methods such as these will comprise backbone linkages replacing $5'CH_2\text{-}(NH)\text{-}PO_2\text{-}O\text{-}3'C$ for the usual $5'CH_2\text{-}O\text{-}PO_2\text{-}O\text{-}3'C$. Under conditions where this backbone nitrogen is not further protonated (i.e. under mildly basic conditions), hybridization properties of oligonucleotides with such modifications are expected to be similar to those of unmodified oligonucleotides of identical sequence. The logic and considerations pertaining to the spatial patterns chosen and order of comonomer addition to yield an array with a desired oligonucleotide array sequence compositional

configuration are substantially identical to those pertaining to the methods of S.P.A. Fodor et al.,[46] or other similar methods.

Note further that such methods could activate regions of surfaces to the immobilization of small-molecule haptens or ligands, for use in production of arrays useful for parallel assays for the determination of affinity or binding of receptors, immunoglobulins, enzymes, cells, ribozymes or other macromolecules. Likewise, biomimetic copolymers which may be prepared from monomer precursors comprising nitrile groups may likewise be thus prepared.

Copolymers of particular interest which may be produced by this method are backbone-modified-DNA, backbone-modified-RNA, peptide-nucleic acids, peptides and the various peptide analogues or mimics, and obvious variations upon all of these.

Copolymer and Modified Oligonucleotide Synthesis by Patterned Activation with Metal Catalyst AFM Tip Surface:

The same synthesis methods used above with patterned metal catalyst surface controlled activation patterning may be used to synthesize copolymer arrays with nanometer resolution by the use of scanning probe positioning means (e.g. a scanning probe microscope) to position a catalyst metal coated probe (e.g. an AFM tip), as was done by W.T. Müller, P.G. Schultz, et al.[47] but otherwise according to the above oligonucleotide synthetic methods. In other words, activation of initiator nitriles or nucleotide or oligonucleotide nitriles, to form amines which will react as donors with phosphate derivative groups on nucleotide monomer reactants is performed here by the positioning and contacting of the platinum on an AFM or other similar tip with said substrate surface, but synthesis is otherwise conducted as above.

Arrays used with the present invention may be used to bind polynucleotide species labeled with one among several different affinity labels, which affinity labels may be detected with beads comprising receptors or antibodies or other binding means specific for each of said different affinity labels. Each of said receptors or antibodies or other binding means is placed in communication with a perceptibly distinct bead (e.g. streptavidin with red beads and antidigoxigenin with yellow beads).

Note that such tag-probe arrays may be synthesized on high dielectric, low dielectric or conducting materials with appropriate initiators (e.g. on Au(111) with linkers comprising terminal -SH groups) or on membranes of polymeric composition (e.g. nylon, latex, etc.,) chosen or treated so as to be inert to synthesis conditions and reagents, where transfer of species from a gel by fluid absorption (as with the transfer of polynucleotides from gels to membranes in Southern and Northern blots) is desired, which choice will be determined by the desired method of transfer of fractionated species in each separation channel of a separation channel array to said tag-probe arrays. Probe-tag arrays situated on surfaces of such compositions favorable for blotting transfer may be produced by the methods and means disclosed herein, by those of known art including those photolithography based methods of S.P.A. Fodor et al. or any other equivalent means compatible with such surfaces, appropriate initiators and linkers being chosen.

The foregoing embodiments and examples have been presented for illustration rather than limitation, with the many possible variations, changes and alternative embodiments which will be obvious to those of ordinary skill in the relevant arts included within the scope and spirit of the present invention. Thus, the scope and the breadth of the present invention is intended to be limited only by the appended claims rather than the foregoing description;

---

[1]Maxam, A.M. ; Gilbert, W.; **1977**. *Proc. Natl. Acad. Sci. U.S.A.* 74:560.

[2]Sanger, F.; Nicklen, S.; Coulson, A.R.; **1977**. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463.

[3]See, for example, a standard laboratory manual is: Sambrook, J.; Fritsch, E.F.; Maniatis, T.; **1989**. *Molecular Cloning: A Laboratory Manual*, Second Edition. Cold Spring Harbor Laboratory Press.

[4]For example, low viscosity media as taught U.S. Patent Number 5,374,527.

[5]Church, G.M.; Gilbert, W.; **1984**. *Proc. Natl. Acad. Sci. U.S.A.* 81:1991.

[6]Cherry, J.L.; et al.; **1994**. *Genomics, 20*:68.

[7]Church, G.M.; Kieffer-Higgins, S.; **1992**. U.S. Patent Number 5,149,625.

[8]Church, G.M.; **1992**. U.S. Patent Number 5,149,625.

[9]Rickwood, D.; Hames, B.D., eds.; **1982**. *Gel Electrophoresis of Nucleic Acids: A Practical Approach*. I.R.L. Press, Washington, D.C.,.

[10]Sambrook, J.; Fritsch, E.F.; Maniatis, T.; **1989**. *Molecular Cloning: A Laboratory Manual*, Second Edition. Cold Spring Harbor Laboratory Press.

[11]Lagerkvist, A.; et al.; **1994**. *Proc. Natl. Acad. Sci., U.S.A., 91*:2245.

[12]This difficulty has been noted by both E.S. Yeung and L.S. Smith. See news article, Borman, S.; 1995. Chemical and Engineering News, 24 July 1995, p.37-39.

[13]Tonucci, R.J.; et al.; **1992**. *Science, 258*:783.

[14]Woolley, A.T.; Mathies, R.A.; **1994**. *Proc. Natl. Acad. Sci., U.S.A., 91(24)*:11348.

[15]Fung, N.E.; Yeung, E.S.; **1995**. *Analyt. Chem., 67(13)*:1913.

[16]Huang, X.C.; Quesada, M.A.; Mathies, R.A.; **1992**. *Analyt. Chem., 64(18)*:2149.

[17]Aumiller, G.D.; Chandross, E.A.; Tomlinson, W.J.; Weber, H.P.; **1974**. *Jour. Appl. Phys., 45(10)*:4557.

[18]Becker, E.W.; Ehrfeld, W.; et al.; **1986**. *Microelectronic Engineering, 4*:35.

[19]Abraham, M; Ehrfeld, W.; et al.; **1994**. *Proc. SPIE, 2213*:48.

[20]Gale, M.T.; et al.; **1994**. *Proc. SPIE, 2213*:2.

[21]Neyer, A.; et al.; **1993**. *Electronics Letters, 29(4)*:399.

[22]Klein, R.; Neyer A.; **1994**. *Electronics Letters, 30(20)*:1672.

[23]Bednar, B. ;Kralicek, J.; Zachoval J.; **1993**. *Resists in Microlithography and Printing*. Elsevier. See pages 289-294.

[24]A recent example of this work may be found in Wilbur, J.M.; Whitesides, G.M.; et al.; **1995**. *Advanced Materials, 7(7)*:649.

[25]Müller, W.T.; Schultz, P.G.; et al.; **1995**. *Science, 268*:272.

[26]McIntyre, B.J.; Salmeron, M.; and G.A. Somorjai; **1994**. *Science, 265*:1415.

[27]For example, see Beavis; R.C.; Chait; B.T.; **1994**. U.S. Patent Number 5,288,644.

[28]For example, see Jett, J.H.; Keller, R.A.; *et al.*; **1990**. U.S. Patent Number 4,962,037.

[29]For examples, see: Driscol, R.J.; Youngquist, M.G.; Baldeschwieler, J.D.; **1990**. *Nature, 346*:294;

Dunlap, D.D.; Bustamante, C.; **1989**. *Nature, 342*:204;

Allison, D.P.; *et al.*; **1992**. *Proc. Natl. Acad. Sci., U.S.A. 89*::10129;

and for a review, see: Stine, W.B.; Rabani, E.M.; Smith, D.W.; *et al.*; **1995**. "Imaging of DNA Bases and Oligodeoxynucleotides using Scanning Tunneling Microscopy," in: *Frontiers in Biological Physics*. Academic Press, San Diego.

[30]Lindsay, S.M.; Philipp, M.; **1992**. U.S. Patent Number 5,106,729.

[31]Fodor, S.P.A.; et al.; **1991**. *Science, 251*:767.

[32]Bains, W.; **1994**. *Genome Analysis, Techniques and Applications, 11(3)*:49.

[33]Broude, N.E.; et al.; **1994**. *Proc. Natl. Acad. Sci. U.S.A., 91*:3072.

[34]Broude, N.E.; et al.; **1994**. *Proc. Natl. Acad. Sci. U.S.A., 91*:3072.

[35]Fodor, S.P.A.; et al.; **1991**. *Science, 251*:767.

[36]Lagerkvist, A.; et al.; **1994**. *Proc. Natl. Acad. Sci., U.S.A., 91*:2245.

[37]See, for example, Ford, J.E.; Esener, S.; et al.; **1993**? *Proc. SPIE, 2026*:604.

[38]Tonucci, R.J.; et al.; **1992**. *Science, 258*:783.

[39]Ulman, A.; **1991**. *An Introduction to Ultrathin Organic Films, From Langmuir-Blodgett to Self-Assembly*. Academic Press, Inc.,. See especially pages 237-304.

[40]Weinberger, Robert; **1993**. *Practical Capillary Electrophoresis*. Academic Press, Inc., p.56-59.

[41]Cherry, J.L.; et al.; **1994**. *Genomics, 20*:68.

[42]Church, G.M.; Kieffer-Higgins, S.; **1992**. U.S. Patent Number 5,149,625.

[43]Church, G.M.; **1992**. U.S. Patent Number 5,149,625.

[44]Bednar, B. ;Kralicek, J.; Zachoval J.; **1993**. *Resists in Microlithography and Printing*. Elsevier. See pages 289-294.

[45]See, for example, Roberts, J.D.; Caserio, M.C.; 1977. Basic Principles of Organic Chemistry, Second Edition. W.A. Benjamin, Inc.,. p. 1146.

[46]Fodor, S.P.A.; et al.; **1991**. *Science, 251*:767.

[47]Müller, W.T.; Schultz, P.G.; et al.; **1995**. *Science, 268*:272.

whereforth, I claim:

5

Claims:

1. Method for determining the sequences of plural polynucleotide samples comprising one or more of: manifold sample transfer; loading of two or more distinct separatory channels or means or regions thereof each comprising one or more surface immobilized oligonucleotide sequences at or near the input of said separatory channels or means or regions thereof and comprising a step of hybridizing said polynucleotide samples to said surface immobilized oligonucleotide sequences; loading of products resulting from the amplification of polynucleotide molecules randomly distributed into one or more bin arrays;
and further comprising one of: transfer of fractionated polynucleotide molecular species to a separate surface or a surface region not part of the surface of a separatory channel, detection probe array or membrane subsequent to separation or as separation is conducted; transfer of fractionated polynucleotide molecular species to a tag-probe array during or after fractionation; transfer of fractionated polynucleotide molecular species to a tag-probe array after fractionation by means of an array comprising fraction transfer grooves and two or more surface immobilized tag-probes and electrodes for lateral electroelution.

2. The method of claim 1 further comprising separation steps with separation means selected from the group consisting of: parallel capillary array, parallel separatory channel array, channel array glass, microfabricated separatory channels.

3. Method for determining the sequence of two or more different polynucleotide samples simultaneously according to the method of claim 1 comprising sample demultiplexing means situated at or near the entry port of each of two or more separatory channels.

4. Method for determining the sequence of two or more different polynucleotide samples simultaneously according to the method of claim 3 comprising two or more separatory channels which may be opened along their length and tag-probe or oligonucleotide arrays

comprising two or more different sequences with patterns aligned along or repeating periodically perpendicularly to the direction of said separatory channels.

5. Method according to claim 1 further comprising the use of optical detection means for detecting tag-probe array or parallel detection array bound species .

6. An automated device for sequencing polynucleotides according to the method of claim 5, comprising one of:

one or more manifolds for sample transfer; two or more distinct separatory channels or means or regions thereof each comprising one or more surface immobilized oligonucleotide sequences at or near the input of said separatory channels or means or regions thereof; one or more bin arrays for the amplification and/or transfer of single polynucleotide molecules.

and further comprising one of:

one or more separate surfaces or surface regions not part of the surfaces of one or more separatory channels, detection probe arrays or membranes, for transfer of fractionated polynucleotide molecular species subsequent to separation or as separation is conducted; one or more tag-probe arrays for the receiving of fractionated polynucleotide molecular species during or after fractionation; one or more tag-probe arrays for receiving fractionated polynucleotide molecular species after fractionation of said polynucleotide molecular species by means of an array comprising fraction transfer grooves and two or more surface immobilized tag-probes and electrodes for lateral electroelution.

7. Method for the parallel loading of parallel capillary electrophoresis apparatus comprising the use of one or more one- or two-dimensional comb or comb-like structures to individual teeth of which individual polynucleotide samples are reversibly bound or immobilized.

8. Method for the parallel loading of parallel capillary electrophoresis apparatus according to claim 7 where said one or more one- or two-dimensional comb or comb-like structures further comprise on said individual teeth, or on the surface thereof or otherwise in communication with said surface, an oligonucleotide probe such that said one or more one- or two-dimensional comb or comb-like structures comprises on its surface an oligonucleotide probe array, in configurational register with the array of said teeth, to which said teeth individual polynucleotide samples are reversibly bound or immobilized by hybridization with predetermined stringency.

9. Method for the parallel loading of a parallel capillary electrophoresis apparatus according to claim 7 where said one or more one- or two-dimensional comb or comb-like structures are employed to transfer sample species from one of: polynucleotide molecules resulting from the random distribution of polynucleotide molecules into a bin array, such that many bins commonly contain exactly one polynucleotide molecule, followed by a step amplifying said polynucleotide molecule in the bins of said bin array; the random distribution of polynucleotide molecules, such that many bins commonly contain exactly one polynucleotide molecule, into two or more bin arrays followed by a step amplifying said polynucleotide molecules in said bin arrays and a step selectively transferring, on an array element by element basis those products produced in said amplifying step arising from a single molecule in a single element of said two or more bins into the corresponding element of a third bin array or the corresponding channel of a separatory channel array, such that each said corresponding element of said third bin array or each said channel of said separatory channel array receives material amplified from exactly one molecule.

10. Method for the parallel loading of a parallel capillary array comprising the use of one or more oligonucleotide probe arrays the elements of which occur in a spatial configuration complementary to that of the input ports of said parallel capillary array, to which sample mixture of plural polynucleotide

molecule species are bound by hybridization in a first demultiplexation step, after which said one or more oligonucleotide probe arrays are then juxtaposed to the said input ports of said parallel capillary array, followed by a denaturation step to release said plural polynucleotide molecule species such that only said polynucleotide molecule species bound by a single said element of said oligonucleotide probe arrays are released into the single juxtaposed said input ports of said parallel capillary array, in a second transfer step, for subsequent electrophoretic separation and detection of the polynucleotide molecular species thus separated.

11. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising the steps of:

(a) producing one or more lithographic plates, the surface of each comprising at least a first portion which is wetted by a reactant solution and a second portion which repels said reactant solution, and where said first portion or said second portion comprises and is defined by a polymeric coating which differs in composition from the underlying material of said lithographic plate;

(b) uniformly applying a first reactant solution to one said lithographic plate of step (a);

(c) juxtaposingly contacting said lithographic plate with applied reactant solution of step (b) to said solid surface, in proper register and alignment with said solid surface;

(d) performing any necessary chemical reactions necessary to the synthesis of the desired copolymers which need not be under spatial control to obtain the desired surface pattern on said solid surface;

(e) repeating steps (b) through (d) a sufficient number of times to obtain a surface decorated with a predetermined spatial pattern of two or more distinct copolymer sequences.

12. The method of claim **11** where said copolymer array comprises oligomers of composition selected from the group consisting of: DNA; RNA; peptide-nucleic acids; polypeptides; and, polypeptide mimics.

13. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising the steps of:

(a) producing one or more catalytic surfaces the surface configuration of which is known;

(b) reacting initiator molecules to a solid surface;

(c) uniformly applying a first reactant solution to one said catalytic relief of step (a);

(d) juxtaposingly contacting one of said catalytic surfaces, in proper register and alignment with, and relative position to said solid surface, and under conditions favorable to catalysis and the desired chemical reaction, to said solid surface;

(e) performing any necessary chemical reactions necessary to the synthesis of the desired copolymers which need not be under spatial control to obtain the desired surface pattern on said solid surface;

(f) repeating steps (b) through (f) a sufficient number of times to obtain a surface decorated with a predetermined spatial pattern of two or more distinct copolymer sequences.

14. The method of claim 13 where said copolymer array comprises oligomers of composition selected from the group consisting of: backbone-modified DNA; backbone-modified RNA; peptide-nucleic acids; polypeptides; and, polypeptide mimics.

15. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising according to claim 13 where said catalytic surfaces, which are of known configuration, are catalytic reliefs comprising predetermined surface patterns.

16. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising according to claim 13 where said catalytic surfaces are the surfaces of one or more atomic force microscope or scanning probe microscope tips.

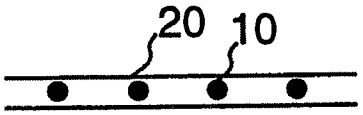


Fig. 1a

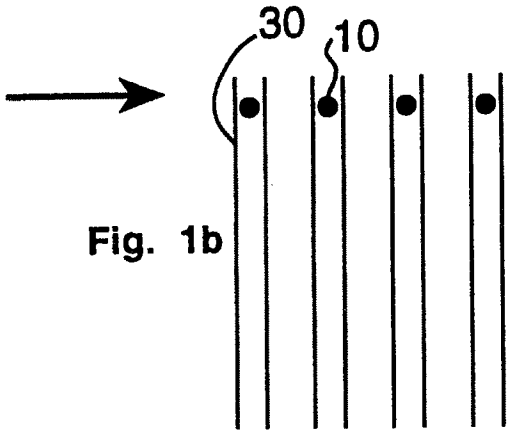


Fig. 1b

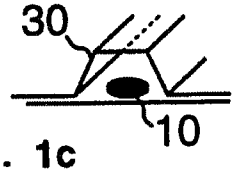


Fig. 1c

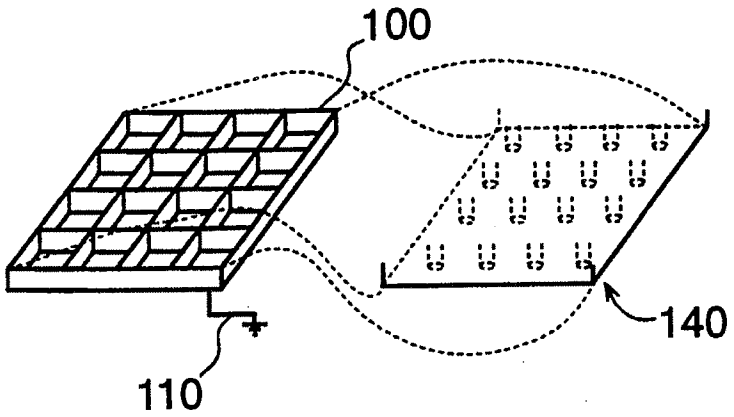


Fig. 2

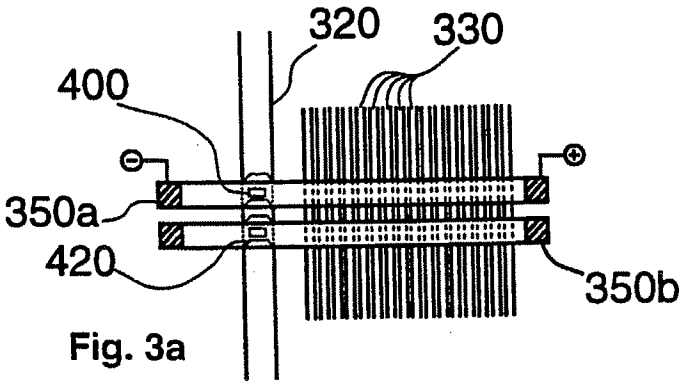


Fig. 3a

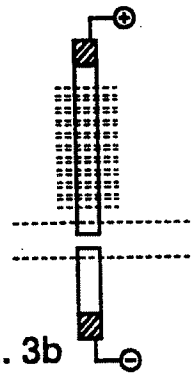


Fig. 3b

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/13313

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12Q 1/68; G01N 27/26; C07H 21/00
US CL :435/6; 204/182.8; 536/25.3
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6; 204/182.8; 536/25.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAPLUS, MEDLINE, EMBASE, BIOSIS, INPADOC
search terms: Multiplex, sequencing, parallel, polynucleotide, nucleic acid, DNA, array, oligonucleotide, synthesis

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,324,401 A (YEUNG ET AL.) 28 June 1994, columns 1-4, 10, and 16.	7-10
X ---- Y	LAGERKVIST et al. Manifold Sequencing: Efficient Processing of Large Sets of Sequencing Reactions. Proceedings of the National Academy of Sciences USA. March 1994, Vol. 91, pages 2245-2249, especially pages 2245 and 2247.	1-6 ---- 7-10
X	WO 94/05394 A1 (ARRIS PHARMACEUTICAL CORPORATION) 17 March 1994, pages 6-17.	11-16

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 05 DECEMBER 1996	Date of mailing of the international search report 24 DEC 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer AMY ATZEL Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/13313

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/13313

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-6, drawn to methods and devices for determining sequences.

Group II, claims 7-10, drawn to methods of parallel landing.

Group III, claims 11-16, drawn to methods of synthesis of patterned arrays.

The inventions listed as Groups I-III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Groups I and III do not pertain to capillary electrophoresis as does Group II; and Groups I and II do not require lithographic techniques as does Group III.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
G01N 33/50, 21/62, 21/77, 21/64

A1

(11) International Publication Number: **WO 99/18434**

(43) International Publication Date: 15 April 1999 (15.04.99)

(21) International Application Number: PCT/US98/21193

(22) International Filing Date: 6 October 1998 (06.10.98)

(30) Priority Data:
08/944,850 6 October 1997 (06.10.97) US

(71) Applicant *(for all designated States except US)*: TRUSTEES OF TUFTS COLLEGE [US/US]; Tufts University, Ballou Hall, Medford, MA 02155 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WALT, David, R. [US/US]; 4 Candlewick Close, Lexington, MA 02178 (US). DICKINSON, Todd, A. [US/US]; Apartment 1133, 3435 Lebon Drive, San Diego, CA 92122 (US).

(74) Agent: CREEHAN, R., Dennis; P.O. Box 750070, Arlington Heights, MA 02175–0070 (US).

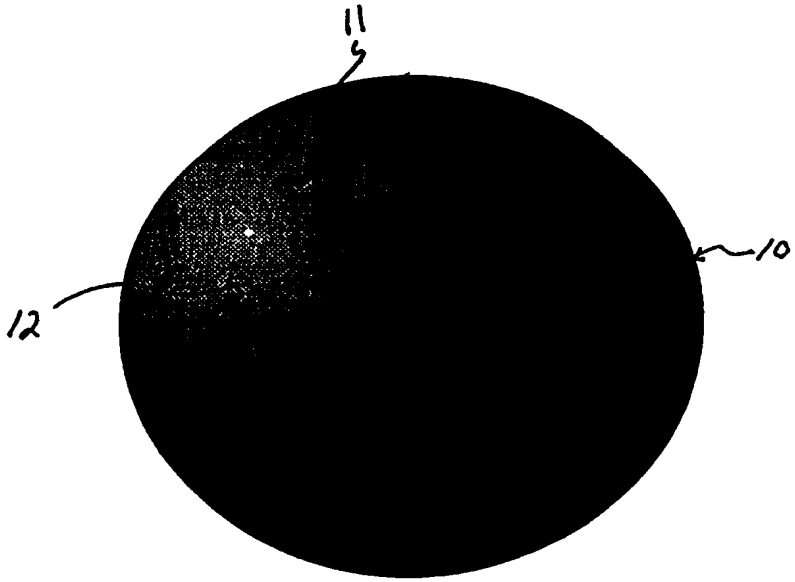
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SELF–ENCODING FIBER OPTIC SENSOR

(57) Abstract

A microsphere–based analytic chemistry system is disclosed in which self–encoding microspheres having distinct characteristic optical response signatures to specific target analytes may be mixed together while the ability is retained to identify the sensor type and location of each sensor in a random dispersion of large numbrs of such sensors in a sensor array using an optically interrogatable encoding scheme. An optical fiber bundle sensor is also disclosed in which individual microsphere sensors are disposed in microwells at a distal end of the fiber bundle and are optically coupled to discrete fibers or groups of fibers within the bundle. The identifies of the individual sensors in the array are self–encoded by exposing the array to a reference analyte while illuminating the array with excitation light energy. A single sensor array may carry thousands of discrete sensing elements whose combined signal provides for substantial improvements in sensor detection limits, response times and signal–to–noise ratios.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SELF-ENCODING FIBER OPTIC SENSOR

FIELD OF THE INVENTION

The present invention is generally concerned with chemical sensors, sensor arrays and sensing apparatus for the detection of gaseous and liquid analytes. More particularly, the invention is directed to optical chemical sensors and the detection and evaluation of optical data generated by sensing receptor units.

BACKGROUND OF THE INVENTION

The use of optical fibers and optical fiber strands in combination with light absorbing dyes for chemical analytical determinations has undergone rapid development, particularly within the last decade. The use of optical fibers for such purposes and techniques is described by Milanovich et al., "Novel Optical Fiber Techniques For Medical Application", Proceedings of the SPIE 28th Annual International Technical Symposium On Optics and Electro-Optics, Volume 494, 1980; Seitz, W.R., "Chemical Sensors Based On Immobilized Indicators and Fiber Optics" in *C.R.C. Critical Reviews In Analytical Chemistry,* Vol. 19, 1988, pp. 135-173; Wolfbeis, O.S., "Fiber Optical Fluorosensors In Analytical Chemistry" in *Molecular Luminescence Spectroscopy, Methods and Applications* (S. G. Schulman, editor), Wiley & Sons, New York (1988); Angel, S.M., *Spectroscopy* 2(4):38 (1987); Walt, et al., "Chemical Sensors and Microinstrumentation", *ACS Symposium Series,* Vol. 403, 1989, p. 252, and Wolfbeis, O.S., *Fiber Optic Chemical Sensors*, Ed. CRC Press, Boca Raton, FL, 1991, 2nd Volume.

When using an optical fiber in an *in vitro/in vivo* sensor, one or more light absorbing dyes are located near its distal end. Typically, light from an appropriate source is used to illuminate the dyes through the fiber's proximal end. The light propagates along the length of the optical fiber; and a portion of this propagated light exits the distal end and is absorbed by the dyes. The light absorbing dye may or may not be immobilized; may or may not be directly attached to the optical fiber itself; may or may not be suspended in a fluid sample containing one or more analytes of interest; and may or may not be retainable for subsequent use in a second optical determination.
Once the light has been absorbed by the dye, some light of varying wavelength and intensity returns, conveyed through either the same fiber or collection fiber(s) to a detection system where it is observed and measured. The interactions between the light conveyed by the optical fiber and the properties of the light absorbing dye provide an optical basis for both qualitative and quantitative determinations.

Of the many different classes of light absorbing dyes which conventionally are employed with bundles of fiber strands and optical fibers for different analytical purposes are those more common compositions that emit light after absorption termed "fluorophores" and those which absorb light and internally convert the absorbed light to heat, rather than emit it as light, termed "chromophores."

Fluorescence is a physical phenomenon based upon the ability of some molecules to absorb light (photons) at specified wavelengths and then emit light of a longer wavelength and at a lower energy. Substances able to fluoresce share a number of common characteristics: the ability to absorb light energy at one wavelength; reach an excited energy state; and subsequently emit light at another light wavelength. The absorption and fluorescence emission spectra are individual for each fluorophore and are often graphically represented as two separate curves that are slightly overlapping. The same fluorescence emission spectrum is generally observed irrespective of the wavelength of the exciting light and, accordingly, the wavelength and energy of the exciting light may be varied within limits; but the light emitted by the fluorophore will always provide the same emission spectrum. Finally, the strength of the fluorescence signal may be measured as the quantum yield of light emitted. The fluorescence quantum yield is the ratio of the number of photons emitted in comparison to the number of photons initially absorbed by the fluorophore. For more detailed information regarding each of these characteristics, the following references are recommended: Lakowicz, J. R., *Principles of Fluorescence Spectroscopy,* Plenum Press, New York, 1983; Freifelder, D., Physical Biochemistry, second edition, W. H. Freeman and Company, New York, 1982; "Molecular Luminescence Spectroscopy Methods and Applications: Part I" (S.G. Schulman, editor) in *Chemical Analysis,* vol. 77, Wiley & Sons, Inc., 1985; *The Theory of Luminescence,* Stepanov and Gribkovskii, Iliffe Books, Ltd., London, 1968.

Many of the recent improvements employing optical fiber sensors in both qualitative and quantitative analytical determinations concern the desirability of depositing and/or immobilizing various light absorbing dyes at the distal end of the optical fiber. In this manner, a variety of different optical fiber chemical sensors and methods have been reported for specific analytical determinations and applications such as pH measurement, oxygen detection, and carbon dioxide analyses. These developments are exemplified by the following publications: Freeman, et al., *Anal Chem.* 53:98 (1983); Lippitsch et al., *Anal. Chem. Acta.* 205:1, (1988); Wolfbeis et al., *Anal. Chem.* 60:2028 (1988); Jordan, et al., *Anal. Chem.* 59:437 (1987); Lubbers et al., *Sens. Actuators* 1983; Munkholm et al., *Talanta* 35:109 (1988); Munkholm et al., *Anal. Chem.* 58:1427 (1986); Seitz, W. R., *Anal. Chem.* 56:16A-34A (1984); Peterson, et al., *Anal. Chem.* 52:864 (1980): Saari, et al., *Anal. Chem.* 54:821 (1982); Saari, et al., *Anal. Chem.* 55:667 (1983); Zhujun et al., *Anal. Chem. Acta.* 160:47 (1984); Schwab, et al., *Anal. Chem.* 56:2199 (1984); Wolfbeis, O.S., "Fiber Optic Chemical Sensors", *Ed. CRC Press, Boca Raton, FL,* 1991, 2nd Volume; and Pantano, P., Walt, D.R., *Anal. Chem.*, 481A-487A, Vol. 67, (1995).

More recently, fiber optic sensors have been constructed that permit the use of multiple dyes with a single, discrete fiber optic bundle. U.S. Pat. Nos. 5,244,636 and 5,250,264 to Walt, *et al.* disclose systems for affixing multiple, different dyes on the distal end of the bundle, the teachings of each of these patents being incorporated herein by this reference. The disclosed configurations enable separate optical fibers of the bundle to optically access individual dyes. This avoids the problem of deconvolving the separate signals in the returning light from each dye, which arises when the signals from two or more dyes are combined, each dye being sensitive to a different analyte, and there is significant overlap in the dyes' emission spectra.

Most recently, fiber optic sensors have been employed in arrays of semi-selective chemical sensors and patttern recognition schemes to discriminate and quantify odors. Such approaches have been useful in implementing the principles of biological olfaction in the design of sensing devices or systems. In this field of biomimetry, various technologies have been applied to the sensor transduction mechanism. For example, surface acoustic wave, conducting polymer, metal oxide sensor field-effect transistor (MOSFET), piezo-electric, and quartz crystal microbalance sensor arrays have been pursued.

While such technologies provide inventive approaches utilizing a variety of physical and chemical phenomena to odor sensing, there are a number of limitations to these methods which restrict the efficacy of such devices. Firstly, element-to-element reproducibility both within a single array and between sensor arrays is typically unsatisfactory and thus requires recalibration and network retraining from sensor to sensor. Secondly, most of these methods have a relatively slow response time, frequently requiring several minutes to respond to the presence of an odor. Thirdly, such methods have relatively high detection limits and low sensitivity, typically not functioning at odor levels below 10ppm. Fourthly, devices which embody such technologies typically require a relatively large inherent size, thereby restricting miniaturization of the sensor array for use in remote sensing applications. Finally, construction of multi-sensor arrays by these methods is complex and involves expensive and tedious preparation and placement of individual sensors within a well-defined array.

Most recently, many of these shortcomings have been overcome through the application of fiber optic sensor arrays in an artificial nose sensor device and system. U.S. Pat. Nos. 5,320,814 and 5,512,490 to Walt,et al., the teachings of each of these patents being incorporated herein by reference, disclose a fiber optic array formed of heterogeneous, semi-selective thin films which function as sensing receptor units and are able to detect a variety of different analytes and ligands using spectral recognition patterns. This technology has been applied to a vapor-sensing system which utilizes arrays of polymer-dye combinations which coat the ends of select optical fibers in a fiber optic bundle. These developments are further described in Dickinson, et al, *Nature* 382:697 (1996) and White, et al, *Anal. Chem.* 68:2191 (1996).

An innovative feature of the four previously referenced patents to Walt, et al., was the placement of multiple chemical functionalities at the end of a single optical fiber bundle sensor. This configuration yielded an analytic chemistry sensor that could be remotely monitored via the typically small bundle. The drawback, however, was the difficulty in applying the various chemistries associated with the chemical functionalities at the sensor's end; the functionalities were built on the sensor's end in a step-wise serial fashion. This was a slow process, and in practice, only tens of functionalities could be applied.

U.S. Pat. Appl. No. 08/818,199 to Walt, et al, the teachings of which are incorporated herein by this reference, discloses the use of dye infiltrated polymer microspheres as a substitute for polymer-dye coating layers in optical fiber array sensors. With this approach, a fiber optic bundle serves as a substrate for dye-polymer microsphere array which contains a variety of microsphere bead sensors having different chemical and optical responses to the presence of target analytes. One innovative feature of this invention is in providing for a bead-based analytic chemistry system in which beads or microspheres carrying different chemical functionalities may be mixed together while retaining the ability to identify the functionality of each bead using an optically interrogatable encoding scheme. Additionally, this invention provides for an optical fiber bundle sensor in which the separate beads or microspheres may be optically coupled to discrete fibers or groups of fibers within the bundle. While the innovative features of this invention have separate applications, when implemented together, the invention provides for an optical fiber sensor that can support large numbers, thousands or more, of separate chemical sensor elements, which can be incorporated into a chemical sensor array and chemical analysis system. This approach provides for rapid fabrication and assembly of individual sensors and complex sensor arrays containing a multitude of discrete sensor types. The method also provides for a high degree of reproducibility and conformity within a batch of sensors and sensor arrays. Additional advantages are realized due to the ultrafine sizing available in microspheres. The overall size of the sensor array can be substantially reduced to submillimeter scale. This reduction in scale is particularly advantageous for remote sensing arrays.

While the method of applying microsphere sensor elements in chemical sensor arrays as taught in U.S. Pat. Appl. No. 08/818,199 to Walt, et al, has many innovative features, this method has certain limitations. The method requires a complex multi-step bead encoding process to identify the type and location of bead subpopulations used in the sensor array. Beads are encoded by employing combinations of fluorescent dyes in varying ratio. The choice of encoding dyes is limited to those dyes which emit light at different wavelengths upon exposure to excitation light energy. While combinations of dyes in different ratios provide for encoding subpopulations of beads, the number of dye ratios available for encoding beads with a given dye pair or combination is significantly limited due to crowding the emission spectrum from peak overlap. In addition, a separate reporting dye is necessary for obtaining a unique characteristic optical response signature for a target analyte. Thus, the

encoding dye choice is further limited by selecting dyes whose emission wavelengths do not overlap or interfere with the reporting dye which is uniquely responsive to the presence of an analyte.

Another limiting feature of this invention is that the process of encoding beads requires a series of measurements which calibrate and train the sensors and the sensor array. Encoding is initially accomplished by first illuminating the beads with excitation light energy and monitoring and recording the type and location of the specific bead subpopulation within the sensor array having a given encoding dye ratio. Next, the array is exposed to an analyte while illuminating the array with excitation light energy in the presence of a reporter dye. Those beads which are responsive to the analyte in the presence of the reporter dye are monitored and mapped on the sensor array. In addition, the characteristic optical response signature is stored in a library. This step is repeated for each analyte of interest in combination with a reporter dye. Once all bead subpopulations are encoded and their response characteristics monitored and recorded, the entire sensor array must be decoded for each analyte by indexing each sensor element with the stored optical response signature for each analyte. This process of decoding individual subpopulations of beads may be require additional steps when a large number of subpopulations are deployed in the array, thereby increasing the training time required for each array.

Other alternative approaches to bead encoding, utilizing molecular tagging, capillary gas chromatography and electron capture detection have been disclosed by Still, et al, *Acc. Chem. Res.* 29:155 (1996). However, such methods are limited in scope and have been applied only to a narrow class of bead materials having specific chemical functionality and molecular tags which are readily analyzable.

SUMMARY OF THE INVENTION

In general, the invention provides self-encoding analytic chemical sensor arrays comprising a substrate with a surface comprising discrete sites and a population of microspheres comprising at least a first and a second subpopulation, wherein each subpopulation comprises at least one reporter dye. The reporting dye has a first characteristic optical response signature when subjected to excitation light energy in the presence of a reference analyte, and the microspheres are distributed on the surface. The beads may further comprise a bioactive agent or, alternatively, a chemical functionality which interacts, associates, or binds with analytes to be detected.

In an additional aspect, the invention provides methods of detecting a target analyte in a sample comprising contacting the sample with an sensor array. The sensor array comprises a substrate with a surface comprising discrete sites and a population of microspheres. The microspheres comprise at least a first and a second subpopulation, each subpopulation comprising a bioactive agent and at least one reporter dye. The reporting dye has a first characteristic optical response signature when

subjected to excitation light energy in the presence of a reference analyte and the microspheres are distributed on the surface. The presence or absence (or quantity) of the analyte is then detected. The methods may further comprise identifying the location of each bioactive agent on said substrate by adding the reference analyte.

In a further aspect, the invention provides methods for reducing the signal-to-noise ratio in the characteristic optical response signature of a sensor array having a subpopulations of array elements. The methods comprise decoding the array so as to identify the location of each sensor element within each sensor subpopulation within the array and measuring the characteristic optical response signature of each sensor element in the array. The baseline of the optical response signature is then adjusted for each sensor element in said array, and the baseline-adjusted characteristic optical response signature of all sensor elements within each of the sensor subpopulations is summed. The characteristic optical response signature of each sensor subpopulation as a summation of said baseline-adjusted characteristic optical response signatures of all sensor elements within each of said subpopulations is then reported.

In an additional aspect, the invention provides methods for amplifying the characteristic optical response signature of a sensor array having subpopulations of array elements. The methods comprise decoding the array so as to identify the location of each sensor element within each sensor subpopulation within the array and measuring a characteristic optical response signature of each sensor element in the array. The baseline of the optical response signature for each sensor element in said array is then adjusted. The baseline-adjusted characteristic optical response signature of all sensor elements within each of the sensor subpopulations is then summed and the characteristic optical response signature of each sensor subpopulation as a summation of the baseline-adjusted characteristic optical response signatures of all sensor elements within each of the subpopulations is reported.

The above and other features of the invention including various novel details of construction and combinations of parts, and other advantages, will now be more particularly described with reference to the accompanying drawings and pointed out in the claims. It will be understood that the particular method and device embodying the invention are shown by way of illustration and not as a limitation of the invention. The principles and features of this invention may be employed in various and numerous embodiments without departing from the scope of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

In the accompanying drawings, reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale; emphasis has instead been placed upon illustrating the principles of the invention. Of the drawings:

Fig. 1 is a schematic diagram illustrating the self-encoding microsphere sensor according to the present invention;

Fig. 2 is a process flow diagram of the preparation, encoding and incorporation of microspheres into a sensor array of the present invention;

Figs. 3A and 3B is a schematic process diagram illustrating the preparation and placement of self-encoded microsphere subpopulations in fiber optic sensor array of the present invention;

Fig. 4 is a process flow diagram illustrating microwell formation in the fiber optic bundle and placement of the microspheres in the microwells according to the method of the present invention;

Figs. 5A and 5B are micrographs illustrating the microwells formed on the distal end of a fiber optic bundle and microspheres inserted in the microwell cavities;

Figs. 6A and 6B are micrographs showing the array of microspheres in their corresponding microwells prior to and subsequent to agitation by tapping and an air pulse, demonstrating the electrostatic binding of the microspheres in the microwell cavities;

Fig. 7 is a schematic diagram of the inventive fiber optic sensor and associated instrumentation and control system;

Fig. 8 is a schematic diagram illustrating the experimental apparatus used in the optical measurements of Examples 7 through 17;

Fig. 9 illustrates the characteristic optical response signature of porous silica beads infiltrated with Nile Red dye upon exposure to toluene vapor;

Fig. 10 illustrates the characteristic optical response signature of PMS beads infiltrated with Nile Red dye upon exposure to methanol vapor;

Figs. 11A and 11B illustrate the characteristic optical response signature of a PS802 coated porous silica bead infiltrated with Nile Red dye upon exposure to toluene and methanol vapor;

Figs. 12A and 12B illustrate the characteristic optical response signature of a PDPO coated porous silica beads infiltrated with Nile Red dye upon exposure to toluene and methanol vapor;

Fig. 13 illustrates the characteristic optical response signature of porous silica beads infiltrated with Nile Red dye upon exposure to ethyl acetate vapor;

Fig. 14 illustrates the innovation of optical response signal summing for reducing signal-to-noise ratios in Nile Red infiltrated PMS bead subpopulation measurements of methanol vapor;

Fig. 15 illustrates the innovation of optical response signal summing for signal enhancement in PMS bead subpopulation measurements of methanol vapor;

Fig. 16 compares the characteristic optical response signatures of two PS802 coated porous silica beads infiltrated with Nile Red dye upon exposure to toluene and methanol vapor;

Fig. 17 compares the characteristic optical response signatures to methanol vapor which are used for decoding Nile Red infiltrated porous silica and PMS bead subpopulations in a self-encoded fiber optic sensor array of the present invention;

Fig. 18 compares the characteristic optical response signatures of Nile Red infiltrated porous silica and PMS bead subpopulations to n-proponal vapor in a self-encoded fiber optic sensor array of the present invention;

Fig. 19 compares the characteristic optical response signatures of Nile Red infiltrated porous silica and PMS bead subpopulations to toluene vapor in a self-encoded fiber optic sensor array of the present invention; and

Fig. 20 compares the differences in bead swelling response of PS802 coated porous silica, poly methyl styrene, and poly methyl styrene/divinyl benzene bead subpopulations upon exposure to toluene vapor.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention provides an analytic chemistry system that comprises a self-encoding sensor array comprising a population of beads or microspheres on discrete locations on the surface of a substrate. Within the bead population are separate bead subpopulations, each of which provides a characteristic optical response signature when illuminated by excitation light energy in the presence of a reference analyte, which may in some cases be the target analyte. Although the subpopulations may be randomly mixed together, the identity and location of each bead is determined via a characteristic optical response signature when illuminated by excitation light energy in the presence of a reference analyte.

This allows the decoding of the array, i.e. the identification of the location of each subpopulation of beads on an array, to proceed very simply. In a preferred embodiment, the beads are encoded with

one or more reporter dyes that exhibit a characteristic, i.e. unique, optical response signature to a reference analyte, generally a fluid such as a vapor. Thus, in this embodiment, exposure of the entire array to a reference analyte will allow the identification of the location of each bead of each subpopulation. As a result, by comparing the response of the entire sensor array to a known analyte, the individual sensor elements of the array are conveniently decoded simultaneous in one simple measurement. The self-encoding feature of the present invention eliminates the need for a more complex, multi-step encoding system.

The sensor array can then be used to detect the presence of target analytes, for example when the beads also comprise bioactive agents such as oligonucleotides or, alternatively, a chemical functional group, by looking for changes in the optical signature of the beads upon interaction or binding with a the target analyte, for example a substantially complementary labelled oligonucleotide, or, alternatively, a target analyte of interest. As will be appreciated by those in the art, this may be done in a variety of ways, generally through the use of a change in an optical signal. This change can occur via many different mechanisms. A few examples include the binding of a dye-tagged analyte to the bead, the production of a dye species on or near the beads, the destruction of an existing dye species, a change in the optical signature upon analyte interaction with dye on bead, or any other optical interrogatable event. Thus, once the location of each species of an oligonucleotide probe, or alternatively, a bead having a particular chemical functionality, has been identified, the array can then be used to detect the presence of unknowns that will preferably specifically associate with the bioactive agents or chemical functionality on the beads.

In an alternate preferred embodiment, when the target analyte is not labeled, the optical response of each element in the array can be compared to a library of characteristic optical response signatures for its corresponding bead subpopulation type, where the characteristic optical response signature to various analytes has been previously measured and recorded, and either the identity of the unknown can be determined or the sensor array can be trained to associate the measured response with a particular analyte which is then added to the library of response signatures.

The present invention overcomes certain limitations of the current art by embodying the innovation of a self-encoding sensor array wherein a characteristic optical response signature is produced by the interaction of specific bead subpopulation compositions with a reporter dye. In the self-encoding sensor array of the present invention, the response signal to a target analyte serves both as a response signature for the target analyte and as the encoding signal for the entire sensor array and subpopulations within the array. The decoding of the array is thus accomplished in a one-step process during the array response measurement of a target analyte and utilizes the very response which is used to identify the target analyte. The bead encoding is thus incorporated into the array by the nature of the bead subpopulation responses to target analytes.

In the present invention, each bead-dye combination of a subpopulation has a characteristic optical response signature when exposed to a given fluid, such as a liquid or a vapor. The self-encoding concept is provided by the unique response characteristics of the dye in combination with a specific bead matrix material. Thus the bead subpopulations which are randomly dispersed in a sensor array can be rapidly identified and located after placement in the array simply by exposing the sensor array to a known test fluid and matching the resulting optical response signature to those obtained for each bead subpopulation. With this approach, the beads are self encoding and the response characteristics of the entire sensor array are rapidly determined and stored for measurement of a target analyte. The method of the present invention is particularly useful in applications of sensor arrays containing thousands of sensors having distinctive optical response signature characteristics.

An additional benefit of the present invention is that it allows the synthesis of the bioactive agents (i.e. compounds such as nucleic acids and antibodies) to be separated from their placement on an array, i.e. the bioactive agents may be synthesized on the beads, and then the beads are randomly distributed on a patterned surface. Since the beads are self-encoded by having dyes present that have known responses to a reference analyte, this means that the array can later be "decoded", i.e. after the array is made, a correlation of the location of an individual site on the array with the bead or bioactive agent at that particular site can be made. This means that the beads may be randomly distributed on the array, a fast and inexpensive process as compared to either the *in situ* synthesis or spotting techniques of the prior art. Once the array is loaded with the beads, the array can be decoded, or can be used, with full or partial decoding occuring after testing, as is more fully outlined below.

In a preferred embodiment of the present invention, ultra-fine, porous microbeads or microspheres are utilized as individual sensors. The utilization of porous micron-scale sensors provides for improved sensor response and sensitivity. The reduction in sensor dimension substantially reduces the diffusion length and time for analyte interaction with individual sensors and significantly shortens the sensor response time, while simultaneously enhancing sensor sensitivity and lowering detection limits.

In another preferred embodiment of the present invention, the sensor array is comprised of subpopulations of beads or microspheres which are disposed on a distal end of an optical fiber bundle wherein the separate beads or microspheres may be optically coupled to discrete fibers or groups of fibers within the bundle. Since typically, such fiber optic bundles comprise thousands of discrete fibers, the present invention thus provides for an optical fiber sensor which can support a large number, thousands or more, of sensor array elements of distinct and varying subpopulations each having a characteristic optical response signature when exposed to an analyte while being illuminated by excitation light energy.

In one preferred embodiment, the distal end of a fiber optic bundle substrate is chemically etched so as to create a cavity or micro-well at the end of a discrete fiber. In the preferred embodiment, each one of the beads is located within separate microwells formed at terminal ends of optical fibers of the bundle. These microwells are formed by anisotropic etching of the cores of the optical fibers with respect to the cladding. The resultant etched cavity is dimensioned for accommodating an individual microbead sensor and for providing optical coupling of the individual bead sensor with the discrete optical fiber in the fiber bundle. Since typical fiber optic bundles contain thousands of discrete fibers, this embodiment provides for the individual optical coupling of thousands of sensors in a sensor array, thereby providing for a large number of independent sensor measurements for each bead subpopulation within the array.

Due to both the large number of bead sensor subpopulations available and the correspondingly large number of sensor elements within each subpopulation, a significant innovation of the present invention is in providing for thousands of independent sensor response measurements in a single sensor array. This enables another significant innovation of the present invention by providing for the summing and amplification of the characteristic optical response signatures of multiple independent measurements taken from sensor beads within each sensor array bead subpopulation. This approach directly mimics the actual behavior of the human olfactory where the combined signals from thousands of receptor cells in each of grouping of nearly a thousand different receptor cell types found in the epithelium layer, none of which are particularly sensitive in themselves, lead to a highly amplified sensory response to odors [see J.S. Kauer, *Trends Neurosci.* 14:79-95(1991)].

The present invention thus embodies the evolutionary scent amplification process found in the human olfactory system in order to significantly enhance sensor array sensitivity to analytes by summing the low-level responses of a large number of sensor array elements. By summing the responses from several beads at low vapor concentrations, a substantial improvement in signal-to-noise ratios is achieved, exceeding a factor of ten or more. This innovation has led to reducing the detection limit of the sensor array by over an order of magnitude. The enhancement in sensitivity provided by the sensor array of the present invention is generally known to be directly proportional to the square root of the number of independent sensor bead responses available for summing. With such enhancements, detection limits approaching parts per billion are achievable.

In preferred embodiments, the sensor beads are self-encoded using a reporter dye that is preferably infiltrated or entrapped within the beads. The reporter dye may be a chromophore or phosphor but is preferably a fluorescent dye, which due to characteristically strong optical signals provide a good signal-to-noise ratio for decoding. Although not necessary, the self-encoding can also be accomplished by utilizing the ratios of two or more reporting dyes having characteristic and discrete emission peaks and measuring the peak intensity ratios upon illumination with excitation light energy.

According to another embodiment, the invention also concerns a chemical sensor array designed with a predetermined chemical specificity. In this embodiment, additional chemical functionality can be incorporated into each sensor subpopulation by attaching a desired moiety to the surfaces of the beads. In another embodiment, the sensor array has a population of beads carrying chemical functionality at, on or near, a distal end of the bundle. The ability to monitor optical signature changes associated with individual or multiple beads interacting with a target analyte is provided by optically coupling those signature changes into separate optical fibers or groups of fibers of a fiber optic bundle for transmission to the proximal end where analysis is performed either manually, by the user, or automatically, using image processing techniques.

Although each sensor is different insofar that it has a different distribution of the subpopulations of beads within its microwells, only those beads that exhibit a positive optical response or signature change to a target analyte of interest need to be decoded. Therefore, the burden is placed on the analysis rather than on sensor manufacture. Moreover, since the beads and fibers in the array can be monodisperse, the fluorescent regions arising from signal generation are extremely uniform and can be analyzed automatically using commercially available microscopy analysis software. Such image processing software is capable of defining different spectral regions automatically and counting the number of segments within each region in several seconds.

Accordingly, the present invention provides array compositions comprising at least a first substrate with a surface comprising individual sites. By "array" herein is meant a plurality of bioactive agents, or alternatively, chemical functionalities, in an array format; the size of the array will depend on the composition and end use of the array. Arrays containing from about 2 different bioactive agents (i.e. different beads) to many millions can be made, with very large fiber optic arrays being possible. Similar arrays comprising other chemical functionalities may also be constructed. Generally, the array will comprise from two to as many as a billion or more bioactive agents or chemical functionalities, depending on the size of the beads and the substrate, as well as the end use of the array, thus very high density, high density, moderate density, low density and very low density arrays may be made. Preferred ranges for very high density arrays are from about 10,000,000 to about 2,000,000,000 elements, with from about 100,000,000 to about 1,000,000,000 being preferred. High density arrays range about 100,000 to about 10,000,000 elements, with from about 1,000,000 to about 5,000,000 being particularly preferred. Moderate density arrays range from about 10,000 to about 50,000 elements being particularly preferred, and from about 20,000 to about 30,000 being especially preferred. Low density arrays are generally less than 10,000 elements, with from about 1,000 to about 5,000 being preferred. Very low density arrays are less than 1,000 elements, with from about 10 to about 1000 being preferred, and from about 100 to about 500 being particularly preferred. In some embodiments, the compositions of the invention may not be in array format; that is, for some embodiments, compositions comprising a single bioactive agent may be made as well. In addition, in

some arrays, multiple substrates may be used, either of different or identical compositions. Thus for example, large arrays may comprise a plurality of smaller substrates.

In addition, one advantage of the present compositions is that particularly through the use of fiber optic technology, extremely high density arrays can be made. Thus for example, because beads of 200 nm can be used, and very small fibers are known, it is possible to have as many as 250,000 different fibers and beads in a 1 mm^2 fiber optic bundle, with densities of greater than 15,000,000 individual beads and fibers per 0.5 cm^2 obtainable.

The compositions comprise a substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant any material that can be modified to contain discrete individual sites appropriate for the attachment or association of beads and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, TeflonJ, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, optical fiber bundles, and a variety of other polymers. In general, the substrates allow optical detection and do not appreciably fluorescese.

Generally the substrate is flat or planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well; for example, three dimensional configurations can be used, for example by embedding the beads in a porous block of plastic that allows sample access to the beads and using a confocal microscope for detection. Similarly, the beads may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Preferred substrates include optical fiber bundles as discussed below, and flat planar substrates such as glass, polystyrene and other plastics and acrylics.

At least one surface of the substrate is modified to contain discrete, individual sites for later association of microspheres. These sites may comprise physically altered sites, i.e. physical configurations such as wells or small depressions in the substrate that can retain the beads, such that a microsphere can rest in the well, or the use of other forces (magnetic or compressive), or chemically altered or active sites, such as chemically functionalized sites, electrostatically altered sites, hydrophobically/ hydrophilically functionalized sites, spots of adhesive, etc.

The sites may be a pattern, i.e. a regular design or configuration, or randomly distributed. A preferred embodiment utilizes a regular pattern of sites such that the sites may be addressed in the X-Y coordinate plane. "Pattern" in this sense includes a repeating unit cell, preferably one that allows a high density of beads on the substrate. However, it should be noted that these sites may not be

discrete sites. That is, it is possible to use a uniform surface of adhesive or chemical functionalities, for example, that allows the attachment of beads at any position. That is, the surface of the substrate is modified to allow attachment of the microspheres at individual sites, whether or not those sites are contiguous or non-contiguous with other sites. Thus, the surface of the substrate may be modified such that discrete sites are formed that can only have a single associated bead, or alternatively, the surface of the substrate is modified and beads may go down anywhere, but they end up at discrete sites.

In a preferred embodiment, the surface of the substrate is modified to contain wells, i.e. depressions in the surface of the substrate. This may be done as is generally known in the art using a variety of techniques, including, but not limited to, photolithography, stamping techniques, molding techniques and microetching techniques. As will be appreciated by those in the art, the technique used will depend on the composition and shape of the substrate.

In a preferred embodiment, physical alterations are made in a surface of the substrate to produce the sites. In a preferred embodiment, the substrate is a fiber optic bundle and the surface of the substrate is a terminal end of the fiber bundle. In this embodiment, wells are made in a terminal or distal end of a fiber optic bundle comprising individual fibers. In this embodiment, the cores of the individual fibers are etched, with respect to the cladding, such that small wells or depressions are formed at one end of the fibers. The required depth of the wells will depend on the size of the beads to be added to the wells.

Generally in this embodiment, the microspheres are non-covalently associated in the wells, although the wells may additionally be chemically functionalized as is generally described below, cross-linking agents may be used, or a physical barrier may be used, i.e. a film or membrane over the beads.

In a preferred embodiment, the surface of the substrate is modified to contain chemically modified sites, that can be used to attach, either covalently or non-covalently, the microspheres of the invention to the discrete sites or locations on the substrate. "Chemically modified sites" in this context includes, but is not limited to, the addition of a pattern of chemical functional groups including amino groups, carboxy groups, oxo groups and thiol groups, that can be used to covalently attach microspheres, which generally also contain corresponding reactive functional groups; the addition of a pattern of adhesive that can be used to bind the microspheres (either by prior chemical functionalization for the addition of the adhesive or direct addition of the adhesive); the addition of a pattern of charged groups (similar to the chemical functionalities) for the electrostatic attachment of the microspheres, i.e. when the microspheres comprise charged groups opposite to the sites; the addition of a pattern of chemical functional groups that renders the sites differentially hydrophobic or hydrophilic, such that the addition of similarly hydrophobic or hydrophilic microspheres under suitable experimental conditions will result in association of the microspheres to the sites on the basis of hydroaffinity. For example, the use of

hydrophobic sites with hydrophobic beads, in an aqueous system, drives the association of the beads preferentially onto the sites. As outlined above, "pattern" in this sense includes the use of a uniform treatment of the surface to allow attachment of the beads at discrete sites, as well as treatment of the surface resulting in discrete sites. As will be appreciated by those in the art, this may be accomplished in a variety of ways.

The compositions of the invention further comprise a population of microspheres. By "population" herein is meant a plurality of beads as outlined above for arrays. Within the population are separate subpopulations, which can be a single microsphere or multiple identical microspheres. That is, in some embodiments, as is more fully outlined below, the array may contain only a single bead for each bioactive agent; preferred embodiments utilize a plurality of beads of each type.

By "microspheres" or "beads" or "particles" or grammatical equivalents herein is meant small discrete particles. The composition of the beads will vary, depending on the class of bioactive agent and the method of synthesis. Suitable bead compositions include those used in peptide, nucleic acid and organic moiety synthesis, including, but not limited to, plastics, ceramics, glass, polystyrene, methylstyrene, acrylic polymers, paramagnetic materials, thoria sol, carbon graphited, titanium dioxide, latex or cross-linked dextrans such as Sepharose, cellulose, nylon, cross-linked micelles and teflon may all be used.

Synthetic beads may be fabricated by polymerizing or copolymerizing a variety of condensation or vinyl precursor monomers or by way of combinatorial polymer synthesis. Such polymers can be further modified by the addition of plasticizers, such as tritolyl phosphate (TTP), triphenyl phospate (TTP) or dibutyl phthalate (DBP). Particularly useful dye-encoding bead candidates for use in sensor array subpopulations are polymer and copolymer materials which exhibit either a characteristic swelling upon exposure to various vapor analytes, a characteristic polarity difference due to their chemical structure, or a characteristic chemical adsorption response with various vapor analytes. In prescreening candidate polymers as bead materials and evaluating candidates based on desirable swelling, polarity and adsorption characteristics, two particularly useful references are: R.A. McGill, et al., *Chemtech*, September 24, 1996,p27-37 and J.W. Grate, et al., *Anal. Chem.* 68:913-7 (1996).

A variety of bead chemistries may be utilized in fabricating a wide diversity of sensor bead subpopulations. For example, the following compositions have been found to be particularly useful as candidate bead materials: silica, poly(ethylene glycol), polycaprolactone, poly(1,4-butylene adipate), PDPO [poly(2,6-dimethyl-1,4-phenyleneoxide)], PS078.5 [triethoxysilyl-modified polybutadiene (50% in toluene)], PS078.8 [diethoxymethylsilyl-modified polybutadiene in toluene], CPS2067 [acryloxypropylmethyl-cyclosiloxane], PS802 [(80-85%) dimethyl-(15-20%) (acryloxypropyl) methylsiloxane copolymer], PS901.5 poly(acryloxypropyl-methyl)siloxane], PS851 [(97-98%) dimethyl-(2-3%) (methacryloxypropyl) methylsiloxane copolymer], PABS [poly(acrylonitrile-butadiene-styrene)],

poly(methyl methacrylate), poly(styrene-acrylonitrile 75:25), acryloxypropylmethylsiloxane-dimethylsiloxane copolymer, methylstyrene, polystyrene, acrylic polymers, and poly(methyl styrene/divinyl benzene). Other adsorbents, such as commercially available silica beads adapted with a variety of bonded phases for use in phenomenex columns, such as beads comprising C_8, C_{18} and phenyl hexyl, are useful as sensor bead matrices. Inorganic materials such as alumina and zeolites may also be utilized. Other polymers and copolymers having distinguishable and suitable swelling behavior, polarity and chemical adsorption characteristics are also anticipated as likely bead candidate materials. Particularly useful bead candidate materials include the polymers, copolymers, and polymerized monomers listed in Table 7, Table 8 and Table 10 of U.S. Patent 5,512,490 to Walt, et al, which are herein incorporated by reference. In alternative embodiments, any synthesized or commercially available bead materials may be further modified by applying either a surface treatment or coating to modify the characteristic optical response signature. For example, where porous silica beads are utilized, N-octadecyltriethyoxysilane or 3-(trimethoxysilyl)propyl methacrylate may be applied as a silanization treatment. In general, *"Microsphere Detection Guide"* from Bangs Laboratories, Fishers IN is a helpful guide in selection of bead materials.

The choice of subpopulations used to form the sensor array elements in a particular sensor array is primarily determined based on the analytical purposes of the sensor and the specific analytes which are targeted for detection. Typically, bead subpopulations are selected based on distinguishable differences in their characteristic optical response signatures when illuminated by excitation light energy in the presence of a target analyte. In fabricating self-encoding sensor arrays, bead subpopulations are selected which have characteristic optical response signatures when infiltrated with a reporting dye and illuminated by excitation light energy in the presence of both a reference analyte and target analyte. Thus, preferred bead materials for the sensor array are preselected based on either physical or chemical differences in bead subpopulations which produce a characteristic optical response signature in the presence of the analyte when illuminated by excitation light energy.

Features such bead material polarity, chemical structure, chemical functionality, bead surface area, bead pore size, bead swelling characteristics, or chemical adsorption behavior , either separately or in combination, contribute to the characteristic optical response signature of a given bead subpopulation. In one embodiment, bead materials which are permeable or semi-permeable to fluids including vapors and liquid analytes are preferred. In another embodiment, bead materials that swell upon contact with fluids such as vapor or liquid analytes are preferred. In general, bead materials which have unique polarity, structure, pore size, surface area, fuctionality or adsorption characteristics are particularly useful for sensor bead matrices of the present invention.

The microspheres comprise a reporting dye that, in combination with the characteristic bead matrix material, provides an optical response signature that can be used to identify the bead, and thus the attached bioactive agent, upon exposure to a reference analyte. That is, each subpopulation of

microspheres (i.e. each sensor element) comprises a unique optical response signature or optical tag, that can be used to identify the unique bioactive agent or chemical functionality of that subpopulation of microspheres; a bead comprising the unique optical response signature may be distinguished from beads at other locations with different optical response signatures. As is outlined herein, each bioactive agent will have an associated unique optical response signature such that any microspheres comprising that bioactive agent will be identifiable on the basis of the signature upon exposure to a reference analyte or fluid. As is more fully outlined below, it is possible to reuse or duplicate optical response signatures within an array, for example, when another level of identification is used, for example when beads of different sizes are used, or when the array is loaded sequentially with different batches of beads.

The selection of chemical dye indicators is equally important to the design of a fiber optic sensor array system of the present invention. In the preferred embodiment (see Fig. 1), at least one dye 11 is incorporated into the microsphere 10. In the preferred embodiment, this dye 11 acts as both an encoding dye, for identifying the bead subpopulation location in the sensor array, and a reporting dye, for detecting a target analyte of interest. In an alternative embodiment, two or more dyes may be utilized as encoding-reporter dyes. In a preferred embodiment, at least one dye is used solely as an encoding dye and a separate dye is added during analysis as a reporting dye. In one embodiment, where two or more encoding dyes are used, the ratio of peak intensities for dye pairs may be used for encoding the bead subpopulation and a separate reporter dye may be added during analysis. In an alternative embodiment, conjugated dyes, such as acrlyoyl fluorescein and others, may be utilized where it is desirable to incorporate the dye directly into a synthesized polymer or copolymer bead material.

While the reporter dye 11 may be either a chromophore-type or a fluorophore-type, a fluorescent dye is preferred because the strength of the fluorescent signal provides a better signal-to-noise ratio when decoding. In the most preferred embodiment, polarity-sensitive dyes or solvatochromic dyes are utilized. Solvatochromic dyes are dyes whose absorption or emission spectra are sensitive to and altered by the polarity of their surrounding environment. Typically, these dyes exhibit a shift in peak emission wavelength due to a change in local polarity. Polarity changes which cause such wavelength shifts can be introduced by the bead matrix used for a particular sensor bead subpopulation or, the presence of a target analyte. The change in polarity creates a characteristic optical response signature which is useful for both encoding subpopulations of bead types and for detecting specific target analytes. One preferred solvatochromic dye, Nile Red (Eastman Kodak, Rochester,NY), exhibits large shifts in its emission wavelength peak with changes in the local environment polarity. In addition, Nile Red is soluble in a wide range of solvents, is photochemically stable, and has a relatively strong fluorescence peak. Additional dyes which are conventionally known in the art and may be used as dyes in the present invention are those found in U.S. Patent 5,512,490 to Walt, et al., of which Table 3, Table 4, Table 5, Table 6 and Table 11 are incorporated herein by reference.

Different subpopulations of bead sensing elements can be fabricated for the sensor array of the present invention by immobilizing Nile Red in polymer matrices of varying composition. By incorporating Nile Red in bead subpopulations made from different polymer matrices of varying polarity, hydophobicity, pore size, flexibility and swelling tendency, unique subpopulations of sensor beads are produced that react differently with molecules of individual fluids, giving rise to different fluorescence responses when exposed to organic fluids. This results in each bead subpopulation having a characteristic optical response signature when exposed to a variety of analytes.

In a preferred embodiment, the dyes are covalently attached to the surface of the beads. This may be done as is generally outlined below for the attachment of the bioactive agents, using functional groups on the surface of the beads. As will be appreciated by those in the art, these attachments are done to minimize the effect on the dye.

In a preferred embodiment, the dyes are non-covalently associated with the beads, generally by entrapping the dyes in the bead matrix or pores of the beads.

In one embodiment, the dyes are added to the bioactive agent, rather than the beads, although this is generally not preferred.

Fig. 2 is a process diagram illustrating the preparation of the sensor bead subpopulations and sensor bead array. In step 50, suspensions of the various bead subpopulations are individually prepared from either commercial bead materials or synthesized bead materials which have been made from preferred polymeric materials. In this step, the beads may be prewashed, surface treated with a coupling agent, such as a silanizing solution as used in Example 2 and Example 3, or treated with a plasticizer, such as TTP, TPP or DBP as used in Example 6. In preparing the bead subpopulations, each bead grouping is typically dispersed in an appropriate solvent which may comprise additions of surfactants or dispersants to enhance dispersion. For example, Tween 20 (J.T.Baker, Cleveland, OH), a polyoxyethylenesorbitan monolaurate, has been found to be particularly useful as a surfactant.

A dye solution is prepared 51 for tagging or encoding each of the bead subpopulations for subsequent identification and indexing subpopulations in the sensor array in a later decoding step. In the most preferred embodiment, a single dye serves both as a sensor bead subpopulation encoding dye and as an analyte reporting dye that is used to detect the presence of a target analyte. In another embodiment, the dye serves solely to encode the sensor bead subpopulation and an additional dye is used as a reporter dye for detection of a target analyte. In one embodiment, two or more dyes may be incorporated into the bead subpopulation and the peak intensity ratios of dye pairs may be used for encoding the sensor bead subpopulation. Typically, a single solvatochromic dye is used as both the encoding dye and reporting dye. In a preferred embodiment, Nile Red dye (Aldrich, Milwaukee, WI) is

used. For incorporating dye into each bead subpopulation, suspensions of the beads prepared in step 50 are mixed in step 52 with dye solutions prepared in step 51. Preferably, in step 52, the beads or microspheres are placed in a dye solution comprising dye dissolved in an organic solvent that will swell the microspheres. In step 54, the beads are washed, centrifuged or filtered to remove excess dye. The microspheres are typically washed in water, methanol, or any suitable solvent that does not swell the microspheres, but in which the dyes are still soluble. This allows the residual dye to be rinsed off without rinsing the dye out of the microspheres. In an alternative embodiment, a chemical moiety or functional group may be attached to the bead surface following removal of excess dye.

The beads need not be spherical; irregular particles may be used. While both porous and non-porous beads may be utilized, porous beads are preferred for infiltrating the reporter dye and enhancing the responsivity and sensitivity of the microsphere sensor due to an increase in surface area for attachment of the reporter dye, bioactive agents, etc. The bead sizes range from nanometers, i.e. 100 nm, to millimeters, i.e. 1 mm, with beads from about 0.2 micron to about 200 microns being preferred, and from about 0.5 to about 5 micron being particularly preferred, although in some embodiments smaller beads may be used. Fig. 1 illustrates the construction of a typical bead or microsphere sensor 10 comprising a reporting dye 11 entrapped within bead pores 12.

It should be noted that a key component of the invention is the use of a substrate/bead pairing that allows the association or attachment of the beads at discrete sites on the surface of the substrate, such that the beads do not move during the course of the assay.

The present invention embodies a bead-based analytical chemistry system in which beads or microspheres are fabricated from various inorganic or organic materials wherein each material can be identified by a characteristic temporal optical response signature which enables verifying both the identity and location of a particular bead within a sensor array upon exposure to a reference analyte while illuminating with excitation light energy. The invention provides for utilization of any source of excitation light energy and is not limited to a specific wavelength. The principal requirement of the excitation light is that it produces emitted light of a characteristic wavelength upon illumination of a reporter dye associated with a given bead composition.

In a preferred embodiment, the microspheres further comprise a bioactive agent. By "candidate bioactive agent" or "bioactive agent" or "chemical functionality" or "binding ligand" herein is meant as used herein describes any molecule, *e.g.*, protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc. which can be attached to the microspheres of the invention. It should be understood that the compositions of the invention have two primary uses. In a preferred embodiment, as is more fully outlined below, the compositions are used to detect the presence of a particular target analyte; for example, the presence or absence of a particular nucleotide sequence or a particular protein, such as an enzyme, an antibody or an antigen. In an alternate preferred

embodiment, the compositions are used to screen bioactive agents, i.e. drug candidates, for binding to a particular target analyte.

Bioactive agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Bioactive agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The bioactive agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Bioactive agents are also found among biomolecules including peptides, nucleic acids, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are nucleic acids and proteins.

Bioactive agents can be obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification and/or amidification to produce structural analogs.

In a preferred embodiment, the bioactive agents are proteins. By "protein" herein is meant at least two covalently attached amino acids, which includes proteins, polypeptides, oligopeptides and peptides. The protein may be made up of naturally occurring amino acids and peptide bonds, or synthetic peptidomimetic structures. Thus "amino acid", or "peptide residue", as used herein means both naturally occurring and synthetic amino acids. For example, homo-phenylalanine, citrulline and norleucine are considered amino acids for the purposes of the invention. The side chains may be in either the (R) or the (S) configuration. In the preferred embodiment, the amino acids are in the (S) or L-configuration. If non-naturally occurring side chains are used, non-amino acid substituents may be used, for example to prevent or retard *in vivo* degradations.

In one preferred embodiment, the bioactive agents are naturally occurring proteins or fragments of naturally occuring proteins. Thus, for example, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of procaryotic and eukaryotic proteins may be made for screening in the systems described herein. Particularly

preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred.

In a preferred embodiment, the bioactive agents are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized bioactive proteinaceous agents.

In a preferred embodiment, a library of bioactive agents are used. The library should provide a sufficiently structurally diverse population of bioactive agents to effect a probabilistically sufficient range of binding to target analytes. Accordingly, an interaction library must be large enough so that at least one of its members will have a structure that gives it affinity for the target analyte. Although it is difficult to gauge the required absolute size of an interaction library, nature provides a hint with the immune response: a diversity of 10^7-10^8 different antibodies provides at least one combination with sufficient affinity to interact with most potential antigens faced by an organism. Published in vitro selection techniques have also shown that a library size of 10^7 to 10^8 is sufficient to find structures with affinity for the target. Thus, in a preferred embodiment, at least 10^6, preferably at least 10^7, more preferably at least 10^8 and most preferably at least 10^9 different bioactive agents are simultaneously analyzed in the subject methods. Preferred methods maximize library size and diversity.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, for example, of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

In a preferred embodiment, the bioactive agents are nucleic acids (generally called "probe nucleic acids" or "candidate probes" herein). By "nucleic acid" or "oligonucleotide" or grammatical equivalents herein means at least two nucleotides covalently linked together. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, as outlined below, nucleic acid analogs are included that may have alternate backbones, comprising, for example,

phosphoramide (Beaucage, *et al.*, Tetrahedron, **49**(10):1925 (1993) and references therein; Letsinger, J. Org. Chem., **35**:3800 (1970); Sprinzl, *et al.*, Eur. J. Biochem., **81**:579 (1977); Letsinger, *et al.*, Nucl. Acids Res., **14**:3487 (1986); Sawai, *et al.*, Chem. Lett., 805 (1984), Letsinger, *et al.*, J. Am. Chem. Soc., **110**:4470 (1988); and Pauwels, *et al.*, Chemica Scripta, **26**:141 (1986)), phosphorothioate (Mag, *et al.*, Nucleic Acids Res., **19**:1437 (1991); and U.S. Patent No. 5,644,048), phosphorodithioate (Briu, *et al.*, J. Am. Chem. Soc., **111**:2321 (1989)), O-methylphophoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm, J. Am. Chem. Soc., **114**:1895 (1992); Meier, *et al.*, Chem. Int. Ed. Engl., **31**:1008 (1992); Nielsen, Nature, **365**:566 (1993); Carlsson, *et al.*, Nature, **380**:207 (1996), all of which are incorporated by reference)). Other analog nucleic acids include those with positive backbones (Denpcy, *et al.*, Proc. Natl. Acad. Sci. USA, **92**:6097 (1995)); non-ionic backbones (U.S. Patent Nos. 5,386,023; 5,637,684; 5,602,240; 5,216,141; and 4,469,863; Kiedrowshi, *et al.*, Angew. Chem. Intl. Ed. English, **30**:423 (1991); Letsinger, *et al.*, J. Am. Chem. Soc., **110**:4470 (1988); Letsinger, *et al.*, Nucleosides & Nucleotides, **13**:1597 (1994); Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker, *et al.*, Bioorganic & Medicinal Chem. Lett., **4**:395 (1994); Jeffs, *et al.*, J. Biomolecular NMR, **34**:17 (1994); Tetrahedron Lett., **37**:743 (1996)) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within the definition of nucleic acids (see Jenkins, *et al.*, Chem. Soc. Rev., (1995) pp. 169-176). Several nucleic acid analogs are described in Rawls, C & E News, June 2, 1997, page 35. All of these references are hereby expressly incorporated by reference. These modifications of the ribose-phosphate backbone may be done to facilitate the addition of additional moieties such as labels, or to increase the stability and half-life of such molecules in physiological environments. In addition, mixtures of naturally occurring nucleic acids and analogs can be made. Alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made. The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid contains any combination of deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthanine, hypoxanthanine, isocytosine, isoguanine, and basepair analogs such as nitropyrrole and nitroindole, etc.

As described above generally for proteins, nucleic acid bioactive agents may be naturally occuring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of procaryotic or eukaryotic genomes may be used as is outlined above for proteins.

In general, probes of the present invention are designed to be complementary to a target sequence (either the target analyte sequence of the sample or to other probe sequences, as is described

herein), such that hybridization of the target and the probes of the present invention occurs. This complementarity need not be perfect; there may be any number of base pair mismatches that will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under the selected reaction conditions. High stringency conditions are known in the art; see for example Maniatis et al., Molecular Cloning: A Laboratory Manual, 2d Edition, 1989, and Short Protocols in Molecular Biology, ed. Ausubel, et al., both of which are hereby incorporated by reference. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g. 10 to 50 nucleotides) and at least about 60°C for long probes (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. In another embodiment, less stringent hybridization conditions are used; for example, moderate or low stringency conditions may be used, as are known in the art; see Maniatis and Ausubel, supra, and Tijssen, supra.

The term 'target sequence" or grammatical equivalents herein means a nucleic acid sequence on a single strand of nucleic acid. The target sequence may be a portion of a gene, a regulatory sequence, genomic DNA, cDNA, RNA including mRNA and rRNA, or others. It may be any length, with the understanding that longer sequences are more specific. As will be appreciated by those in the art, the complementary target sequence may take many forms. For example, it may be contained within a larger nucleic acid sequence, i.e. all or part of a gene or mRNA, a restriction fragment of a plasmid or genomic DNA, among others. As is outlined more fully below, probes are made to hybridize to target sequences to determine the presence or absence of the target sequence in a sample. Generally speaking, this term will be understood by those skilled in the art.

In a preferred embodiment, the bioactive agents are organic chemical moieties, a wide variety of which are available in the literature.

In a preferred embodiment, each bead comprises a single type of bioactive agent, although a plurality of individual bioactive agents are preferably attached to each bead. Similarly, preferred embodiments utilize more than one microsphere containing a unique bioactive agent; that is, there is redundancy built into the system by the use of subpopulations of microspheres, each microsphere in the subpopulation containing the same bioactive agent.

As will be appreciated by those in the art, the bioactive agents may either be synthesized directly on the beads, or they may be made and then attached after synthesis. In a preferred embodiment, linkers are used to attach the bioactive agents to the beads, to allow both good attachment, sufficient flexibility to allow good interaction with the target molecule, and to avoid undesirable binding reactions.

In a preferred embodiment, the bioactive agents are synthesized directly on the beads. As is known in the art, many classes of chemical compounds are currently synthesized on solid supports, such as peptides, organic moieties, and nucleic acids. It is a relatively straightforward matter to adjust the current synthetic techniques to use beads.

In a preferred embodiment, the bioactive agents are synthesized first, and then covalently attached to the beads. As will be appreciated by those in the art, this will be done depending on the composition of the bioactive agents and the beads. The functionalization of solid support surfaces such as certain polymers with chemically reactive groups such as thiols, amines, carboxyls, etc. is generally known in the art. Accordingly, "blank" microspheres may be used that have surface chemistries that facilitate the attachment of the desired functionality by the user. Some examples of these surface chemistries for blank microspheres are listed in Table I.

Table I

Surface chemistry	Name:
NH_2	Amine
COOH	Carboxylic Acid
CHO	Aldehyde
CH_2-NH_2	Aliphalic Amine
CO NH_2	Amide
CH_2-C1	Chloromethyl
CONH-NH_2	Hydrazide

OH	Hydroxyl
SO_4	Sulfate
SO_3	Sulfonate
Ar NH_2	Aromatic Amine

These functional groups can be used to add any number of different bioactive agents to the beads, generally using known chemistries. For example, bioactive agents containing carbohydrates may be attached to an amino-functionalized support; the aldehyde of the carbohydrate is made using standard techniques, and then the aldehyde is reacted with an amino group on the surface. In an alternative embodiment, a sulfhydryl linker may be used. There are a number of sulfhydryl reactive linkers known in the art such as SPDP, maleimides, α-haloacetyls, and pyridyl disulfides (see for example the 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference) which can be used to attach cysteine containing proteinaceous agents to the support. Alternatively, an amino group on the bioactive agent may be used for attachment to an amino group on the surface. For example, a large number of stable bifunctional groups are well known in the art, including homobifunctional and heterobifunctional linkers (see Pierce Catalog and Handbook, pages 155-200). In an additional embodiment, carboxyl groups (either from the surface or from the bioactive agent) may be derivatized using well known linkers (see the Pierce catalog). For example, carbodiimides activate carboxyl groups for attack by good nucleophiles such as amines (see Torchilin et al., Critical Rev. Therapeutic Drug Carrier Systems, 7(4):275-308 (1991), expressly incorporated herein). Proteinaceous bioactive agents may also be attached using other techniques known in the art, for example for the attachment of antibodies to polymers; see Slinkin et al., Bioconj. Chem. 2:342-348 (1991); Torchilin et al., *supra*; Trubetskoy et al., Bioconj. Chem. 3:323-327 (1992); King et al., Cancer Res. 54:6176-6185 (1994); and Wilbur et al., Bioconjugate Chem. 5:220-235 (1994), all of which are hereby expressly incorporated by reference). It should be understood that the bioactive agents may be attached in a variety of ways, including those listed above. What is important is that manner of attachment does not significantly alter the functionality of the bioactive agent; that is, the bioactive agent should be attached in such a flexible manner as to allow its interaction with a target.

Specific techniques for immobilizing enzymes on microspheres are known in the prior art. In one case, NH_2 surface chemistry microspheres are used. Surface activation is achieved with a 2.5% glutaraldehyde in phosphate buffered saline (10 mM) providing a pH of 6.9. (138 mM NaCl, 2.7 mM, KCl). This is stirred on a stir bed for approximately 2 hours at room temperature. The microspheres are then rinsed with ultrapure water plus 0.01% tween 20 (surfactant) -0.02%, and rinsed again with a pH 7.7 PBS plus 0.01% tween 20. Finally, the enzyme is added to the solution, preferably after being prefiltered using a 0.45μm amicon micropure filter.

After the desired number of bead subpopulations are prepared by the method of steps 50 through 54, discussed above, the subpopulations are typically combined in step 55 to provide a random mixture of subpopulations for use as sensor array elements prior to dispersing the subpopulation mixture on the array substrate in step 56. In a preferred embodiment, Fig. 3. shows a schematic process diagram which illustrates the preparation and placement of self-encoded sensor bead subpopulations in fiber optic bundle sensor array . In an alternative embodiment, step 55 may be omitted and each of the sensor bead subpopulations may be separately and sequentially positioned on the array substrate in predetermined locations.

Thus, once the microspheres are made comprising at least one bioactive agent and a self-encoding reporter dye, the microspheres are added to discrete sites on the surface of the substrate. This can be done in a number of ways, but generally comprises adding the beads to the surface under conditions that will allow the association of the microspheres on or at the discrete sites. The association of the beads on the surface may comprise a covalent bonding of the bead to the surface, for example when chemical attachment sites are added to both the substrate and the bead; an electrostatic or hydroaffinity, when charge, hydrophobicity or hydrophilicity is used as the basis of the binding; a physical yet non-covalent attachment such as the use of an adhesive; or a spatial attachment, for example the localization of a bead within a well. In some embodiments it may be preferable to effect a more permanent attachment after the initial localization, for example through the use of cross-linking agents, a film or membrane over the array.

The microsphere system may be attached to the distal end of the optical fiber bundle using a variety of compatible processes as outlined below. It is important that the microspheres are located close to the end of the bundle. This ensures that the light returning in each optical fiber predominantly comes from only a single microsphere. This feature is necessary to enable the interrogation of the optical response signature of individual microspheres to identify reactions involving the microsphere's functionality and also to decode the dye/bead sets contained in those microspheres. The adhesion or affixing technique, however, must not chemically insulate the microspheres from the analyte.

Fig. 7 is a schematic block diagram showing the inventive optical fiber sensor 200 and associated control system 210. The fiber optic sensor 200 comprises a fiber optic bundle 202 (Galileo Electro-Optics, Sturbridge, MA), that is typically constructed from many thousands of separately clad discrete fibers, each only a few microns in diameter, so that light does not mix between the individual fibers. Any suitable fiber optic bundle 202 may be employed having a range in the number of individual fibers or a range of individual fiber diameters. The microsphere or bead sensor array 100 is attached to the bundle's distal end 212, with the proximal end 214 being received by a conventional z-translation microscope stage 216, for vertical positioning of the array 100 for focusing, and an x-y micropositioner 218 (Burleigh Instruments, Fishers, NY), for horizontal alignment of the array 100 with the optical train. These two components act in concert to properly position the proximal end 214 of the bundle 202 for

a microscope objective lens 220. Light collected by the objective lens 220 is passed to a reflected light fluorescence attachment with a four position dichromic cube wheel 222. The attachment 222 allows insertion of light from a 75 Watt Xenon arc lamp 224 through the objective lens 220 to be coupled into the fiber bundle 202. The light from the source 224 is condensed by condensing lens 226, then filtered and/or shuttered by filter and shutter wheel 228.

Light returning from the distal end 212 of the bundle 202 is passed by the attachment 222 and is then shuttered and filtered by a second wheel 234. The light is then imaged on a charge coupled device (CCD) camera 236. A conventional computer 238 executes imaging processing software to process the information from the CCD camera 236 and also possibly control the first and second shutter and filter wheels 228, 234. Either a Macintosh or, alternatively, an IBM-compatible personal computer may be utilized for controlling the instrumentation and data acquisition. The instrumentation and optical apparatus deployed at the proximal end 214 of the fiber optic sensor 200, exclusive of the fiber optic sensor 200, are discussed more completely by Bronk, et al., *Anal. Chem.* 67(17):2750-2752(1995) and Bronk, et al., *Anal. Chem.* 66:3519 (1994).

The bead sensor array 100 may be attached to the distal end of the optical fiber bundle 202 using a variety of compatible processes. It is important that the microspheres 10 are located close to the end of the bundle. This ensures that the light returning in each discrete optical fiber predominantly comes from only a single microsphere. This feature is necessary to decode the self-encoded bead subpopulations for the purpose of identifying both bead type and location, to enable the interrogation of the optical signature of individual microspheres within a bead subpopulation, and to provide for the summing of individual bead responses within each subpopulation for reducing signal to noise and improving signal enhancement. The bead adhesion or affixing technique, however, must not chemically insulate the microspheres from the analyte or interfere with the optical measurement.

Figs. 5A and 5B are micrographs illustrating the preferred method for attaching beads to a sensor array substrate. Microwells 250 are formed on the distal end 212 of a fiber optic bundle 202 and microspheres 10 are inserted in the microwell cavities 250. The microwells 250 are formed at the center of each optical fiber 252 of the fiber optic bundle 202. As shown in Fig. 5B, the size of the microwells 250 are coordinated with the size of the microspheres 10 so that the microspheres 10 can be placed within the microwells 250. Thus, each optical fiber 252 of the bundle 202 conveys light from the single microsphere 10 contained in its well. Consequently, by imaging the end of the bundle 202 onto the CCD array 236, the optical signatures of the microspheres 10 are individually interrogatable.

Fig. 4 illustrates how the microwells 250 are formed and microspheres 10 placed in the microwells. In one embodiment, a 1 mm hexagonally-packed imaging fiber bundle 202 was employed comprising approximately 20,600 individual optical fibers having cores approximately 3.7μm across (Part No. ET26 from Galileo Fibers, Sturbridge, MA). Typically, the cores of each fiber are

hexagonally shaped as a result of glass hardness and drawing during fiber fabrication. In some cases, the shape can be circular, however.

In step 270, both the proximal 214 and distal 212 ends of the fiber bundle 202 are successively polished on 12 μm, 9 μm, 3 μm, 1 μm, and 0.3 μm lapping films. Subsequently, the ends can be inspected for scratches on a conventional atomic force microscope. In step 272, etching is performed on the distal end 212 of the bundle 202. A solution of 0.2 grams NH_4F (ammonium fluoride) with 600 μl dH_20 and 100 μl of HF (hydrofluoric acid), 50% stock solution, may be used. The distal end 212 is etched in this solution for a specified time, preferably approximately 80 seconds.

Upon removal from this solution, the bundle end is immediately placed in deionized water to stop any further etching in step 274. The fiber is then rinsed in running tap water. At this stage, sonication is preferably performed for several minutes to remove any salt products from the reaction. The fiber is then allowed to air dry.

As illustrated in Figs. 5A and 5B, the foregoing procedure produces microwells by the anisotropic etching of the fiber cores 254 favorably with respect to the cladding 256 for each fiber of the bundle 202. The microwells have approximately the diameter of the cores 254, 3.7 μm. This diameter is selected to be slightly larger than the diameters of the microspheres used, 3.1μm, in the example. The preferential etching occurs because the pure silica of the cores 254 etches faster in the presence of hydrofluoric acid than the germanium-doped silica claddings 256.

The microspheres are then placed in the microwells 250 in step 276 according to a number of different techniques. The placement of the microspheres may be accomplished by dripping a solution containing the desired randomly mixed subpopulations of the microspheres over the distal end 212, sonicating the bundle to settle the microspheres in the microwells, and allowing the microsphere solvent to evaporate. Alternatively, the subpopulations could be added serially to the bundle end. Microspheres 10 may then be fixed into the microwells 250 by using a dilute solution of sulfonated Nafion that is dripped over the end. Upon solvent evaporation, a thin film of Nafion was formed over the microspheres which holds them in place. This approach is compatible for fixing microspheres for pH indication that carry FITC functionality. The resulting array of fixed microspheres retains its pH sensitivity due to the permeability of the sulfonated Nafion to hydrogen ions. This approach, however, can not be employed generically as Nafion is impermeable to most water soluble species. A similar approach can be employed with different polymers. For example, solutions of polyethylene glycol, polyacrylamide, or polyhydroxymethyl methacrylate (polyHEMA) can be used in place of Nafion, providing the requisite permeability to aqueous species.

In an another embodiment, an alternative fixation approach employs microsphere swelling to entrap each microsphere 10 in its corresponding microwell 250. In this approach, the microspheres are first

distributed into the microwells 250 by sonicating the microspheres suspended in a non-swelling solvent in the presence of the microwell array on the distal end 212. After placement into the microwells, the microspheres are subsequently exposed to an aqueous buffer in which they swell, thereby physically entrapping them in the microwells. By way of example of this particular embodiment, one commonly known microsphere material is tentagel, a styrene-polyethylene glycol copolymer. These microspheres can be unswollen in nonpolar solvents such as hexane and swell approximately 20-40% in volume upon exposure to a more polar or aqueous media. In certain embodiments, this fixation approach may be desirable since it does not significantly compromise the diffusional or permeability properties of the microspheres themselves.

Figs. 6A and 6B show typical microspheres 10 in microwells 250 after their initial placement and then after tapping and exposure to air pulses. Figs. 6A and 6B illustrate that there is no appreciable loss of microspheres from the microwells due to mechanical agitation even without a specific fixing technique. This effect is probably due to electrostatic forces between the microspheres and the optical fibers. These forces tend to bind the microspheres within the microwells. Thus, in most environments, it may be unnecessary to use any chemical or mechanical fixation for the microspheres.

It should be noted that not all sites of an array may comprise a bead; that is, there may be some sites on the substrate surface which are empty. In addition, there may be some sites that contain more than one bead, although this is not preferred.

In some embodiments, for example when chemical attachment is done, it is possible to attach the beads in a non-random or ordered way. For example, using photoactivatible attachment linkers or photoactivatible adhesives or masks, selected sites on the array may be sequentially rendered suitable for attachment, such that defined populations of beads are laid down.

In addition, since the size of the array will be set by the number of unique optical response signatures, it is possible to "reuse" a set of unique optical response signatures to allow for a greater number of test sites. This may be done in several ways; for example, by using a positional coding scheme within an array; different sub-bundles may reuse the set of optical response signatures. Similarly, one embodiment utilizes bead size as a coding modality, thus allowing the reuse of the set of unique optical response signatures for each bead size. Alternatively, sequential partial loading of arrays with beads can also allow the reuse of optical response signatures.

In a preferred embodiment, a spatial or positional coding system is done. In this embodiment, there are sub-bundles or subarrays (i.e. portions of the total array) that are utilized. By analogy with the telephone system, each subarray is an "area code", that can have the same tags (i.e. telephone numbers) of other subarrays, that are separated by virtue of the location of the subarray. Thus, for example, the same unique dye/bead combinations can be reused from bundle to bundle. Thus, the

use of 50 unique tags in combination with 100 different subarrays can form an array of 5000 different bioactive agents. In this embodiment, it becomes important to be able to identify one bundle from another; in general, this is done either manually or through the use of marker beads, i.e. beads containing unique tags for each subarray.

In alternative embodiments, additional encoding parameters can be added, such as microsphere size. For example, the use of different size beads may also allow the reuse of sets of optical response signatures; that is, it is possible to use microspheres of different sizes to expand the encoding dimensions of the microspheres. Optical fiber arrays can be fabricated containing pixels with different fiber diameters or cross-sections; alternatively, two or more fiber optic bundles, each with different cross-sections of the individual fibers, can be added together to form a larger bundle; or, fiber optic bundles with fiber of the same size cross-sections can be used, but just with different sized beads. With different diameters, the largest wells can be filled with the largest microspheres and then moving onto progressively smaller microspheres in the smaller wells until all size wells are then filled. In this manner, the same dye/bead combinations could be used to encode microspheres of different sizes thereby expanding the number of different oligonucleotide sequences or chemical functionalities present in the array. Although outlined for fiber optic substrates, this as well as the other methods outlined herein can be used with other substrates and with other attachment modalities as well.

In a preferred embodiment, the coding and decoding is accomplished by sequential loading of the microspheres into the array. As outlined above for spatial coding, in this embodiment, the optical response signatures can be "reused". In this embodiment, the library of microspheres each comprising a different bioactive agent (or the subpopulations each comprise a different bioactive agent), is divided into a plurality of sublibraries; for example, depending on the size of the desired array and the number of unique tags, 10 sublibraries each comprising roughly 10% of the total library may be made, with each sublibrary comprising roughly the same unique tags. Then, the first sublibrary is added to the fiber optic bundle comprising the wells, and the location of each bioactive agent is determined, using its optical response signature. The second sublibrary is then added, and the location of each optical response signature is again determined. The signal in this case will comprise the "first" optical response signature and the "second" optical response signature; by comparing the two matrices the location of each bead in each sublibrary can be determined. Similarly, adding the third, fourth, etc. sublibraries sequentially will allow the array to be filled.

Thus, arrays are made of a large spectrum of chemical functionalities utilizing the compositions of invention comprising microspheres and substrates with discrete sites on a surface. Specifically, prior art sensors which can be adapted for use in the present invention include four broad classifications of microsphere sensors: 1) basic indicator chemistry sensors; 2) enzyme-based sensors; 3) immuno-based sensors (both of which are part of a broader general class of protein sensors); and 4) geno-sensors.

In a preferred embodiment, the bioactive agents are used to detect chemical compounds. A large number of basic indicator sensors have been previously demonstrated. Examples include:

Table III

TARGET ANALYTE	Bioactive agent (or chemical functionality)	NOTES ($\lambda_{AB}/\lambda_{EM}$)
pH Sensors based on:	seminaphthofluoresceins	*e.g.*, carboxyl-SNAFL
	seminaphthorhodafluors	*e.g.*, carboxyl-SNARF
	8-hydroxypyrene-1,3,6-trisulfonic acid	
	fluorescein	
CO2 Sensors based On:	seminaphthofluoresceins	*e.g.*, carboxyl-SNAFL
	seminaphthorhodafluors	*e.g.*, carbody-SNARF
	8-hydroxypyrene-1,3,6-trisulfonic acid	
Metal Ions Sensors based on:	desferriozamine B	*e.g.*, Fe
	cyclen derivative	*e.g.*, Cu, Zn
	derivatized peptides	*e.g.*, FITC-Gly-Gly-His, and FITC-Gly His, Cu, Zn
	fluorexon (calcine)	*e.g.*, Ca, Mg, Cu, Pb, Ba
	calcine blue	*e.g.*, Ca, Mg, Cu
	methyl calcine blue	*e.g.*, Ca, Mg, Cu
	ortho-dianisidine tetracetic acid (ODTA)	*e.g.*, Zn
	bis-salicylidene ethylenediamine (SED)	*e.g.*, Al
	N-(6-methozy-8-quinolyl-p-	*e.g.*, Zn

Table III		
	toluenesulfonamine (TSQ)	
	Indo-1	*e.g.*, Mn, Ni
	Fura-2	*e.g.*, Mn, Ni
	Magesium Green	*e.g.*, Mg, Cd, Tb
O_2	Siphenylisobenzofuran	409/476
	Methoxyvinyl pyrene	352/401
Nitrite	diaminonaphthalene	340/377
NO	luminol	355/411
	dihydrohodamine	289/none
Ca^{2+}	Bis-fura	340/380
	Calcium Green	visible light/530
	Fura-2	340/380
	Indo-1	405/485
	Fluo-3	visible light/525
	Rhod-2	visible light/570
Mg^{2+}	Mag-Fura-2	340/380
	Mag-Fura-5	340/380
	Mag-Indo-1	405/485
	Magnesium Green	475/530
	Magnesium Orange	visible light/545
Zn^{2+}	Newport Green	506/535
TSQ	Methoxy-Quinobyl	334/385
Cu^{+}	Phen Green	492/517
Na^{+}	SBFI	339/565

Table III		
	SBFO	354/575
	Sodium Green	506/535
K^+	PBFI	336/557
Cl^-	SPQ	344/443
	MQAE	350/460

Each of the chemicals listed in Table III directly produces an optically interrogatable signal or a change in the optical signature, as is more fully outlined below, in the presence of the targeted analyte.

Enzyme-based microsphere sensors have also been demonstrated and could be manifest on microspheres. Examples include:

Table IV

SENSOR TARGET	Bioactive agent
Glucose Sensor	glucose oxidase (enz.) + O_2-sensitive dye (see Table I)
Penicillin Sensor	penicillinase (enz.) + pH-sensitive dye (see Table I)
Urea Sensor	urease (enz.) + pH-sensitive dye (see Table I)
Acetylcholine Sensor	acetylcholinesterase (enz.) + pH-sensitive dye (see Table I)

Generally, as more fully outlined above, the induced change in the optical signal upon binding of the target analyte due to the presence of the enzyme-sensitive chemical analyte occurs indirectly in this class of chemical functionalities. The microsphere-bound enzyme, *e.g.,* glucose oxidase, decomposes the target analyte, *e.g.,* glucose, consume a co-substrate, *e.g.,* oxygen, or produce some by-product, *e.g.,* hydrogen peroxide. An oxygen sensitive dye is then used to trigger the signal change.

Immuno-based microsphere sensors have been demonstrated for the detection for environmental pollutants such as pesticides, herbicides, PCB's and PAH's. Additionally, these sensors have also been used for diagnostics, such as bacterial (*e.g.,* leprosy, cholera, lyme disease, and tuberculosis), viral (*e.g.,* HIV, herpes simplex, cytomegalovirus), fungal (*e.g.,* aspergillosis, candidiasis,

cryptococcoses), Mycoplasmal (e.g., mycoplasmal pneumonia), Protozoal (e.g., amoebiasis, toxoplasmosis), Rickettsial (e.g., Rocky Mountain spotted fever), and pregnancy tests.

Microsphere genosensors may also be made. These are typically constructed by attaching a probe sequence to the microsphere surface chemistry, typically via an NH_2 group. A fluorescent dye molecule, *e.g.*, fluorescein, is attached to the target sequence, which is in solution. The optically interrogatable signal change occurs with the binding of the target sequences to the microsphere. This produces a higher concentration of dye surrounding the microsphere than in the solution generally. A few demonstrated probe and target sequences, see Ferguson, J.A. et al. *Nature Biotechnology*, Vol. 14, Dec. 1996, are listed below in Table V.

Table V

PROBE SEQUENCES	TARGET SEQUENCES
B-glo(+) (segment of human B-globin)5'-NH_2-$(CH_2)_8$-)TT TTT TTT TCA ACT TCA TCC ACG TTC ACC-3	B-glo(+)-CF 5'-Fluorescein-TC AAC GTG GAT GAA GTT C-3'
IFNG(interferon gamma 1)5'-NH_2-$(CH_2)_8$-T_{12}-TGG CTT CTC TTG GCT GTT ACT-3'	IFNG-CF 5'-Fluorescein-AG TAA CAG CCA AGA GAA CCC AAA-3'
IL2(interleukin-2)5'-NH_2-$(CH_2)_8$-T_{12}-TA ACC GAA TCC CAA ACT CAC CAG-3'	IL2-CF 5'-Fluorescein-CT GGT GAG TTT GGG ATT CTT GTA-3'
IL4(interleukin-4)5'NH_2-$(CH_2)_8$-T_{12}-CC AAC TGC TTC CCC CTC TGT-3'	IL4-CF 5'-Fluorescein-AC AGA GGG GGA AGC AGT TGG-3'
IL6(interleukin-6)5'NH_2-$(CH_2)_8$-T12-GT TGG GTC AGG GGT GGT TAT T-3'	IL6-CF 5'-Fluorescein-AA TAA CCA CCC CTG ACC CAA C-3'

It should be further noted that the genosensors can be based on the use of hybridization indicators as the labels. Hybridization indicators preferentially associate with double stranded nucleic acid, usually reversibly. Hybridization indicators include intercalators and minor and/or major groove binding moieties. In a preferred embodiment, intercalators may be used; since intercalation generally only occurs in the presence of double stranded nucleic acid, only in the presence of target hybridization will the label light up.

The present invention may be used with any or all of these types of sensors. As will be appreciated by those in the art, the type and composition of the sensor will vary widely, depending on the composition

of the target analyte. That is, sensors may be made to detect nucleic acids, proteins (including enzyme sensors and immunosensors), lipids, carbohydrates, etc; similarly, these sensors may include bioactive agents that are nucleic acids, proteins, lipids, carbohydrates, etc. In addition, a single array sensor may contain different binding ligands for multiple types of analytes; for example, an array sensor for HIV may contain multiple nucleic acid probes for direct detection of the viral genome, protein binding ligands for direct detection of the viral particle, immuno-components for the detection of anti-HIV antibodies, etc.

In addition to the beads and the substrate, the compositions of the invention may include other components, such as light sources, optical components such as lenses and filters, detectors, computer components for data analysis, etc.

The arrays of the present invention are constructed such that information about the identity of the bioactive agent is built into the array, such that the random deposition of the beads on the surface of the substrate can be "decoded" to allow identification of the bioactive agent at all positions. This may be done in a variety of ways.

In a preferred embodiment, the beads are loaded onto the substrate and then the array is decoded, prior to running the assay. This is done by detecting the optical response signature associated with the bead at each site on the array upon exposure to a reference analyte. This may be done all at once, if unique optical signatures are used, or sequentially, as is generally outlined above for the "reuse" of sets of optical signatures. Alternatively, full or partial decoding may occur after the assay is run.

Once made and decoded if necessary, the compositions find use in a number of applications. As a preliminary matter, the invention finds use in methods for reducing the signal-to-noise ratio in the characteristic optical response signature of a sensor array having a subpopulations of array elements. The methods comprise a) decoding the array so as to identify the location of each sensor element within each sensor subpopulation within the array; b) measuring the characteristic optical response signature of each sensor element in the array; c) adjusting the baseline of the optical response signature for each sensor element in the array; d) summing the baseline-adjusted characteristic optical response signature of all sensor elements within each of the sensor subpopulations; and e) reporting the characteristic optical response signature of each sensor subpopulation as a summation of the baseline-adjusted characteristic optical response signatures of all sensor elements within each of the subpopulations. This can result in an increase in the signal-to-noise ratio by a factor of at least about ten, with at least about 100 being preferred.

Similarly, the invention provides methods for amplifying the characteristic optical response signature of a sensor array having subpopulations of array elements, comprising: a) decoding the array so as to

identify the location of each sensor element within each sensor subpopulation within the array; b) measuring a characteristic optical response signature of each sensor element in the array; c) adjusting the baseline of the optical response signature for each sensor element in the array; d) summing the baseline-adjusted characteristic optical response signature of all sensor elements within each of the sensor subpopulations; and e) reporting the characteristic optical response signature of each sensor subpopulation as a summation of the baseline-adjusted characteristic optical response signatures of all sensor elements within each of the subpopulations. In a preferred embodiment, the signal is amplified by a factor of at least about fifty and an analyte detection limit is reduced by a factor of at least about 100.

Generally, a sample containing a target analyte (whether for detection of the target analyte or screening for binding partners of the target analyte) is added to the array, under conditions suitable for binding of the target analyte to at least one of the bioactive agents, i.e. generally physiological conditions. The presence or absence of the target analyte is then detected. As will be appreciated by those in the art, this may be done in a variety of ways, generally through the use of a change in an optical signal. This change can occur via many different mechanisms. A few examples include the binding of a dye-tagged analyte to the bead, the production of a dye species on or near the beads, the destruction of an existing dye species, a change in the optical signature upon analyte interaction with dye on bead, or any other optical interrogatable event.

In a preferred embodiment, the change in optical signal occurs as a result of the binding of a target analyte that is labeled, either directly or indirectly, with a detectable label, preferably an optical label such as a fluorochrome. Thus, for example, when a proteinaceous target analyte is used, it may be either directly labeled with a fluor, or indirectly, for example through the use of a labeled antibody. Similarly, nucleic acids are easily labeled with fluorochromes, for example during PCR amplification as is known in the art. Alternatively, upon binding of the target sequences, an intercalating dye (*e.g.*, ethidium bromide) can be added subsequently to signal the presence of the bound target to the probe sequence. Upon binding of the target analyte to a bioactive agent, there is a new optical signal generated at that site, which then may be detected.

Alternatively, in some cases, as discussed above, the target analyte such as an enzyme generates a species (for example, a fluorescent product) that is either directly or indirectly detectable optically.

Furthermore, in some embodiments, a change in the optical response signature may be the basis of the optical signal. For example, the interaction of some chemical target analytes with some fluorescent dyes on the beads may alter the optical response signature, thus generating a different optical signal. For example, fluorophore derivatized receptors may be used in which the binding of the ligand alters the signal.

As will be appreciated by those in the art, in some embodiments, the presence or absence of the target analyte may be done using changes in other optical or non-optical signals, including, but not limited to, surface enhanced Raman spectroscopy, surface plasmon resonance, radioactivity, etc.

The assays may be run under a variety of experimental conditions, as will be appreciated by those in the art. A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding. Various blocking and washing steps may be utilized as is known in the art.

In a preferred embodiment, the compositions are used to probe a sample solution for the presence or absence of a target analyte. By "target analyte" or "analyte" or grammatical equivalents herein is meant any atom, molecule, ion, molecular ion, compound or particle to be either detected or evaluated for binding partners. As will be appreciated by those in the art, a large number of analytes may be used in the present invention; basically, any target analyte can be used which binds a bioactive agent or for which a binding partner (i.e. drug candidate) is sought.

Suitable analytes include organic and inorganic molecules, including biomolecules. When detection of a target analyte is done, suitable target analytes include, but are not limited to, an environmental pollutant (including pesticides, insecticides, toxins, etc.); a chemical (including solvents, polymers, organic materials, etc.); therapeutic molecules (including therapeutic and abused drugs, antibiotics, etc.); biomolecules (including hormones, cytokines, proteins, nucleic acids, lipids, carbohydrates, cellular membrane antigens and receptors (neural, hormonal, nutrient, and cell surface receptors) or their ligands, etc); whole cells (including procaryotic (such as pathogenic bacteria) and eukaryotic cells, including mammalian tumor cells); viruses (including retroviruses, herpesviruses, adenoviruses, lentiviruses, etc.); and spores; etc. Particularly preferred analytes are nucleic acids and proteins.

In a preferred embodiment, the target analyte is a protein. As will be appreciated by those in the art, there are a large number of possible proteinaceous target analytes that may be detected or evaluated for binding partners using the present invention. Suitable protein target analytes include, but are not limited to, (1) immunoglobulins; (2) enzymes (and other proteins); (3) hormones and cytokines (many of which serve as ligands for cellular receptors); and (4) other proteins.

In a preferred embodiment, the target analyte is a nucleic acid. These assays find use in a wide variety of applications.

In a preferred embodiment, the probes are used in genetic diagnosis. For example, probes can be made using the techniques disclosed herein to detect target sequences such as the gene for nonpolyposis colon cancer, the BRCA1 breast cancer gene, P53, which is a gene associated with a variety of cancers, the Apo E4 gene that indicates a greater risk of Alzheimer's disease, allowing for easy presymptomatic screening of patients, mutations in the cystic fibrosis gene, or any of the others well known in the art.

In an additional embodiment, viral and bacterial detection is done using the complexes of the invention. In this embodiment, probes are designed to detect target sequences from a variety of bacteria and viruses. For example, current blood-screening techniques rely on the detection of anti-HIV antibodies. The methods disclosed herein allow for direct screening of clinical samples to detect HIV nucleic acid sequences, particularly highly conserved HIV sequences. In addition, this allows direct monitoring of circulating virus within a patient as an improved method of assessing the efficacy of anti-viral therapies. Similarly, viruses associated with leukemia, HTLV-I and HTLV-II, may be detected in this way. Bacterial infections such as tuberculosis, clymidia and other sexually transmitted diseases, may also be detected.

In a preferred embodiment, the nucleic acids of the invention find use as probes for toxic bacteria in the screening of water and food samples. For example, samples may be treated to lyse the bacteria to release its nucleic acid, and then probes designed to recognize bacterial strains, including, but not limited to, such pathogenic strains as, *Salmonella*, *Campylobacter*, *Vibrio cholerae*, *Leishmania*, enterotoxic strains of *E. coli*, and Legionnaire's disease bacteria. Similarly, bioremediation strategies may be evaluated using the compositions of the invention.

In a further embodiment, the probes are used for forensic "DNA fingerprinting" to match crime-scene DNA against samples taken from victims and suspects.

In an additional embodiment, the probes in an array are used for sequencing by hybridization.

The present invention also finds use as a methodology for the detection of mutations or mismatches in target nucleic acid sequences. For example, recent focus has been on the analysis of the relationship between genetic variation and phenotype by making use of polymorphic DNA markers. Previous work utilized short tandem repeats (STRs) as polymorphic positional markers; however, recent focus is on the use of single nucleotide polymorphisms (SNPs), which occur at an average frequency of more than 1 per kilobase in human genomic DNA. Some SNPs, particularly those in and around coding sequences, are likely to be the direct cause of therapeutically relevant phenotypic variants. There are a number of well known polymorphisms that cause clinically important phenotypes; for example, the apoE2/3/4 variants are associated with different relative risk of Alzheimer's and other diseases (see Cordor et al., Science 261(1993). Multiplex PCR amplification of SNP loci with subsequent

hybridization to oligonucleotide arrays has been shown to be an accurate and reliable method of simultaneously genotyping at least hundreds of SNPs; see Wang et al., Science, 280:1077 (1998); see also Schafer et al., Nature Biotechnology 16:33-39 (1998). The compositions of the present invention may easily be substituted for the arrays of the prior art.

In a preferred embodiment, the compositions of the invention are used to screen bioactive agents to find an agent that will bind, and preferably modify the function of, a target molecule. As above, a wide variety of different assay formats may be run, as will be appreciated by those in the art. Generally, the target analyte for which a binding partner is desired is labeled; binding of the target analyte by the bioactive agent results in the recruitment of the label to the bead, with subsequent detection.

In a preferred embodiment, the binding of the bioactive agent and the target analyte is specific; that is, the bioactive agent specifically binds to the target analyte. By "specifically bind" herein is meant that the agent binds the analyte, with specificity sufficient to differentiate between the analyte and other components or contaminants of the test sample. However, as will be appreciated by those in the art, it will be possible to detect analytes using binding which is not highly specific; for example, the systems may use different binding ligands, for example an array of different ligands, and detection of any particular analyte is via its "signature" of binding to a panel of binding ligands, similar to the manner in which "electronic noses" work. This finds particular utility in the detection of chemical analytes. The binding should be sufficient to remain bound under the conditions of the assay, including wash steps to remove non-specific binding, although in some embodiments, wash steps are not desired; i.e. for detecting low affinity binding partners. In some embodiments, for example in the detection of certain biomolecules, the dissociation constants of the analyte to the binding ligand will be less than about 10^{-4}-10^{-6} M^{-1}, with less than about 10^{-5} to 10^{-9} M^{-1} being preferred and less than about 10^{-7} -10^{-9} M^{-1} being particularly preferred.

The following examples serve to more fully describe the manner of using the above-described invention, as well as to set forth the best modes contemplated for carrying out various aspects of the invention. It is understood that these examples in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All references cited herein are incorporated by reference in their entireity.

EXAMPLES

Characteristic temporal optical response data measurements of sensor bead and sensor array response to specific vapor analytes and excitation light energy were made according to the established method and instrumentation disclosed by White, et al., *Anal. Chem.* 68:2191-2202(1996). In Fig. 8, a schematic diagram illustrates the experimental apparatus and instrumentation used for the data measurements reported in Examples 7 through 17.

In a typical measurement, the proximal end 214 of a fiber optic bundle 202 was placed into a fiber chuck 300 and secured for viewing with an Olympus microscope-based imaging system. In other embodiments, a conventional Olympus microscope slide platform and slide clamp was used for positioning alternative sensor array substrates, such as glass cover slips. An Olympus microscope 320 equipped with an epi-illuminator was utilized for optical measurements. The microscope 320 was equipped with Olympus 20x and 40x and Zeiss 100x objectives. An Omega 560 DCRP dichroic mirror 330 was used to direct filtered exitation light energy from a 75W Xenon arc lamp 340 to the sensor array 100 and to permit the emitted light energy, due to the characteristic optical response signature originating from each of the sensor beads 10 in the sensor array 100, to be recorded by a CCD frame transfer camera 310. The excitation light energy emanating from the arc lamp 340 was filtered by an Omega 535 BP40 integrated excitation light filter/shutter 350. The emission light energy which emitted from the sensor beads 10 of the sensor array 100 was filtered with an Omega 640 BP20 integrated emitted light filter/shutter 360 prior to the CCD frame transfer camera 310.

Experiments generally consist of collecting video camera frames of fluorescence response images of the characteristic optical response signatures of individual sensor beads in the sensor array 100 conveyed by the fiber optic bundle 202 to its proximal end 214. The bead and array images are recorded with a CCD frame transfer camera 310 (Model TE512EFT from Princeton Instruments, Trenton, NJ). A preselected number of image frames are captured and sent to a computer system 400, comprising a Princeton Instruments NUBus camera interface card installed in a 8100AV Macintosh Power PC. Camera frame rates can be set at any desired value and typically range between 80 to 250 ms/frame. The following is a list of frame rates (time between data points) used in acquiring the data shown in for Figs 9-16. The specified frame rate corresponds to a specific time interval between data points.

Figure	Rate (ms/frame)	Total No. of Frames
Fig.9	135	30
Fig.10	183.3	60
Fig.11	103.3	60
Fig.12	190.6	60
Fig.13	133	60
Fig.14	155	60
Fig.15	155	60
Fig.16	124	60

A conventional air dilution olfactometer and vacuum-controlled vapor delivery system 500, as commonly known and used in olfactory research and described in Kauer, et al., *J. Physiol.* 272:495-516 (1977), was used to apply controlled pulses of analyte vapor and air carrier gas to either a sensor

bead substrate or the distal end 212 of a fiber optic sensor array 100 containing an array of sensor beads 10 immobilized in microwells 250.

To produce a saturated vapor sample, generally, a stream of air carrier gas is passed through a 5 ml cartridge containing filter paper saturated with the analyte. Analyte dilutions are produced by adjusting the relative flow rates of saturated vapor and clean carrier gas streams. Typically, a flow rate of 100 ml/min is used for the combined gas flow to the sensor array. At this flow rate, a 2 second pulse would deliver approximately 3.3 ml of analyte vapor with carrier gas. In generally, depending on the analyte vapor pressure and dilution factor, vapor pulses contain between 10^{-7} to 10^{-5} mol of analyte.

The vapor pulse was typically delivered during the 11th through 30th frame, commencing on the 11th frame. The duration of the vapor pulse varied with the specific frame rate utilized and typically ranged between 2 to 3 seconds. Baseline control measurements were performed with high purity, Ultra Zero grade air. The air pulse measurements were performed to account for any bead responses due to the vapor carrier gas.

Data processing: Following the collection of a temporal series of sensor bead or sensor array images, segments are drawn, using IPLab image processing software (Signal Analytics, Vienna, VA), over each pixel which corresponds to an individual fiber where the fiber is coupled to one sensor bead at its distal end. The mean fluorescence intensity was measured for each one of these segments in each frame in the sequence. This is done for both the vapor pulse responses and the baseline air pulse responses. Averages of multiple runs of each may be performed to improve data quality where needed. The air pulse data is then subtracted from the vapor pulse data to subtract the background due to air alone. The resulting data can be plotted to yield temporal intensity responses for all beads of interest. In a preferred embodiment, the sensor array data are used in a neural network analysis according to the method disclosed in White, et al, *Anal. Chem.* 68:2193-2202 (1996).

All data manipulation is performed within the IPLab program environment using simple operator scripts that call standardized image or data processing routines included with the software. These scripts and routines consist of a data collection portion and a data analysis portion.

In the data collection portion, there are three segments or loops as follows:
Loop 1. This establishes the baseline fluorescence of each sensor. This loop can be shortened or extended to adjust to slower or faster response times of specific sensor beads or sensor arrays to certain analytes. For Examples 7 through 17, this loop was set between 5 to 10 frames.

Loop 2. This is the vapor exposure loop. A vapor pulse is applied just before this loop starts by way of a script command that sends a 5 volt pulse to an attached solenoid valve which switches a vacuum

line off, thereby allowing a vapor sample to emit from the end of a nozzle. Typically, this loop is 20 frames in duration. In Example 7, a 10 frame duration was utilized.

Loop 3. This is a sensor recovery loop. Another 5 volt trigger pulse is sent to a solenoid which switches back to its initial position, causing the vacuum system to resume collection of the solvent vapor and carry it off to waste. Typically, this loop is of 30 frames duration. In Example 7, a 15 frame duration was utilized.

Data analysis: In the data analysis portion, pre-selected segments taken from a previously collected "focus" image are transferred to the sequence of images collected. These segments, drawn by the user, allow the mean pixel intensity to be measured in particular regions throughout the image field. Typically, they are drawn over individual pixels of a fiber optic sensor array, each of which contains a bead. The script then enters a loop that steps through each frame, measuring the mean pixel intensity within each segment, and placing the values in data columns. The resulting columns can then be plotted to yield the temporal response of each bead of interest. Before plotting, however, responses are "standardized" by dividing the data for each bead response by its first point. Thus, all responses can be normalized to start at a value of 1.0.

Bead response summing: The optical response signals from a large number of sensor beads within each bead subpopulation can be summed by simply adding the baseline-adjusted intensity values of all responses at each time point, generating a new temporal response comprised of the sum of all bead responses. Signal summing can be performed in real time or during post-data acquisition data reduction and analysis. In one embodiment, signal summing is performed with a commercial spreadsheet program (Excel, Microsoft, Redmond, WA) after optical response data is collected. In a typical procedure, the standardized optical responses are adjusted to start at a value of 0.0 by subtracting the integer 1.0 from all data points. Doing this allows the baseline-loop data to remain at zero even when summed together and the random response signal noise is canceled out. The vapor pulse-loop temporal region, however, exhibits a characteristic change in response, either positive, negative or neutral, prior to the vapor pulse and often requires a baseline adjustment to overcome noise associated with drift in the first few data points due to charge buildup in the CCD camera. If no drift is present, tyically the baseline from the first data point for each bead sensor is subtracted from all the response data for the same bead. If drift is observed, the average baseline from the first ten data points for each bead sensor is substracted from the all the response data for the same bead. By applying this baseline adjustment, when multiple bead responses are added together they can be amplified while the baseline remains at zero. Since all beads respond at the same time to the vapor pulse, they all see the pulse at the exact same time and there is no registering or adjusting needed for overlaying their responses. Cummulative response data is generated by simply adding all data points in successive time intervals. This final column, comprised of the sum of all data points at a particular

time interval, may then be compared or plotted with the individual bead responses to determine the extent of signal enhancement or improved signal-to-noise ratios as shown in Figs. 14 and 15.

Example 1

Preparation of porous silica/Nile Red beads: Approximately 0.5 cm^3 of nominally 3.2 micron diameter commercial porous silica beads were removed from a LUNA column (Phenomenex, Torrance, CA). Sample of beads were placed onto a filter paper and, using vacuum filtration, 0.5 mL of Nile Red (Eastman Kodak, Rochester, NY) solution (1 mg/mL in toluene) was poured over beads. Nile Red was immediately taken up by silica beads, turning them a deep purple color. The beads were washed repeatedly with toluene to remove any excess, non-adsorbed Nile Red.. The beads were dried on a watch glass overnight. Beads were then placed into microwells formed by etching a fiber optic bundle according to the method of the present invention.

Example 2

Preparation of PDPO polymer coated porous silica beads: A silanizing solution was prepared from 20 uL N-octadecyl-triethyoxysilane in 980 uL of ethanol/water (95% ethanol, 5% ultrapure water with pH adjusted to 4.9 with acetic acid). The LUNA porous silica beads of Example 1 were dispersed in an excess of silanizing solution for approximately 10 minutes, vortexing continuously. The particles were rinsed three times with ethanol and dried in a 120◦C oven, overnight for approximately 12 hours.

Stock solution of PDPO, poly(2,6-dimethyl-1,4-phenylene oxide),(Aldrich, Milwaukee, WI) and Nile Red was prepared from 0.09 g PDPO and 1.0 mL chloroform. After complete dissolution of the polymer, a 100 uL aliquot of 1mg/mL nile red in chloroform was added. The resultant solution was vortexed continuously for uniform dispersion.

Excess PDPO/Nile Red was added to a small fraction of the silanized porous beads, approximately 100 uL polymer/dye solution to approximately 1mg of beads. The sample was vortexed for approximately 3 hours then washed. Excess polymer dye was removed and the beads were then washed repeatedly with methylene chloride, two to three times, followed by a washing with 0.01% polyoxyethylene-sorbitan monolaurate, Tween 20 (J.T.Baker, Cleveland, OH), in water. The washed beads were collected in a solution of 0.01% Tween 20/ultrapure water. A single, small drop was placed on a microscope coverslip and allowed to dry protected from light.

Example 3

Preparation of non-porous silica/Nile Red beads coated with polysiloxane polymer: Commercially available non-porous 3.1 um silica beads (Bangs Laboratory, Fishers, IN) were first

silanized in excess silanizing solution, a 10% solution by volume of 3-(trimethoxysilyl)propyl methacrylate (Aldrich, Milwaukee, WI) in acetone, overnight. Excess silanizing solution was decanted and the beads were rinsed repeatedly, two to three times, with ultrapure acetone, vortexing and centrifuging between washes. The beads were soaked in excess Nile Red solution (1mg/ml in toluene) for approximately 3 hours while vortexing so as to fully saturate the surface. The bead solution was centrifuged and excess dye solution was decanted. A mixture of 7.9 mg benzoin ethyl ether (Polysciences Inc., Warrington, PA), 250 microliters stock Nile Red in toluene and 250 microliters (15-20% acryloxypropyl-methylsiloxane) 80-85% dimethylsiloxane copolymer (Gelest Inc., Tullytown PA) were then added to the beads. The bead suspension was vortexed to uniformly coat the particles. The resultant suspension mixture was added dropwise to approximately 100 mL 0.1% Tween 20 in ultrapure water stirring at approximately 350 rpm. Polymerization was accomplished by ultraviolet excitation for 10 second durations for a total exposure of 30 seconds. The sample solution was stirred over night. The suspension was passed through a 230 micron sieve, followed by a 5 um sieve. The filtrate was centrifuged at 3000 rpm for approximately 5 minutes and the beads were collected into centrifuge tubes and washed with 0.01% Tween 20 in ultrapure water. A single small drop was placed on a microscope coverslip and allowed to dry protected from light.

Example 4

Preparation of (15-20% acryloxypropylmethylsiloxane) 80-85% dimethylsiloxane copolymer beads with nile red: Approximately 25mL of ultrapure water plus 25 mL ethanol were placed in a 100mL round bottom flask and stirred with a stirbar at approximately 350 rpm. A mixture of 500 uL (15-20% acryloxypropylmethylsiloxane) 80-85% dimethylsiloxane copolymer, 200 uL Nile Red solution (1 mg/ml in chloroform) and 250 uL methylene chloride was made and added dropwise to the stirred water/ethanol solution. A solution of 5.5 mg AIBN, N,N'-azobis-isobutyl nitrile (2,2'-azobis-2-methylproprio-nitrile)(Phaltz & Bauer, Inc.), in methylene chloride was added to the stirring dispersion. The mixture was degassed with argon for approximately one hour and then heated to approximately 70 degrees celcius. After approximately three hours of heating, 20 mL of 0.01% Tween 20 in ultrapure water was added to the mixture. Heating and stirring was continued for approximately 12 hours. The mixture was passed through 230 micron sieve, then solids collected from centrifugation at up to 5000 rpm. The solids were washed twice with methanol and then washed with 0.01% Tween 20 in ultrapure water. The resultant beads were collected in a solution of 0.01% Tween 20 in ultrapure water. A single drop of the bead suspension was placed on a microscope coverslip and allowed to dry protected from light.

Example 5

Nile red dyed poly(methylstyrene/divinyl benzene) beads: Approximately 1 mg of commercially available 3.15 um polymer beads, 87% methyl styrene, 13% divinyl benzene with amine functionalized

surface (Bangs Laboratories, Fishers, IN), was washed in 1 ml of methanol by vortexing, centrifuging at approximately 3000 rpm and decanting the solvent. The beads were transferred to brown vial and approximately 100 uL of Nile Red solution (1 mg/ml in toluene) was added. The sample was vortexed and placed on a wrist shaker to agitate overnight. The suspension was transferred to a microcentrifuge tube and washed with methanol until the decanted solvent was clear. The beads were collected in approximately 0.5 mL of a solution of 0.01% Tween 20 in ultrapure water. A single drop placed on a microscope coverslip and allowed to dry protected from light.

Example 6

Plasticizer modified poly(methylstyrene/divinyl benzene) beads with nile red incorporated:

Approximately 1 mg of commercially available 3.15 um polymer beads, 87% methyl styrene, 13% divinyl benzene with amine functionalized surface (Bangs Laboratories, Fishers, IN), were rinsed with methanol according to Example 5 and transferred to a brown vial. Approximately 2-40% by wt plasticizer to polymer solutions of plasticizers, tritolyl phosphate (TTP), triphenyl phosphate (TPP),and dibutyl phthalate (DBP) (Aldrich, Milwaukee, WI), with nile red solution (1 mg/mL in toluene) were added to samples of beads, covered, vortexed then shaken on wrist shaker for approximately 12 hours. The beads were transferred to microcentifuge tubes and washed with Nile Red in methanol, then repeatedly with methanol until the decanted solvent was clear. The beads were collected in a solution of 0.01% Tween 20 in ultrapure water. A single drop of the suspension was placed on a microscope coverslip and allowed to dry protected from light.

Example 7

The porous silica beads prepared by the method of Example 1 were evaluated to determine their characteristic optical response signature to toluene vapor following the experimental method described above. The results are presented in Fig. 9 where the temporal optical response of 62 individual bead sensors to a pulse of toluene vapor is shown.

Example 8

The poly(methylstyrene/divinyl benzene) beads prepared by the method of Example 5 were evaluated to determine their characteristic optical response signature to methanol vapor. The results are presented in Fig. 10 where the temporal optical response of 39 individual bead sensors to a pulse of methanol vapor is shown.

Example 9

The (15-20% acryloxypropylmethylsiloxane) 80-85% dimethylsiloxane copolymer beads prepared by the method of Example 4 were evaluated to determine their characteristic optical response signature to both toluene and methanol vapor. The results are presented in Fig. 11 where the temporal optical responses of an individual bead sensor to a pulse of toluene and a pulse of methanol vapor is shown.

Example 10

The PDPO polymer coated porous silica beads prepared by the method of Example 2 were evaluated to determine their characteristic optical response signature to both toluene and methanol vapor. The results are presented in Fig. 12 where the temporal optical responses of an individual bead sensor to a pulse of toluene and a pulse of methanol vapor is shown.

Example 11

Porous silica beads prepared by the method of Example 1 were incorporated into etched microwells on the distal end of a fiber optic bundle according to the method described above. The resultant sensor array was evaluated to determine the characteristic optical response signature of the bead subpopulation to ethyl acetate vapor. The results are presented in Fig. 13 where the temporal optical response of 218 individual bead sensors to a pulse of ethyl acetate vapor is shown.

Example 12

The signal summing method of the present invention was evaluated in analyzing the experimental measurements made on poly(methylstyrene /divinyl benzene) beads prepared by the method of Example 5 and tested by the method of Example 8. The results are shown in Fig. 14 where the normalized temporal optical response for a single sensor bead, Bead #1, is compared with the summed responses of all 39 beads tested.
As shown by Fig. 14, the signal summing method of the present invention significantly reduces the experimental noise encountered in a single sensor bead measurement and provides a substantial improvement, ten-fold or greater, in the signal-to-noise ratio of analytical measurements.

Example 13

The signal summing method of the present invention was evaluated in analyzing the experimental measurements made on poly(methylstyrene /divinyl benzene) beads prepared by the method of Example 5 and tested by the method of Example 8. The results are shown in Fig. 15 where the actual relative intensities of the temporal optical response for each of the 39 sensor beads is compared to relative intensity of the temporal optical response obtained from signal summing. As shown by Fig.

15, substantial signal enhancement is obtained by signal summing with a correspondingly significant improvement, up to a hundred fold, in the detection limit for target analytes.

Example 14

The polysiloxane coated porous silica beads prepared by the method of Example 3 were evaluated to determine their characteristic optical response signature to both toluene and methanol vapor. The results are presented in Fig. 16 where the temporal optical responses of two bead sensors to both toluene and methanol are shown. The results shown in Fig. 16 demonstrates the capability of this subpopulation of bead sensors to distinguish between two analytes of interest by utilizing the characteristic optical response signatures of the bead sensors to specific analytes.

Example 15

A 50/50 mixture of porous silica beads prepared by the method of Example 1 and poly(methylstyrene / divinyl benzene) beads prepared by the method of Example 5 were randomly dispersed and incorporated into etched microwells on the distal end of a fiber optic bundle according to the method of the present invention as described above. The resultant sensor array was evaluated to determine the characteristic optical response signature of the bead subpopulation to methanol vapor. An 535 nm excitation filter and 600 nm emission filter was used in this experiment. The results are presented in Fig. 17 where the normalized temporal optical response of 3 porous silica bead sensors and 6 PMS bead sensors to a pulse of methanol vapor is shown. In this example, the characteristic emitted light peak shapes of the bead subpopulations provide a distinguishable characteristic response signature for each subpopulation. Fig. 17 demonstrates the innovative self-encoding feature of the present invention where the identity and location of the beads is determined in a single measurement of a reference vapor analyte.

Example 16

The self-encoded fiber optic sensor array produced by the method of Example 15 was evaluated by measuring the characteristic temporal optical response signature of the porous silica and PMS sensor bead subpopulations of the array in response to a pulse of n-propanol vapor. The results are presented in Fig. 18 where the temporal optical response of 3 porous silica bead sensors and 6 PMS bead sensors to a pulse of n-propanol vapor is shown. In this example, the characteristic emitted light intensities of the bead subpopulations provide a distinguishable characteristic response signature for each subpopulation. Fig. 18 demonstrates the advantages of using the distinct characteristic temporal optical response signature of different bead subpopulations to detect a specific analyte of interest. Note that the identity and location of the bead sensors in the sensor array was decoded by the method

of Example 15. By the combination of self-encoding the sensor array by the method of Example 15 and the sensor array measurement made by the method of the current Example 16, the sensor array was trained to detect and distinguish methanol from n-propanol.

Example 17

The self-encoded fiber optic sensor array produced by the method of Example 15 was evaluated by measuring the characteristic temporal optical response signature of the porous silica and PMS sensor bead subpopulations of the array in response to a pulse of toluene vapor. The results are presented in Fig. 19 where the temporal optical response of 3 porous silica bead sensors and 6 PMS bead sensors to a pulse of toluene vapor is shown. Fig. 19 demonstrates the advantages of using the characteristic temporal optical response signature of different bead subpopulations to detect a specific analyte of interest. Note that the identity and location of the bead sensors in the sensor array was decoded by the method of Example 15. By the combination of decoding the self-encoding the sensor array by the method of Example 15, the sensor array measurement made by the method of Example 16, and the sensor array measurement made by the method of the current Example 17, the sensor array was trained to detect and distinguish between the group of target analytes comprising methanol, n-propanol, and toluene.

Example 18

Samples of PS802 bead sensors produced by the method of Example 4, Poly methyl styrene/2% divinyl benzene bead sensors produced by the method of Example 5, and commercially available poly methyl styrene beads (Bangs Laboratory, Fishers, IN) were dispersed on a microscope coverslip substrate. Following equilibration of each bead subpopulation in air, each subpopulation was exposed to a pulse of saturated toluene vapor while illuminating the beads with excitation light energy. The changes in bead dimension due to the swelling response of each polymer type to toluene vapor was monitored using the apparatus of Fig. 7. The response of the bead was recorded by filming the time varying fluorescence image of the beads and capturing changes in bead image dimensions with a CCD camera. Fig. 20 illustrates the differences in swelling response of the three bead subpopulations by comparing the initial fluorescence image of each bead type in air with subsequent image of each bead type following exposure to toluene vapor. Such measurements of the swelling response characteristics of various polymer candidate materials is useful in prescreening bead sensor materials for use as bead sensor elements in the self-encoded sensor array of the present invention.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

PAGE INTENTIONALLY LEFT BLANK

CLAIMS

What is claimed is:

1. A self-encoding analytic chemical sensor array comprising:
    a) a substrate with a surface comprising discrete sites; and
    b) a population of microspheres comprising at least a first and a second subpopulation, wherein each subpopulation comprises at least one reporter dye;
    wherein each subpopulation emits a first characteristic optical response signature when subjected to excitation light energy in the presence of a reference analyte;
    wherein said microspheres are distributed on said surface.

2. A sensor array according to claim 1 wherein each subpopulation further comprises a bioactive agent.

3. A sensor array according to claim 1 or 2 wherein the reporter dye comprises a fluorescent dye.

4. A sensor array according to claim 1 or 2 wherein the reporter dye comprises a solvatochromic dye.

5. A sensor array according to claim 4 wherein the solvatochromic dye comprises Nile Red.

6. A sensor array according to claim 1, 2, 3, 4 or 5 wherein the beads are encoded with a predetermined ratio of at least two reporter dyes.

7. A sensor array according to claim 2 wherein said bioactive agent is selected from the group consisting of nucleic acids and proteins.

8. A sensor array according to claims 1, 2, 3, 4, 5, 6 and 7 wherein said substrate is a fiber optic bundle and said surface is a proximal end of said bundle.

9. A sensor array according to claims 1, 2, 3, 4, 5, 6, 7 and 8 further comprising an excitation light energy source in optical communication with said proximal end.

10. A sensor array according to claims 1, 2, 3, 4, 5, 6, 7, 8 and 9 further comprising an emission light energy detection means in optical communication with said proximal end.

11. A method of detecting a target analyte in a sample comprising:
    a) contacting said sample with an sensor array comprising:
        i) a substrate with a surface comprising discrete sites; and

ii) a population of microspheres comprising at least a first and a second subpopulation, each subpopulation comprising:

1) a bioactive agent; and

2) at least one reporter dye;

wherein said reporting dye has a first characteristic optical response signature when subjected to excitation light energy in the presence of a reference analyte;

wherein said microspheres are distributed on said surface;

b) detecting the presence of said analyte.

12. A method according to claim 11 further comprising identifying the location of each bioactive agent on said substrate by adding said reference analyte.

13. A method according to claim 11 or 12 wherein said detecting is done by detecting the presence of a label attached to said target analyte.

14. A method of decoding an array composition comprising

a) providing an array composition comprising:

i) a substrate with a surface comprising discrete sites; and

ii) a population of microspheres comprising at least a first and a second subpopulation, wherein each subpopulation comprises at least one reporter dye;

wherein said reporting dye has a first characteristic optical response signature when subjected to excitation light energy in the presence of a reference analyte;

wherein said microspheres are distributed on said surface; and

b) adding at least one reference analyte to said array composition to identify the location of at least one subpopulation.

15. A method according to claim 14 wherein the location of each subpopulation is determined.

16. A self-encoding analytic chemical sensor array comprising:

a population of beads dispersed on a substrate, said population encoded with at least one reporting dye;

a plurality of separate subpopulations of beads contained within said population, each subpopulation of beads comprising a characteristic bead matrix material infiltrated with at least one reporting dye, which said reporting dye, in combination with said characteristic bead matrix material, has a first characteristic optical response signature when subjected to excitation light energy in the presence of a reference analyte and, which said reporting dye, in

combination with said characteristic bead matrix material, has a second characteristic optical response signature when subjected to excitation light energy in the presence of an unknown target analyte;

wherein both the identity and the location of each bead comprising said bead subpopulations in said sensor array is determined and recorded upon exposure of said sensor array to said reference analyte while subjecting said bead to excitation light energy.

17. The sensor array described in claim 16, wherein the reporter dye comprises a fluorescent dye.

18. The sensor array described in claim 17, wherein the reporter dye comprises a solvatochromic dye and the dye is entrapped within the bead matrix material.

19. The sensor array described in claim 18, wherein the solvatochromic dye comprises Nile Red.

20. The sensor array described in claim 16, wherein the beads are encoded with a predetermined ratio of at least two reporter dyes.

21. The sensor array described in claim 16, further comprising a chemical functional roup attached to a surface of at least one subpopulation of beads.

22. The sensor array of claim 21 wherein said reporting dye provides only said first characteristic optical response signature to decode said array and said second characteristic optical response signature in the presence of an unknown target analyte is provided by an additional reporter dye which responds only in the presence of an unknown target analyte.

23. The sensor array described in claim 21, wherein the addition of said chemical functional group modifies the characteristic optical response signature of said bead subpopulation, thereby creating a new self-encoding bead subpopulation in said array.

24. The sensor array described in claim 21, wherein the chemical functionalities of the beads support sites for hybridization.

25. The sensor array described in claim 16, wherein the substrate comprises a distal end of an optical fiber bundle.

26. The sensor array of claim 16, wherein the substrate comprises etched microwells at a distal end of individual fibers in an optical fiber bundle.

27. The sensor array of claim 26, wherein the etched wells form a cavity which is larger than the bead diameter and the bead is essentially disposed within said cavity.

28. The sensor array of claim 27 wherein said cavity is no greater than 5 um in diameter.

29. The sensor array of claim 16 wherein the average bead diameters are less than 4 um.

30. The sensor array of claim 16 wherein the characteristic bead matrix material is an organic material selected from the group consisting of polymers and copolymers.

31. The sensor array of claim 30 wherein the organic material swells upon contact to a target analyte.

32. The sensor array of claim 16 wherein the characteristic bead matrix material is an inorganic material selected from the group consisting of porous silicas, aluminas
and zeolites.

33. A method for reducing the signal-to-noise ratio in the characteristic optical response signature of a sensor array having a subpopulations of array elements, comprising:
decoding the array so as to identify the location of each sensor element within each sensor

subpopulation within the array;

measuring the characteristic optical response signature of each sensor element in the array;

adjusting the baseline of said optical response signature for each sensor element in said array;

summing the baseline-adjusted characteristic optical response signature of all sensor elements within each of said sensor subpopulations; and

reporting the characteristic optical response signature of each sensor subpopulation as a summation of said baseline-adjusted characteristic optical response signatures of all sensor elements within each of said subpopulations.

34. The method of claim 33, wherein the signal-to-noise ratio is increased by a factor of at least ten.

35. The method of claim 34, wherein an analyte detection limit is reduced by a factor of at least 100.

36. A method for amplifying the characteristic optical response signature of a sensor array having subpopulations of array elements, comprising:

decoding the array so as to identify the location of each sensor element within each sensor subpopulation within the array;

measuring a characteristic optical response signature of each sensor element in the array;

adjusting the baseline of said optical response signature for each sensor element in said array;

summing the baseline-adjusted characteristic optical response signature of all sensor elements within each of said sensor subpopulations; and

reporting the characteristic optical response signature of each sensor subpopulation as a summation of said baseline-adjusted characteristic optical response signatures of all sensor elements within each of said subpopulations.

37. The method of claim 36, wherein the signal is amplified by a factor of at least fifty.

38. The method of claim 37, wherein an analyte detection limit is reduced by a factor of at least 100.

39. A sensing apparatus comprising:

a sensor array which comprises, a population of beads dispersed on a substrate, said population encoded with at least one reporting dye, a plurality of separate subpopulations of beads contained within said population, each subpopulation of beads comprising a characteristic bead matrix material infiltrated with at least one reporting dye, which said reporting dye, in combination with said characteristic bead matrix material, has a first characteristic optical response signature when subjected to excitation light energy in the presence of a reference analyte and, which said reporting dye, in combination with said characteristic bead matrix material, has a second characteristic optical response signature when subjected to excitation light energy in the presence of an unknown target analyte, wherein both the identity and the location of each bead comprising said bead subpopulations in said sensor array is determined and recorded upon exposure of said sensor array to said reference analyte while subjecting said bead to excitation light energy;

an optical fiber bundle, comprising a plurality of discrete individual fibers, each of said fibers in said bundle optically cooperating with and in optical communication with a discrete sensor bead in said sensor array;

an excitation light energy source in optical communication with a proximal end of said optical fiber bundle, said light source providing excitation light energy to said proximal end and said fiber bundle conveying said excitation light energy through said discrete fibers to a plurality of discrete sensor beads, wherein the light conveyed by each fiber is coupled to a specific bead element in said sensor array; and

an emission light energy detection means in optical communication with said proximal end of said fiber bundle, said detection means measuring light emitted from each fiber in said bundle at said proximal end in response to an analyte, wherein the emitted light which is conveyed by each fiber is produced from the response of a discrete sensor bead coupled to said fiber at a distal end of said fiber.

40. The sensing apparatus of claim 39 further comprising a neural network system for associating the emitted light response of said discrete sensor beads and a characteristic

image of the emitted light response of the entire sensor array to a target analyte for the detection and identification of said analyte.

41. A sensing method comprising:

assembling a sensor array which comprises, a population of beads dispersed on a substrate, said population encoded with at least one reporting dye, a plurality of separate subpopulations of beads contained within said population, each subpopulation of beads comprising a characteristic bead matrix material infiltrated with at least one reporting dye, which said reporting dye, in combination with said characteristic bead matrix material, has a first characteristic optical response signature when subjected to excitation light energy in the presence of a reference analyte and, which said reporting dye, in combination with said characteristic bead matrix material, has a second characteristic optical response signature when subjected to excitation light energy in the presence of an unknown target analyte, wherein both the identity and the location of each bead comprising said bead subpopulations in said sensor array is determined and recorded upon exposure of said sensor array to said reference analyte while subjecting said bead to excitation light energy;

coupling said sensor array with an optical fiber bundle, comprising a plurality of discrete individual fibers, each of said fibers in said bundle optically cooperating with and in optical communication with a discrete sensor bead in said sensor array;

illuminating a proximal end of said optical fiber bundle with an excitation light energy source in optical communication with said proximal end, said light source providing excitation light energy to said proximal end and said fiber bundle;

conveying said excitation light energy through said discrete fibers of said bundle to a plurality of discrete sensor beads, each of said sensor beads positioned at a distal end of each of said fibers, wherein the light conveyed by each fiber is coupled to a specific bead element in said sensor array; and

detecting the light emitted from each fiber in said bundle by an emission light energy detection means in optical communication with said proximal end of said fiber bundle, said detection means measuring light emitted from each fiber in said bundle at said proximal

end in response to an analyte, wherein the emitted light which is conveyed by each fiber is produced from the response of a discrete sensor bead coupled to said fiber at a distal end of said fiber.

42. The sensing method of claim 41 further comprising a applying a neural network system for associating the emitted light response of said discrete sensor beads and a characteristic image of the emitted light response of the entire sensor array to a target analyte for the detection and identification of said analyte.

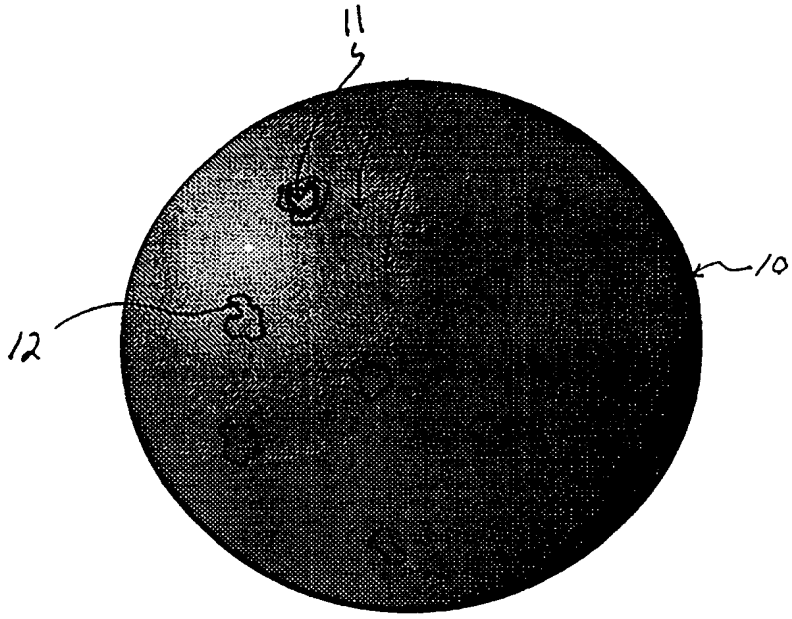


Fig. 1

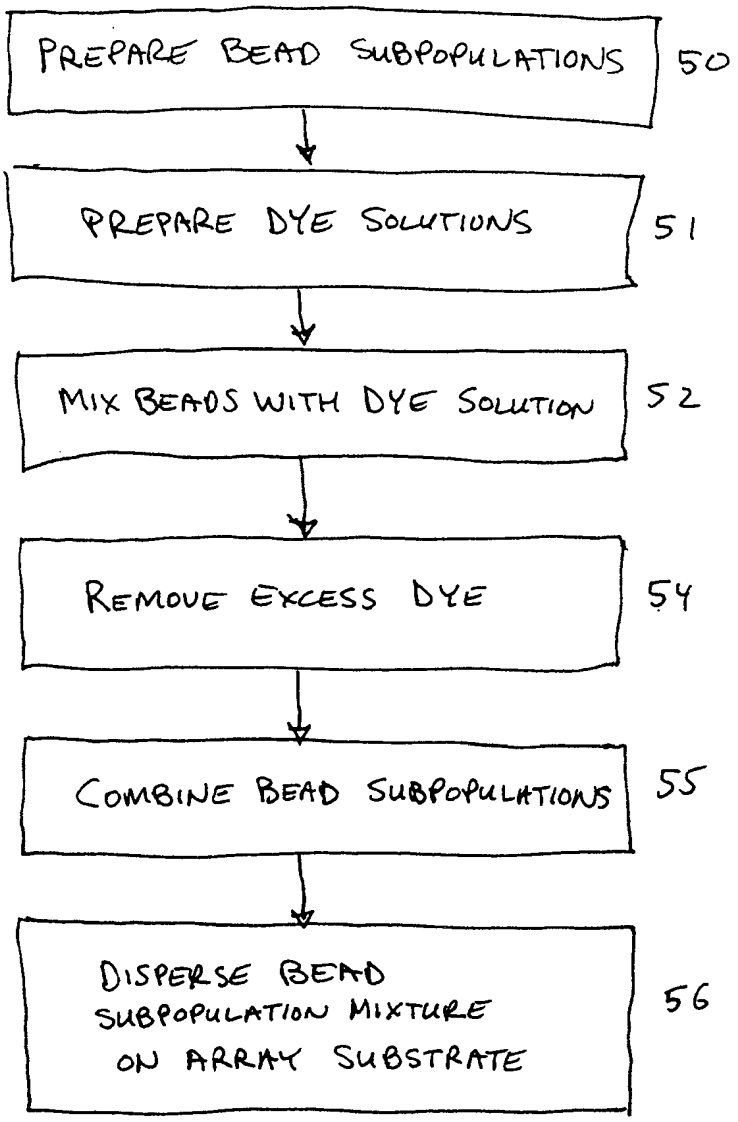

PREPARE BEAD SUBPOPULATIONS
50
PREPARE DYE SOLUTIONS
51
MIX BEADS WITH DYE SOLUTION
52
REMOVE EXCESS DYE
54
COMBINE BEAD SUBPOPULATIONS
55
DISPERSE BEAD SUBPOPULATION MIXTURE ON ARRAY SUBSTRATE
56

Fig. 2

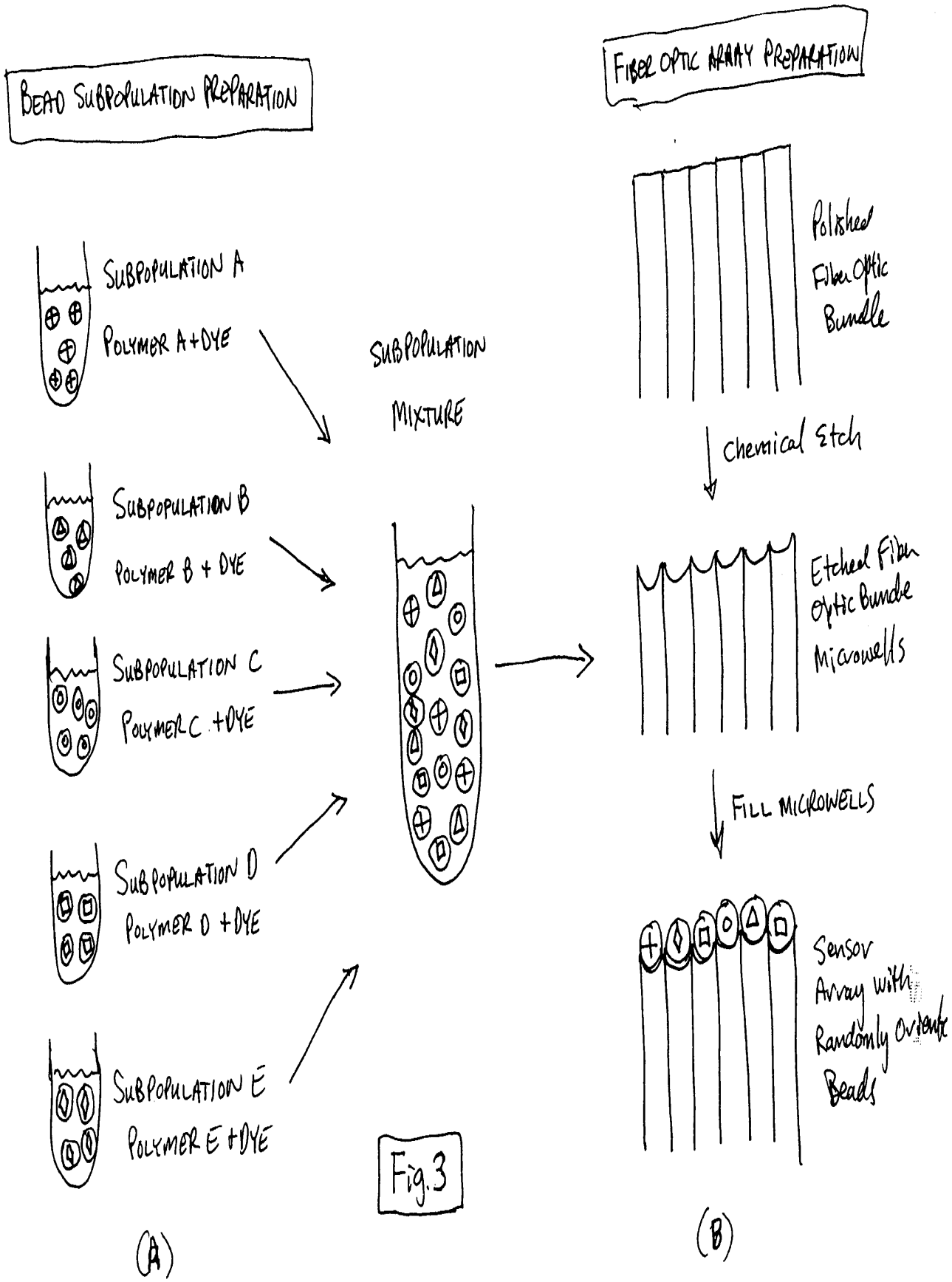

BEAD SUBPOPULATION PREPARATION
SUBPOPULATION A
POLYMER A + DYE
SUBPOPULATION B
POLYMER B + DYE
SUBPOPULATION C
POLYMER C + DYE
SUBPOPULATION D
POLYMER D + DYE
SUBPOPULATION E
POLYMER E + DYE
SUBPOPULATION MIXTURE
(A)
FIBER OPTIC ARRAY PREPARATION
Polished Fiber Optic Bundle
Chemical Etch
Etched Fiber Optic Bundle Microwells
FILL MICROWELLS
Sensor Array with Randomly Distributed Beads
(B)

Fig. 3

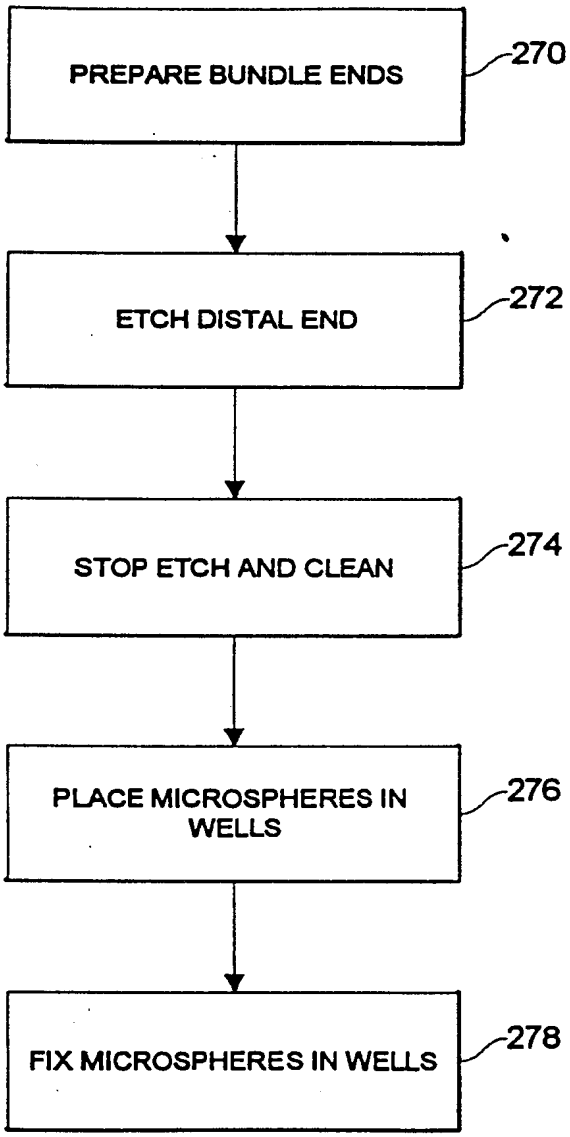

PREPARE BUNDLE ENDS
270
ETCH DISTAL END
272
STOP ETCH AND CLEAN
274
PLACE MICROSPHERES IN WELLS
276
FIX MICROSPHERES IN WELLS
278

Fig. 4

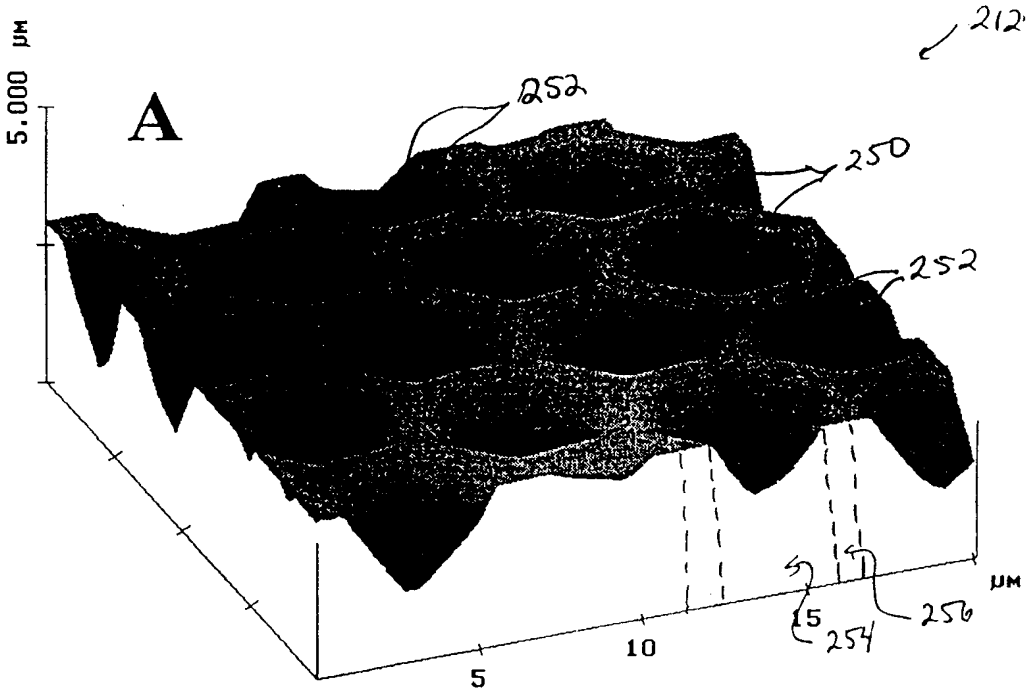

212
A
5.000 µM
252
250
252
15
254
256
5
10
µM

Fig. 5A

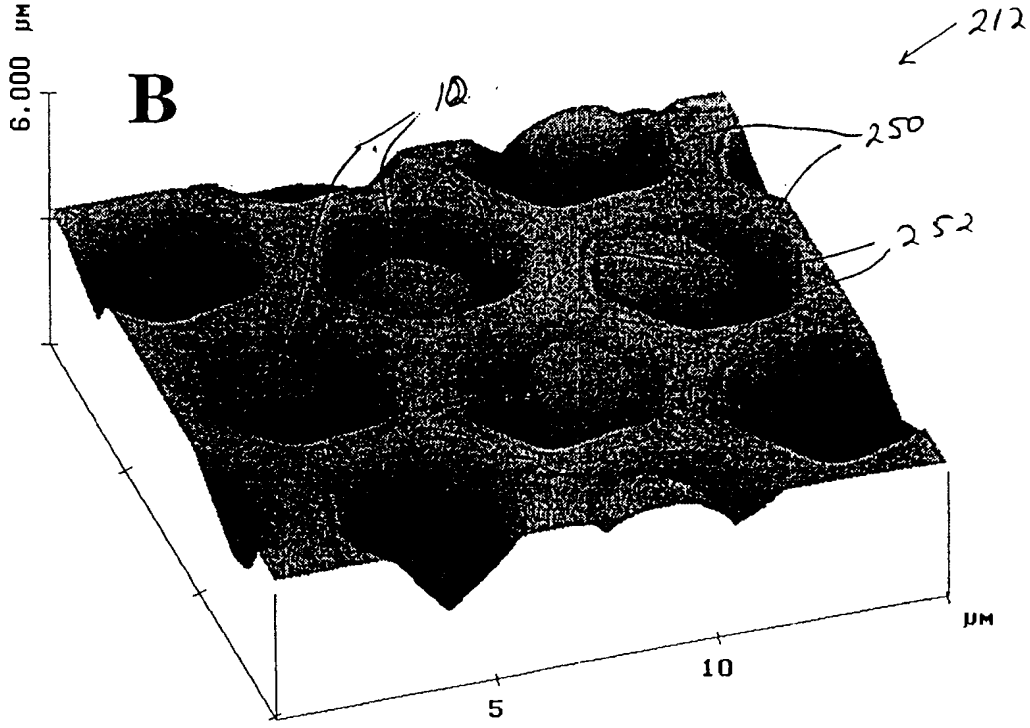
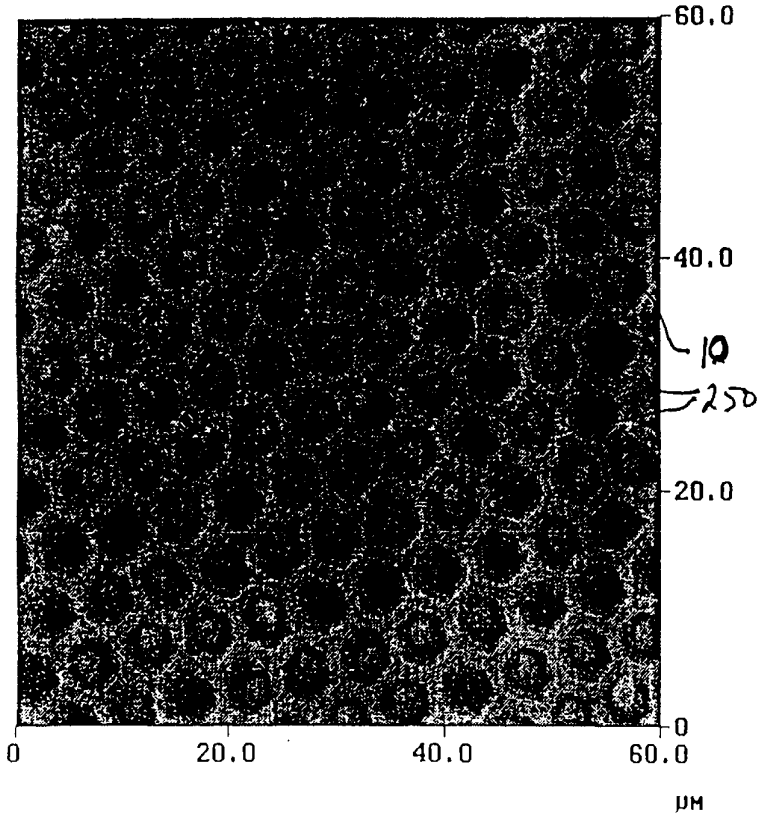

212
B
6.000 µM
10
250
252
5
10
µM

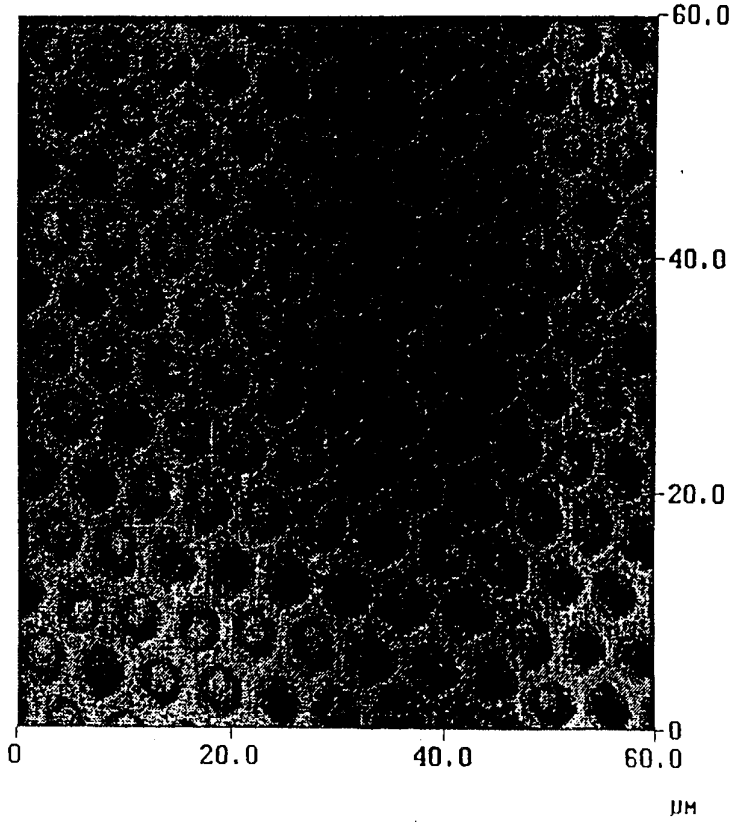
Fig 5B

Polymer Coated Beads in Wells After Air Pulse and Tapping



Before Tapping

Fig. 6A



After Tapping

Fig. 6B

631

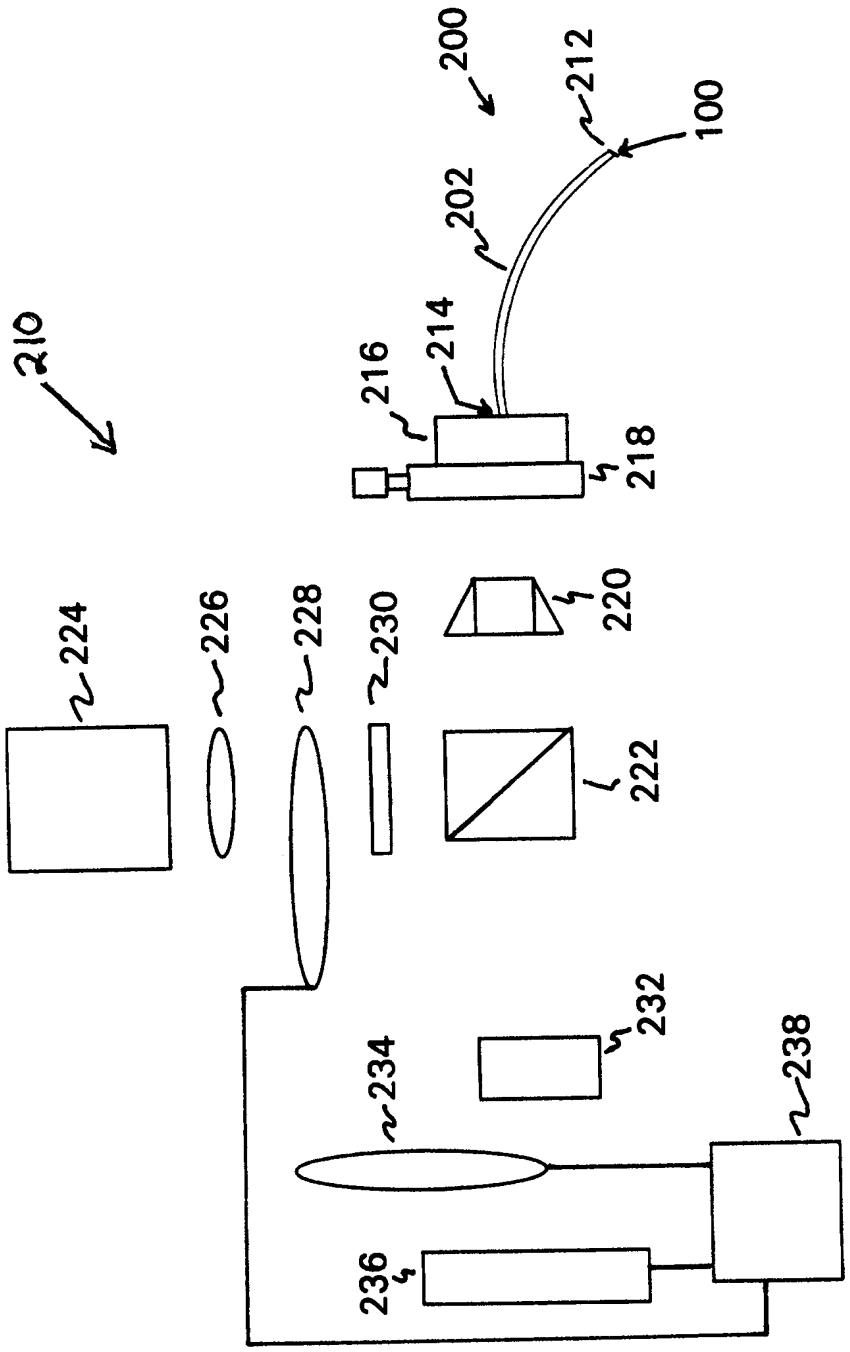

200
212
100
202
214
216
218
210
220
224
226
228
230
222
232
234
238
236

FIG. 7

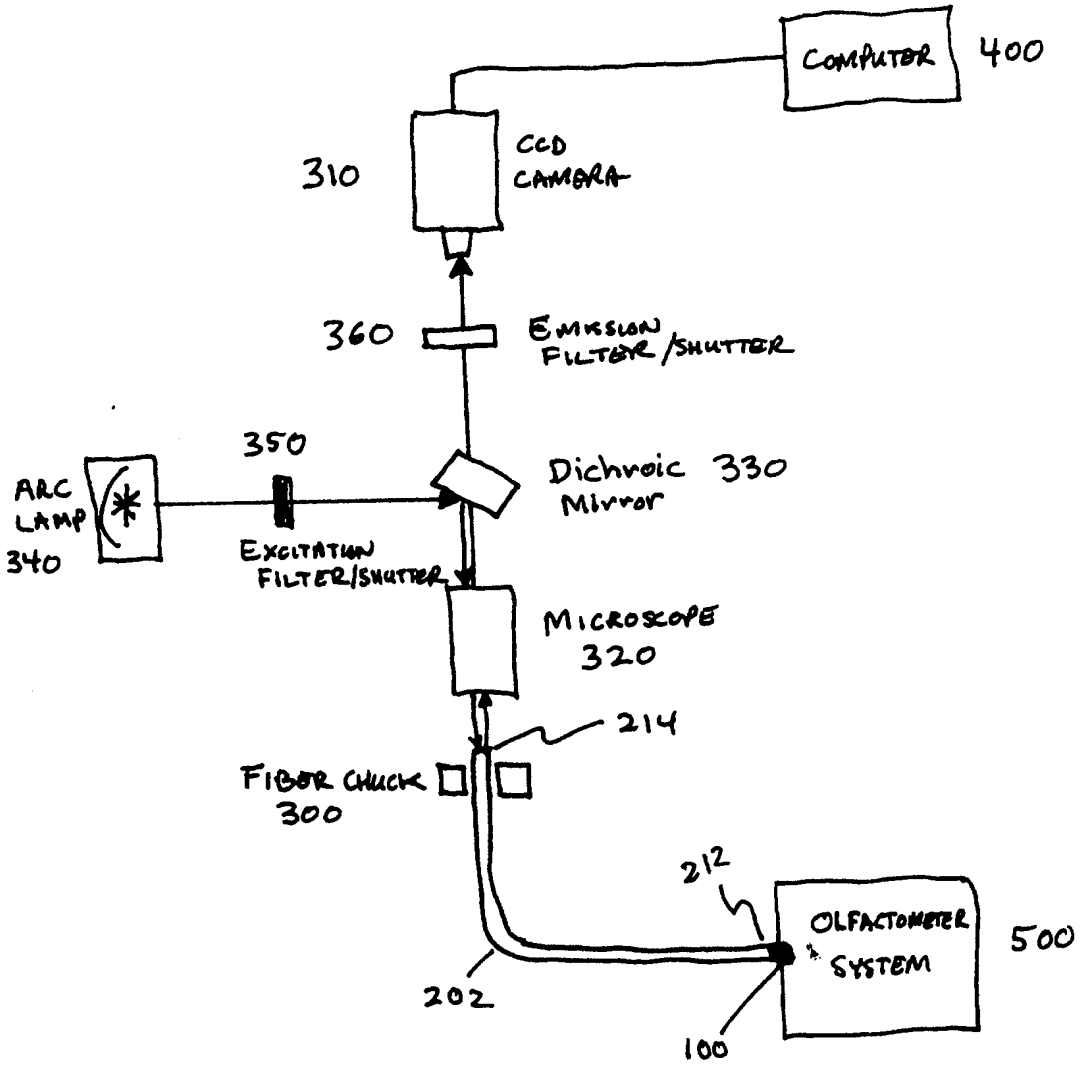

COMPUTER
400
310
CCD CAMERA
360
EMISSION FILTER/SHUTTER
350
ARC LAMP
340
EXCITATION FILTER/SHUTTER
Dichroic Mirror
330
MICROSCOPE
320
214
FIBER CHUCK
300
212
OLFACTOMETER SYSTEM
500
202
100

FIG. 8

Fig. 9

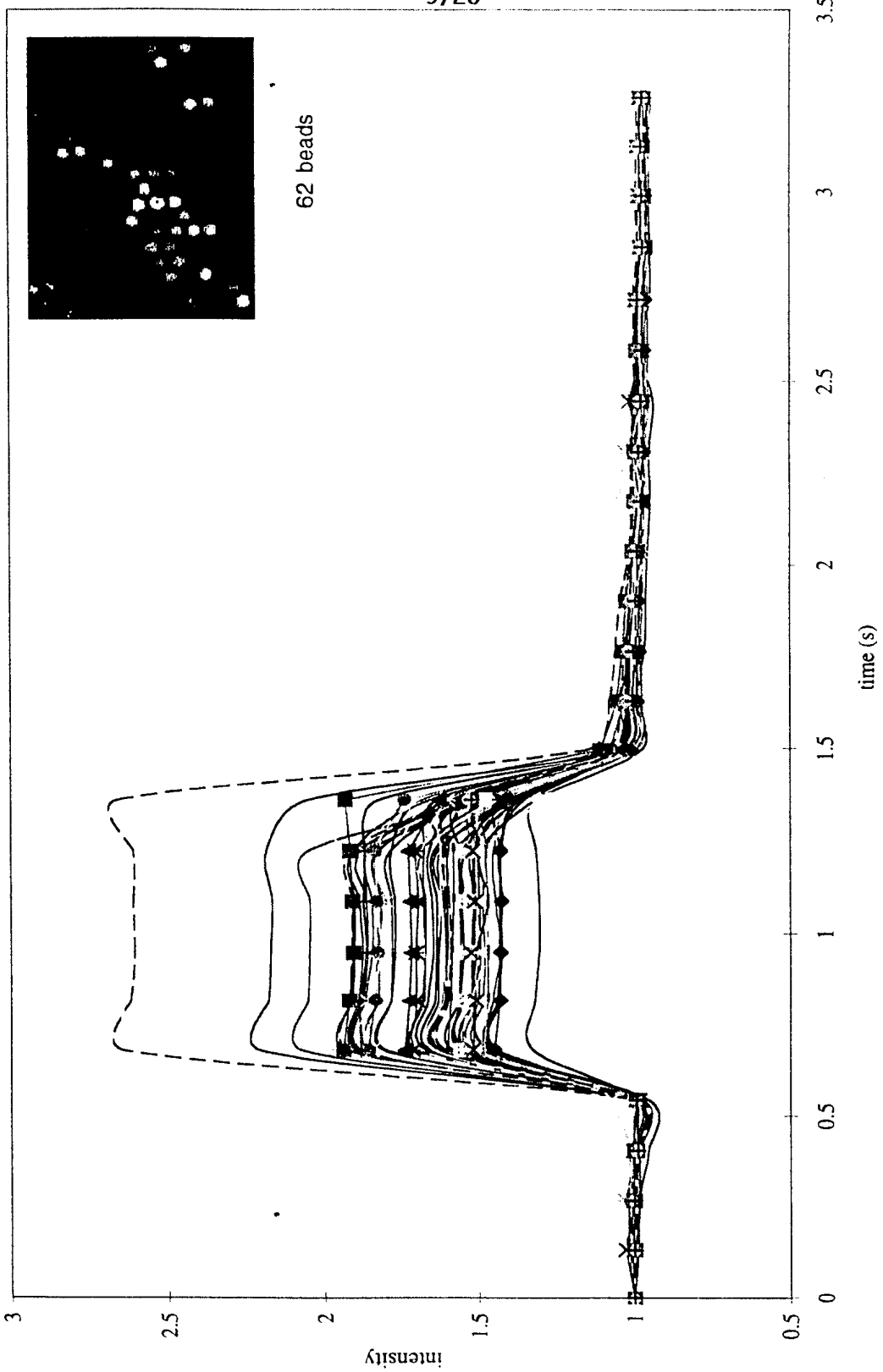

Porous 3µm silica beads high-speed response to Saturated Toluene
62 beads
intensity
time (s)
3
2.5
2
1.5
1
0.5
0
0.5
1
1.5
2
2.5
3
3.5

Fig. 10

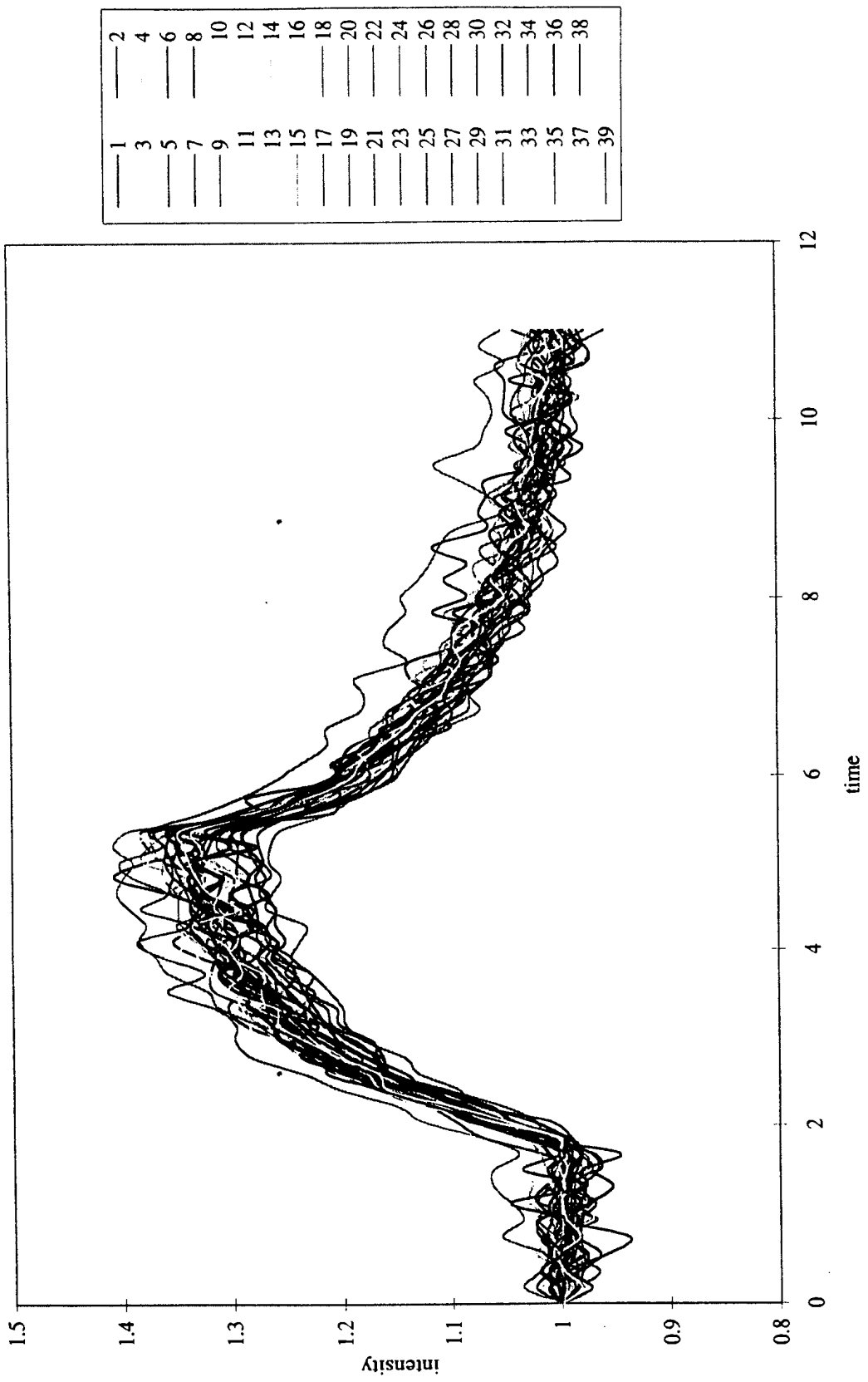

PMS Beads in Fiber: Response to Methanol (sat)
39 beads, mid-fiber region: centroid ovals, standardized responses
intensity
time
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0
2
4
6
8
10
12
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

Fig. 11A/11B

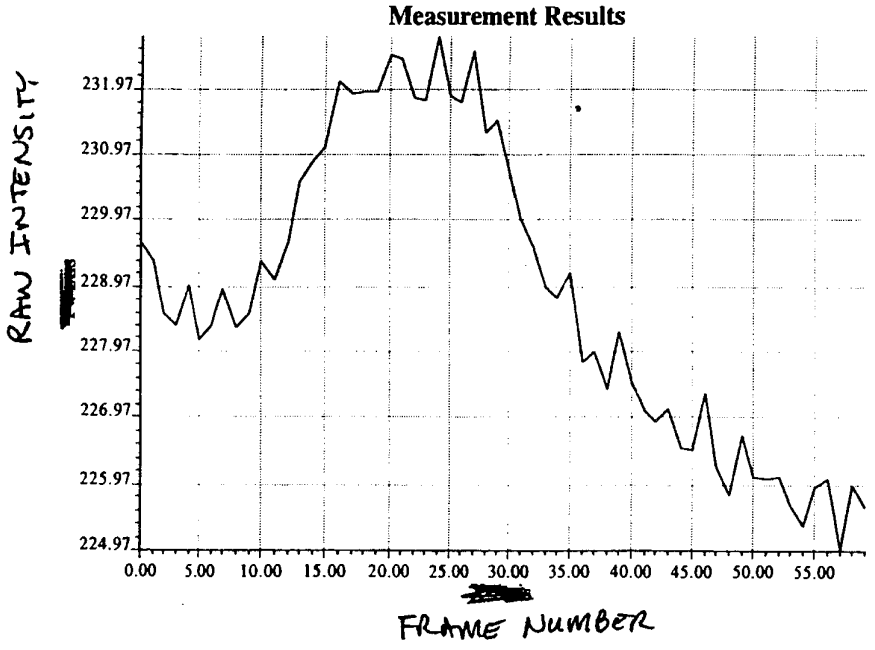
PS802 648.c Beads

bead 1

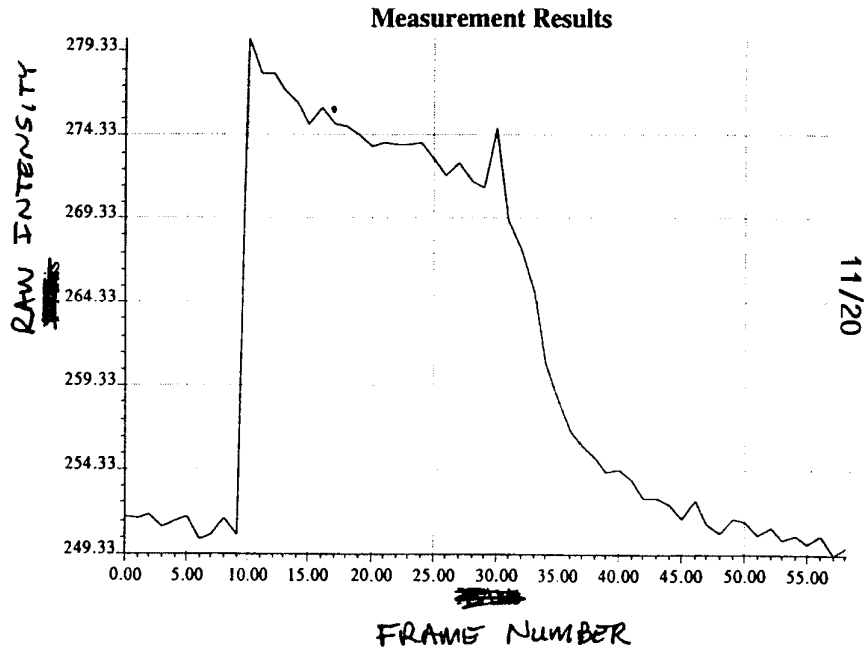
Saturated Toluene



Saturated Methanol



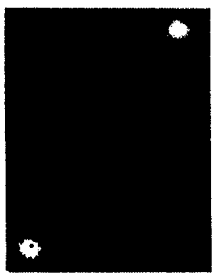
(A)



(B)

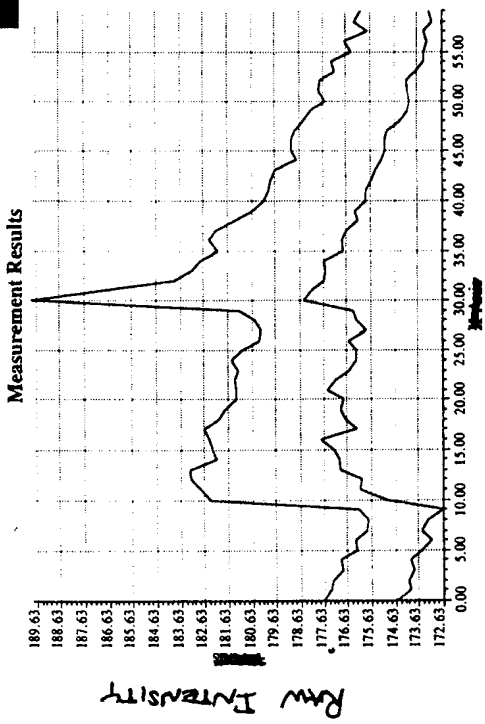
Fig. 12A/12B

PDPO/psil 9/11

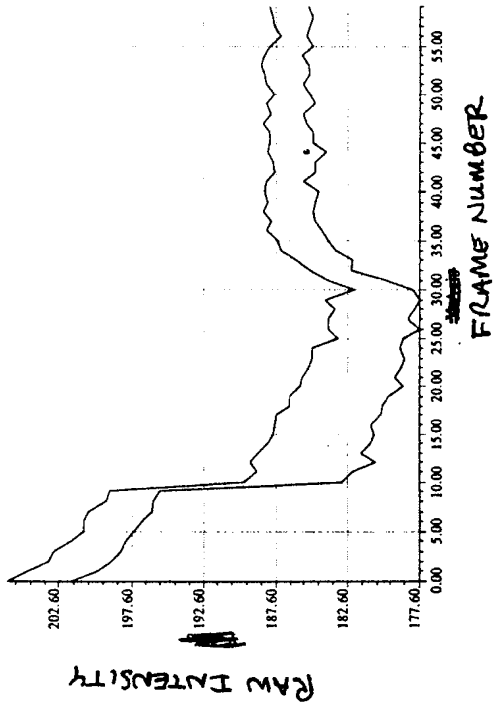


tol

met



(A)



(B)

Fig. 13

Porous 3µm silica beads response to Ethyl Acetate

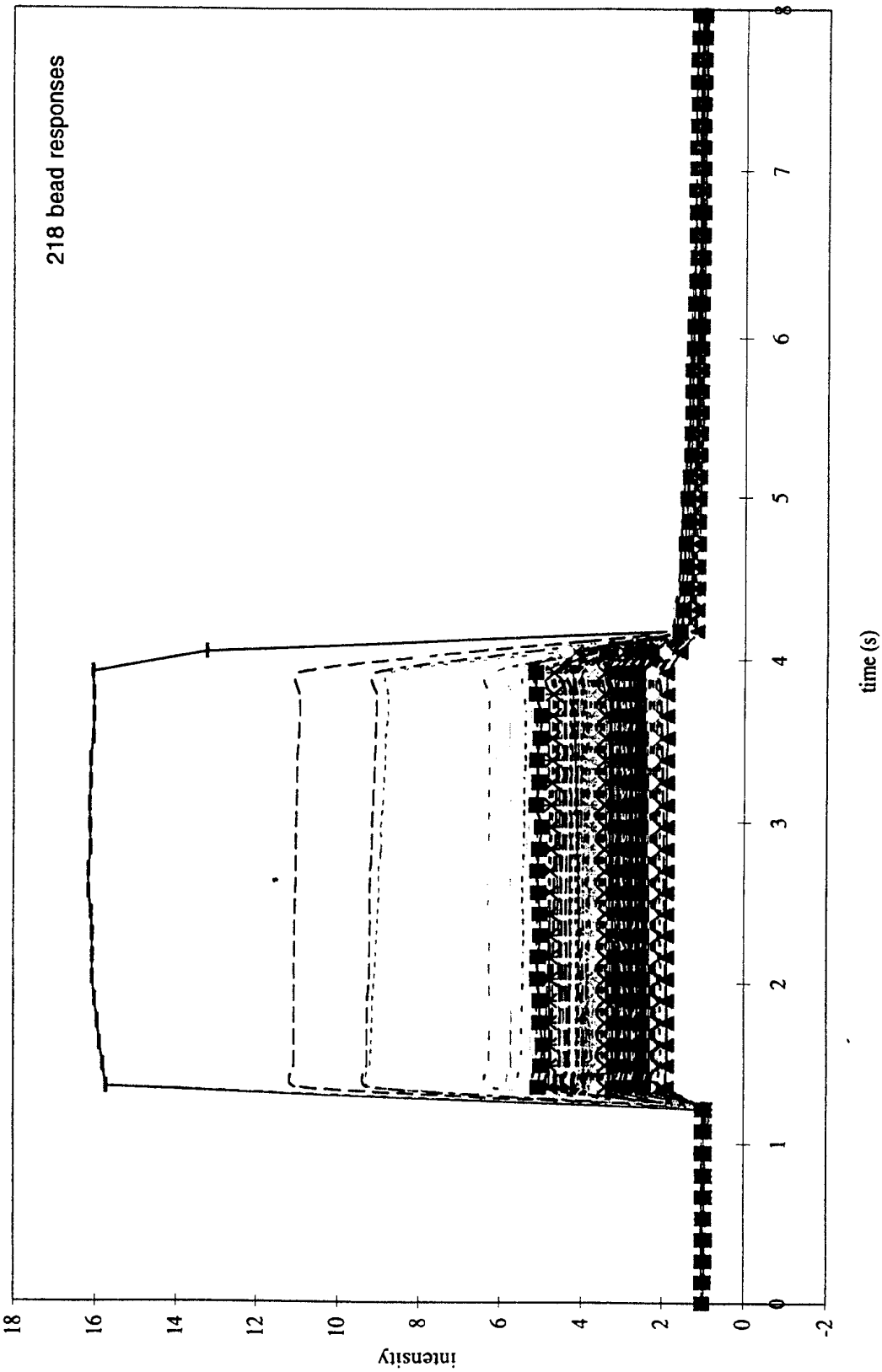


Fig. 14

Normalized signal-to-noise comparison between bead #1 and summed responses of 39 beads

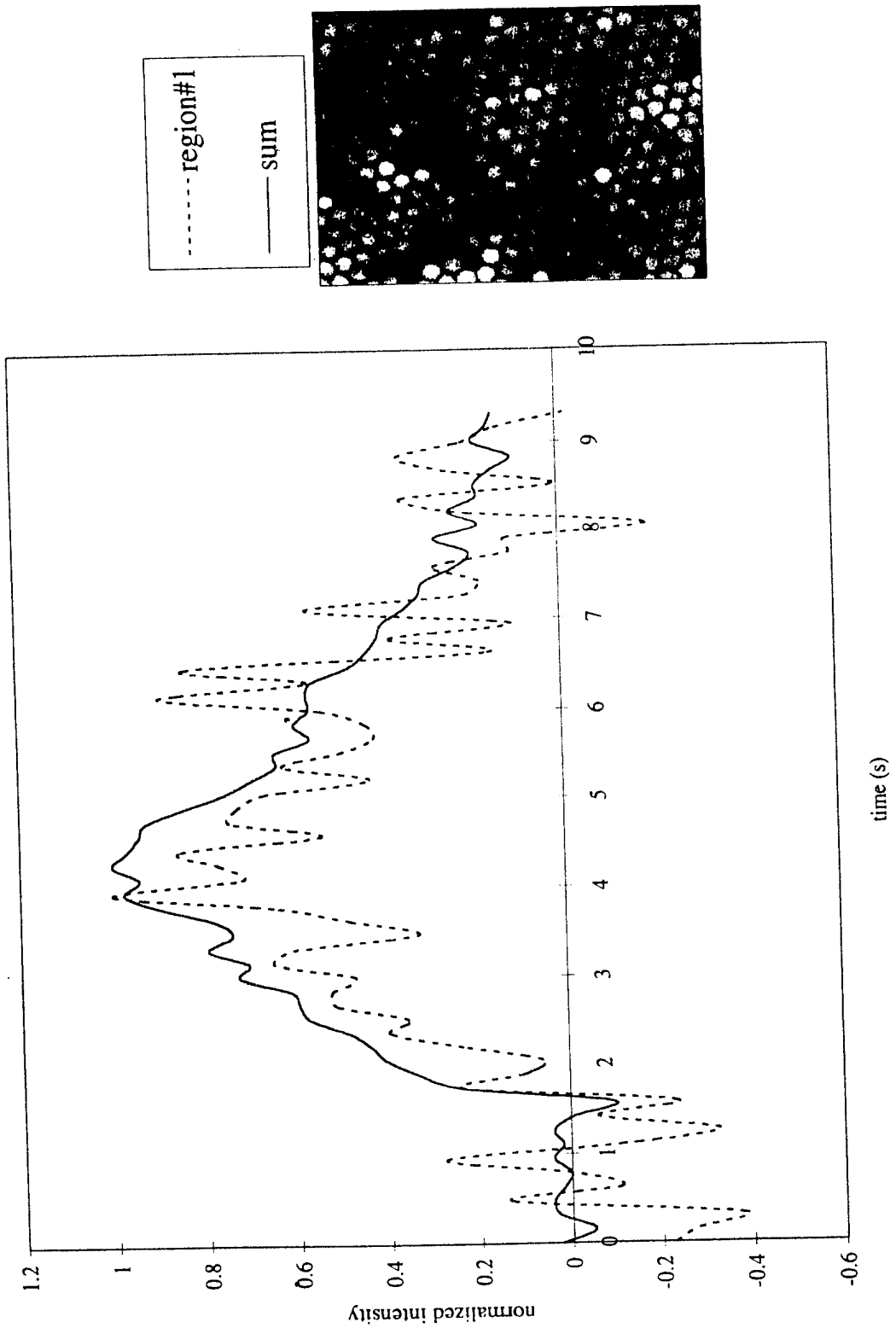


Fig. 15

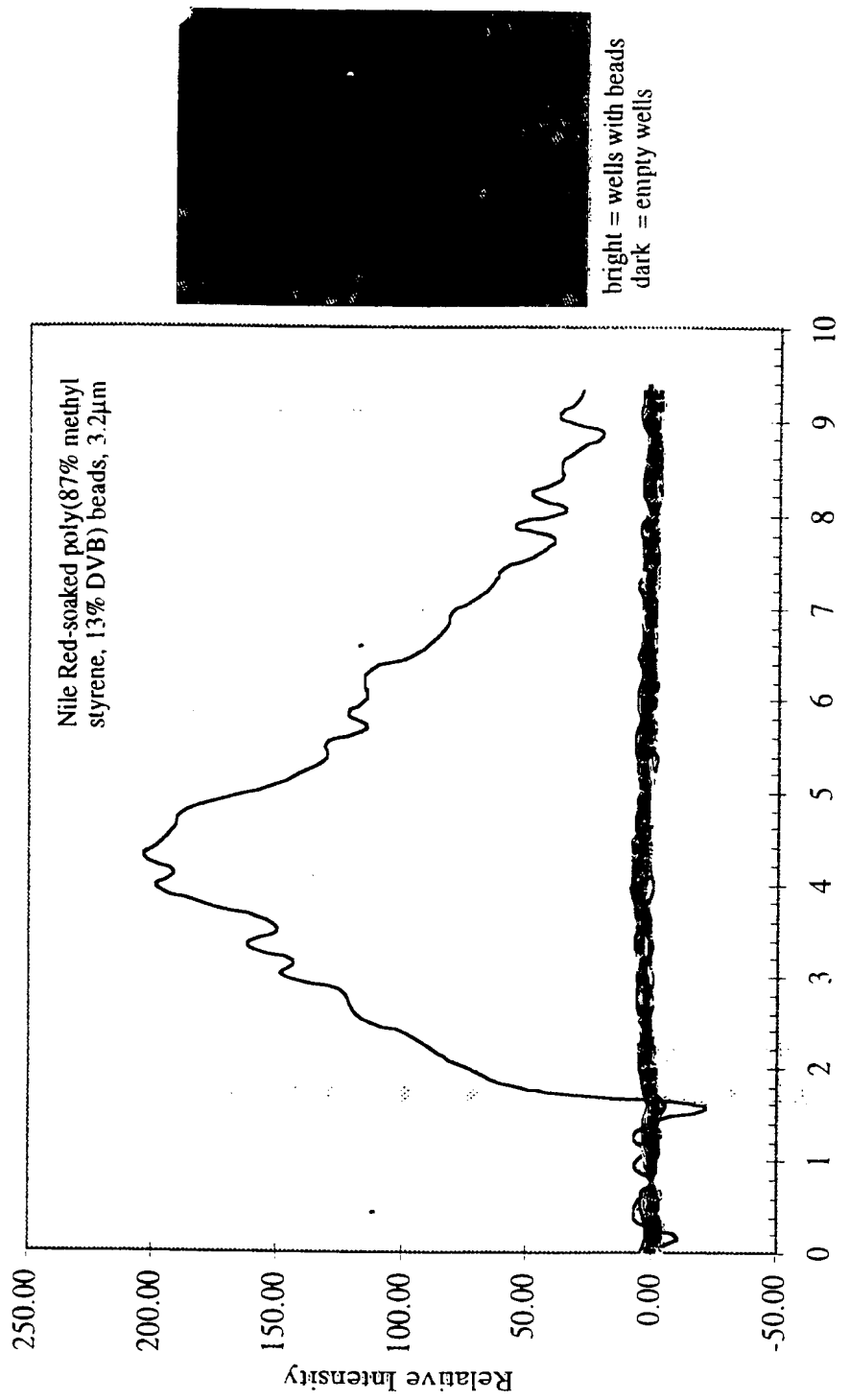
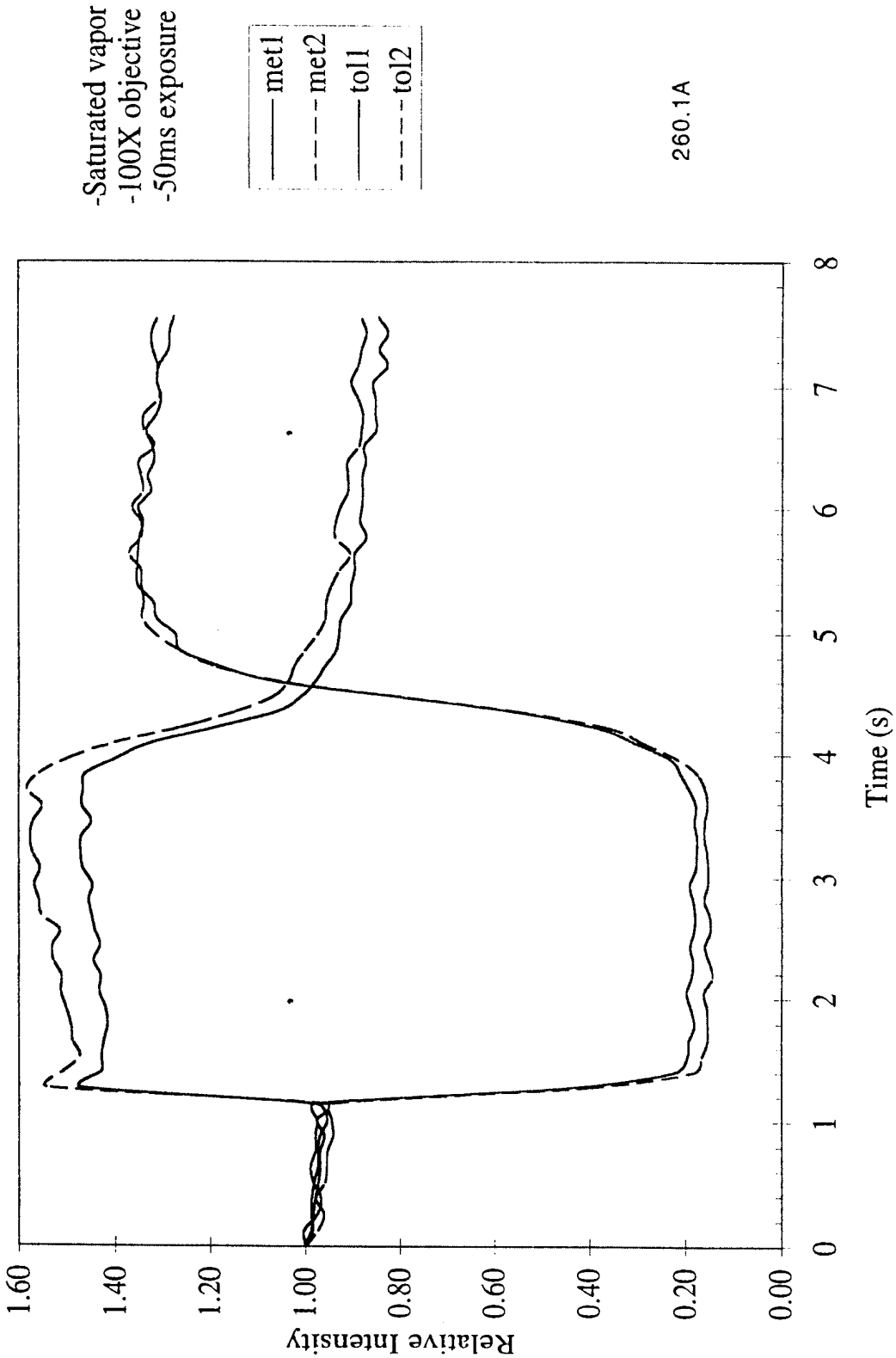

Signal Enhancement Through Multi-bead Response Summing
Nile Red-soaked poly(87% methyl styrene, 13% DVB) beads, 3.2µm
bright = wells with beads
dark = empty wells
Relative Intensity
250.00
200.00
150.00
100.00
50.00
0.00
-50.00
0
1
2
3
4
5
6
7
8
9
10
Time (s)

Fig. 16


"Thick-layer" PS802/Silica 3.2µm Beads
Region 2
-Saturated vapor
-100X objective
-50ms exposure
met1
met2
tol1
tol2
260.1A
Relative Intensity
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
1
2
3
4
5
6
7
8
Time (s)

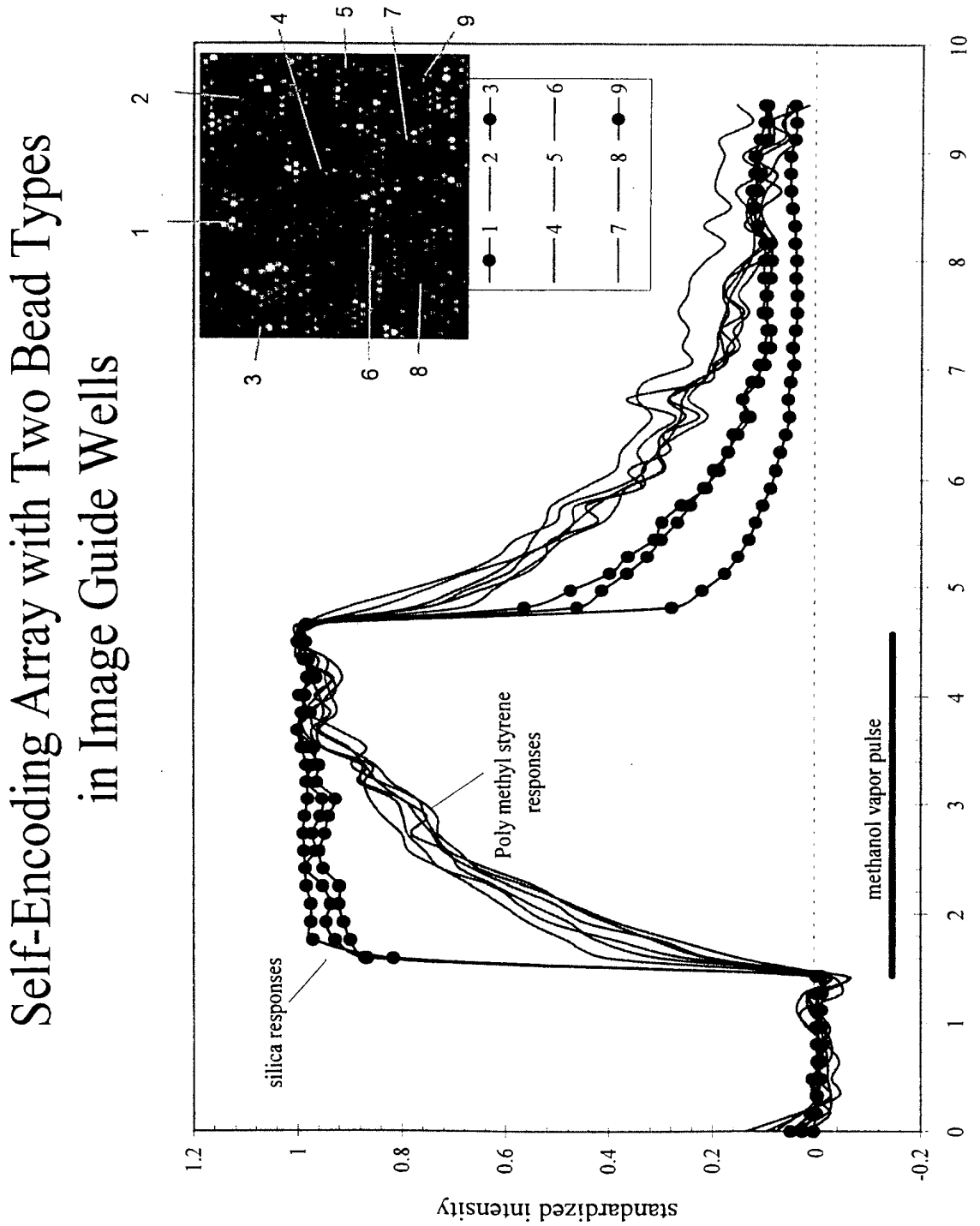

Self-Encoding Array with Two Bead Types
in Image Guide Wells
1
2
3
4
5
6
7
8
9
1 2 3
4 5 6
7 8 9
silica responses
Poly methyl styrene
responses
methanol vapor pulse
standardized intensity
time (s)
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0 1 2 3 4 5 6 7 8 9 10

Fig. 17

PCT/US98/21193

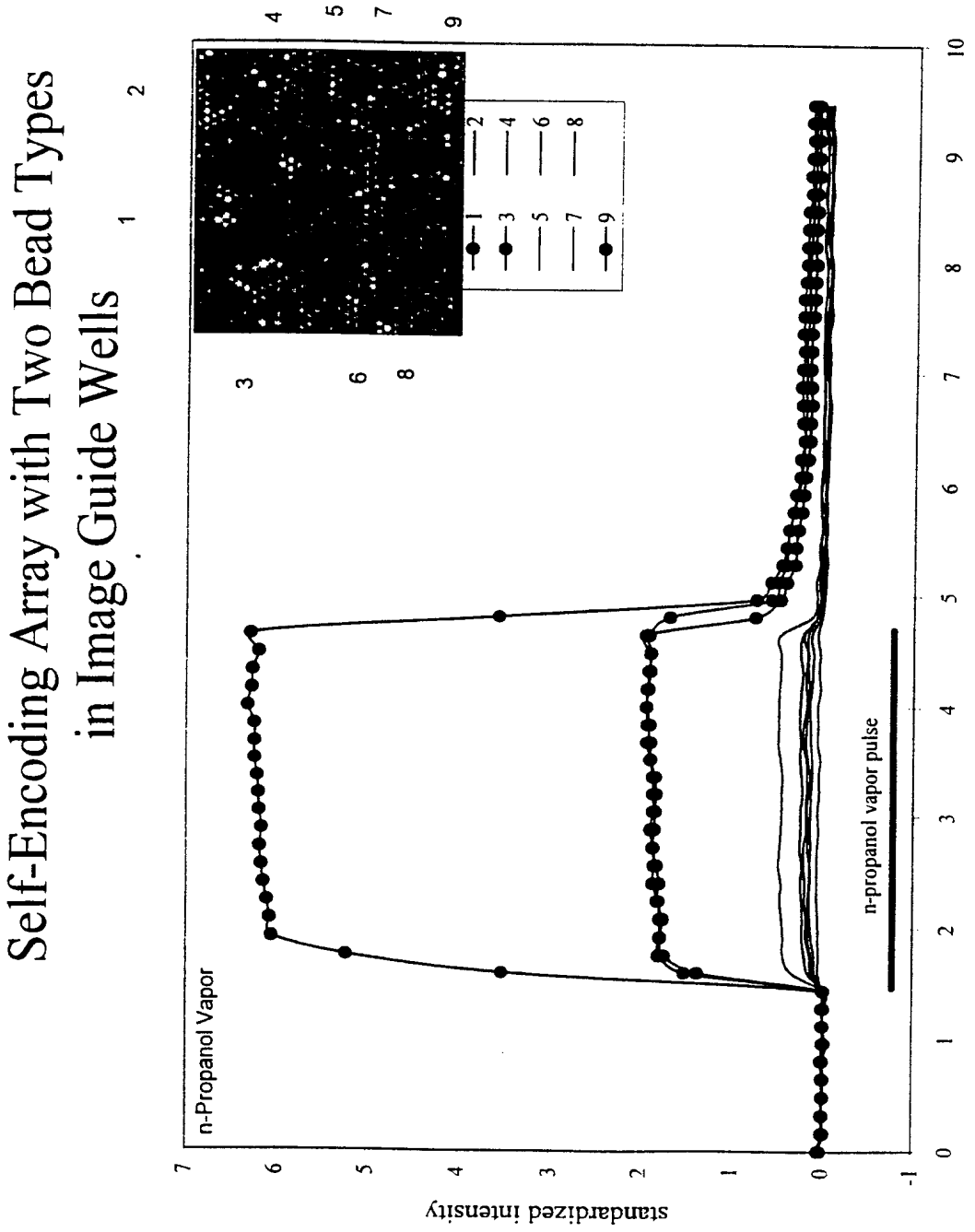


Fig. 18

Chart3

Page 1

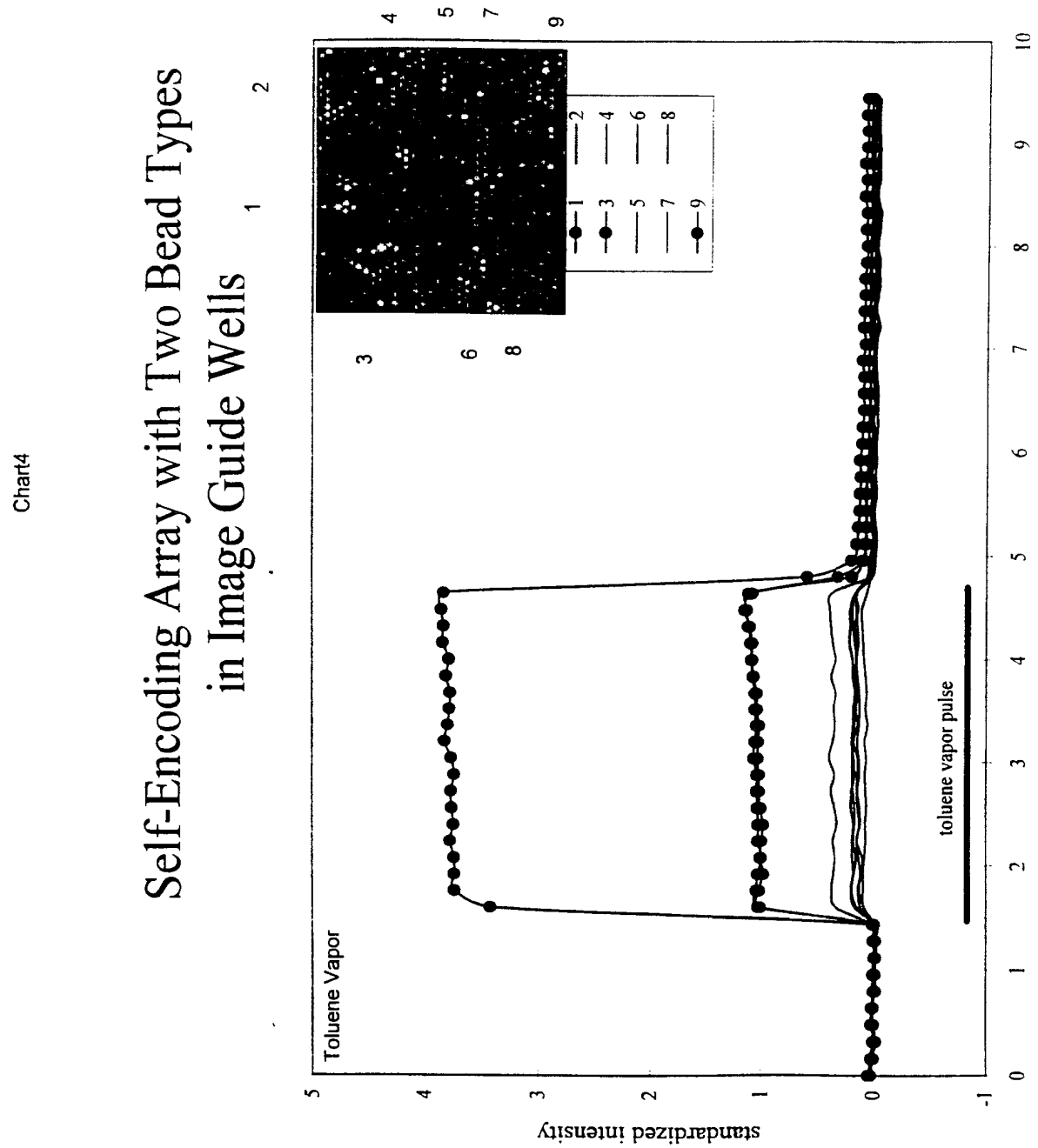
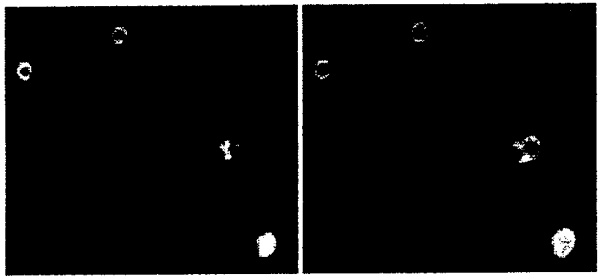

Chart4
Self-Encoding Array with Two Bead Types in Image Guide Wells
Toluene Vapor
standardized intensity
time (s)
toluene vapor pulse
1
2
3
4
5
6
7
8
9

Fig. 19

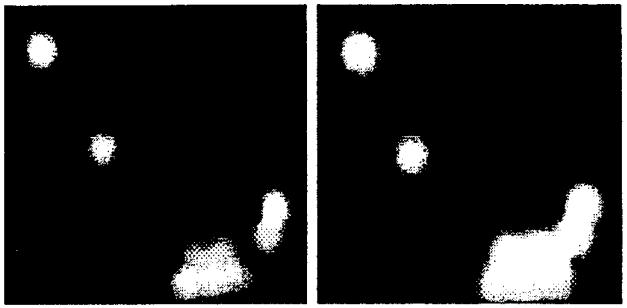
Swelling of three different bead types in presence of saturated toluene vapor

PS802 648.c



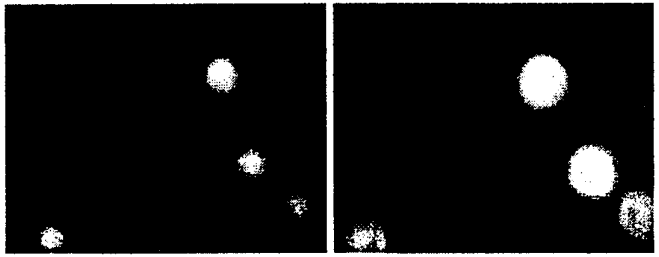
air toluene

Poly methyl styrene/
2% divinyl benzene



air toluene

Poly methyl styrene



air toluene

Fig. 20

INTERNATIONAL SEARCH REPORT

nternational Application No
PCT/US 98/21193

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G01N33/50 G01N21/62 G01N21/77 G01N21/64

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 94 12863 A (TUFTS COLLEGE) 9 June 1994 see the whole document ---	1-42
A	US 4 499 052 A (FULWYLER MACK J) 12 February 1985 see the whole document ---	1,11,14, 16,33, 36,39,41
A	WO 97 12030 A (NANOGEN, INC.) 3 April 1997 see page 17, line 33 - page 53, line 26; figures ---	1,11,14, 16,33, 36,39,41
A	GB 2 294 319 A (CAMBRIDGE IMAGING LTD) 24 April 1996 see the whole document ---	1,11,14, 16,33, 36,39,41
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 March 1999	31/03/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bosma, R

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/21193

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 723 146 A (STANFORD RES INST INT) 24 July 1996 see page 7, line 8 - page 31, line 44; figures	1,11,14, 16,33, 36,39,41
A	US 5 244 636 A (WALT ET AL.) 14 September 1993 see column 8, line 12 - column 31, line 30; figures	1,11,14, 16,33, 36,39,41
A	US 5 028 545 A (SOINI ERKKI J) 2 July 1991 see the whole document	1,11,14, 16,33, 36,39,41
A	US 4 822 746 A (WALT DAVID R) 18 April 1989 see column 5, line 32 - column 188, line 61; figures	1,11,14, 16,33, 36,39,41
A	US 5 308 771 A (ZHOU QUAN ET AL) 3 May 1994 see column 5, line 35 - column 9, line 29; figures	1,11,14, 16,33, 36,39,41

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 98/21193

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9412863 A	09-06-1994	US 5320814 A	14-06-1994
		CA 2128413 A	09-06-1994
		EP 0627074 A	07-12-1994
		JP 7509784 T	26-10-1995
US 4499052 A	12-02-1985	DE 3331017 A	08-03-1984
		FR 2532431 A	02-03-1984
		GB 2126341 A,B	21-03-1984
		JP 1772176 C	14-07-1993
		JP 4047265 B	03-08-1992
		JP 59060261 A	06-04-1984
		US 4717655 A	05-01-1988
WO 9712030 A	03-04-1997	US 5849486 A	15-12-1998
		AU 6968996 A	17-04-1997
		CA 2233238 A	03-04-1997
		CN 1202929 A	23-12-1998
		EP 0852617 A	15-07-1998
GB 2294319 A	24-04-1996	EP 0787288 A	06-08-1997
		WO 9612942 A	02-05-1996
EP 0723146 A	24-07-1996	AT 170004 T	15-09-1998
		CA 2144527 A	31-03-1994
		DE 69320484 D	24-09-1998
		DE 69320484 T	24-12-1998
		EP 0660936 A	05-07-1995
		ES 2123063 T	01-01-1999
		JP 8501632 T	20-02-1996
		WO 9407142 A	31-03-1994
		US 5674698 A	07-10-1997
		US 5736410 A	07-04-1998
US 5244636 A	14-09-1993	US 5244813 A	14-09-1993
		US 5320814 A	14-06-1994
		US 5250264 A	05-10-1993
US 5028545 A	02-07-1991	SE 458968 B	22-05-1989
		DE 3870324 A	27-05-1992
		EP 0296136 A	21-12-1988
		JP 1035268 A	06-02-1989
		JP 2638085 B	06-08-1997
		SE 8702511 A	17-12-1988
US 4822746 A	18-04-1989	US 5143853 A	01-09-1992
		US 5114864 A	19-05-1992
		US 5254477 A	19-10-1993
		US 5252494 A	12-10-1993
US 5308771 A	03-05-1994	NONE	

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number
WO 00/73777 A1

(51) International Patent Classification[7]: G01N 27/00, 27/02, 21/29, 15/06, 27/26, 33/543, 21/03, C12M 1/26, B29D 22/00, H01M 4/36, 6/30, 6/36, 2/38, 6/04, 4/02, B01L 3/00, B01J 19/08

(21) International Application Number: PCT/US00/14957

(22) International Filing Date: 26 May 2000 (26.05.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/320,274 28 May 1999 (28.05.1999) US

(71) Applicant *(for all designated States except US)*: **BIOARRAY SOLUTIONS, LLC** [—/US]; 120 Centennial Avenue, Piscataway, NJ 08854 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SEUL, Michael** [DE/DE]; 84 Pleasant Avenue, Fanwood, NJ 07023 (US). **LI, Alice, Xiang** [CN/CN]; 510 Sheffield Court, Piscataway, NJ 08854 (US).

(74) Agents: **DAVIDSON, Clifford, M.** et al.; Davidson, Davidson & Kappel, LLC, 1140 Avenue of the Americas, 15th floor, New York, NY 10036 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ARRAY CYTOMETRY

(57) Abstract: A method and apparatus for the manipulation of colloidal particulates and biomolecules at the interface between an insulating electrode such as silicon oxide and an electrolyte solution. Light-controlled elektrokinetic assembly of particles near surfaces relies on the combination of three functional elements: the AC electric field-induced assembly of planar aggregates; the patterning of the electrolyte/silicon oxide/silicon interface to exert spatial control over the assembly process; and the real-time control of the assembly process via external illumination. The present invention provides a set of fundamental operations enabling interactive control over the creation and placement of planar arrays of several types of particles and biomolecules and the manipulation of array shape and size. The present invention enables sample preparation and handling for diagnostic assays and biochemical analysis in an array format, and the functional integration of these operations. In addition, the present invention provides a procedure for the creation of material surfaces with desired properties and for the fabrication of surface-mounted optical components. This invention is also for a method and apparatus to direct the lateral motion and induce the assembly into planar arrays of cells on semiconductor surfaces in response to temporally and spatially varying electric fields and to projected patterns of illumination.

WO 00/73777 A1

ARRAY CYTOMETRY

This application is a continuation-in-part application of Application Serial No. 09/171,550, filed on October 26, 1998, corresponding to PCT International Application No. PCT/US97/08159, filed on April 24, 1997, which in turn is based on U.S. Provisional Application Serial No. 60/016,642, filed on April 25, 1996. Applicants hereby claim the priority of these prior applications pursuant to 35 U.S.C. §§ 119 and 120. These prior applications are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention generally relates to the field of materials science and analytical chemistry.

The present invention specifically relates to the realization of a complete, functionally integrated system for the implementation of biochemical analysis in a planar, miniaturized format on the surface of a conductive and/or photoconductive substrate, with applications in pharmaceutical and agricultural drug discovery and in in-vitro or genomic diagnostics. In addition, the method and apparatus of the present invention may be used to create material surfaces exhibiting desirable topographical relief and chemical functionality, and to fabricate surface-mounted optical elements such as lens arrays.

BACKGROUND OF THE INVENTION

I - Ions, Electric Fields and Fluid Flow: Field-induced Formation of Planar Bead Arrays

Electrokinesis refers to a class of phenomena elicited by the action of an electric field on the mobile ions surrounding charged objects in an electrolyte solution. When an object of given surface charge is immersed in a solution containing ions, a

diffuse ion cloud forms to screen the object's surface charge. This arrangement of a layer of (immobile) charges associated with an immersed object and the screening cloud of (mobile) counter-ions in solution is referred to as a "double layer". In this region of small but finite thickness, the fluid is not electroneutral. Consequently, electric fields acting on this region will set in motion ions in the diffuse layer, and these will in turn entrain the surrounding fluid. The resulting flow fields reflect the spatial distribution of ionic current in the fluid. Electroosmosis represents the simplest example of an electrokinetic phenomenon. It arises when an electric field is applied parallel to the surface of a sample container or electrode exhibiting fixed surface charges, as in the case of a silicon oxide electrode (in the range of neutral pH). As counter-ions in the electrode double layer are accelerated by the electric field, they drag along solvent molecules and set up bulk fluid flow. This effect can be very substantial in narrow capillaries and may be used to advantage to devise fluid pumping systems.

Electrophoresis is a related phenomenon which refers to the field-induced transport of charged particles immersed in an electrolyte. As with electroosmosis, an electric field accelerates mobile ions in the double layer of the particle. If, in contrast to the earlier case, the particle itself is mobile, it will compensate for this field-induced motion of ions (and the resulting ionic current) by moving in the opposite direction. Electrophoresis plays an important role in industrial coating processes and, along with electroosmosis, it is of particular interest in connection with the development of capillary electrophoresis into a mainstay of modern bioanalytical separation technology.

In confined geometries, such as that of a shallow experimental chamber in the form of a "sandwich" of two planar electrodes, the surface charge distribution and topography of the bounding electrode surfaces play a particularly important role in determining the nature and spatial structure of electroosmotic flow. Such a "sandwich" electrochemical cell may be formed by a pair of electrodes separated by a shallow gap. Typically, the bottom electrode will be formed by an oxide-capped silicon wafer, while the other electrode is formed by optically transparent, conducting indium tin oxide (ITO). The silicon (Si) wafer represents a thin slice of a single crystal of silicon which is doped to attain suitable levels of electrical conductivity and insulated from the electrolyte solution by a thin layer of silicon oxide (SiOx).

The reversible aggregation of beads into planar aggregates adjacent to an

electrode surface may be induced by a (DC or AC) electric field that is applied normal to the electrode surface. While the phenomenon has been previously observed in a cell formed by a pair of conductive ITO electrodes (Richetti, Prost and Barois, J. Physique Lettr. 45, L-1137 through L-1143 (1984)), the contents of which are incorporated herein by reference, it has been only recently demonstrated that the underlying attractive interaction between beads is mediated by electrokinetic flow (Yeh, Seul and Shraiman, "Assembly of Ordered Colloidal Aggregates by Electric Field Induced Fluid Flow", Nature 386, 57-59 (1997), the contents of which are incorporated herein by reference). This flow reflects the action of lateral non-uniformities in the spatial distribution of the current in the vicinity of the electrode. In the simplest case, such non-uniformities are introduced by the very presence of a colloidal bead near the electrode as a result of the fact that each bead interferes with the motion of ions in the electrolyte. Thus, it has been observed that an individual bead, when placed near the electrode surface, generates a toroidal flow of fluid centered on the bead. Spatial non-uniformities in the properties of the electrode can also be introduced deliberately by several methods to produce lateral fluid flow toward regions of low impedance. These methods are described in subsequent sections below.

Particles embedded in the electrokinetic flow are advected regardless of their specific chemical or biological nature, while simultaneously altering the flow field. As a result, the electric field-induced assembly of planar aggregates and arrays applies to diverse colloidal particles including: beaded polymer resins ("beads"), lipid vesicles, whole chromosomes, cells and biomolecules including proteins and DNA, as well as metal or semiconductor colloids and clusters.

Important for the applications to be described is the fact that the flow-mediated attractive interaction between beads extends to distances far exceeding the characteristic bead dimension. Planar aggregates are formed in response to an externally applied electric field and disassemble when the field is removed. The strength of the applied field determines the strength of the attractive interaction that underlies the array assembly process and thereby selects the specific arrangement adopted by the beads within the array. That is, as a function of increasing applied voltage, beads first form planar aggregates in which particles are mobile and loosely packed, then assume a tighter packing, and finally exhibit a spatial arrangement in the form of a crystalline, or

ordered, array resembling a raft of bubbles. The sequence of transitions between states of increasing internal order is reversible, including complete disassembly of planar aggregates when the applied voltage is removed. In another arrangement, at low initial concentration, beads form small clusters which in turn assume positions within an ordered "superstructure".

II - Patterning of Silicon Oxide Electrode Surfaces

Electrode patterning in accordance with a predetermined design facilitates the quasi-permanent modification of the electrical impedance of the EIS (Electrolyte-Insulator-Semiconductor) structure of interest here. By spatially modulating the EIS impedance, electrode-patterning determines the ionic current in the vicinity of the electrode. Depending on the frequency of the applied electric field, beads either seek out, or avoid, regions of high ionic current. Spatial patterning therefore conveys explicit external control over the placement and shape of bead arrays.

While patterning may be achieved in many ways, two procedures offer particular advantages. First, UV-mediated re-growth of a thin oxide layer on a properly prepared silicon surface is a convenient methodology that avoids photolithographic resist patterning and etching. In the presence of oxygen, UV illumination mediates the conversion of exposed silicon into oxide. Specifically, the thickness of the oxide layer depends on the exposure time and may thus be spatially modulated by placing patterned masks into the UV illumination path. This modulation in thickness, with typical variations of approximately 10 Angstroms, translates into spatial modulations in the impedance of the Si/SiOx interface while leaving a flat and chemically homogeneous top surface exposed to the electrolyte solution. Second, spatial modulations in the distribution of the electrode surface charge may be produced by UV-mediated photochemical oxidation of a suitable chemical species that is first deposited as a monolayer film on the SiOx surface. This method permits fine control over local features of the electrode double layer and thus over the electrokinetic flow.

A variation of this photochemical modulation is the creation of lateral gradients in the EIS impedance and hence in the current generated in response to the applied electric field. For example, this is readily accomplished by controlling the UV exposure so as to introduce a slow lateral variation in the oxide thickness or in the

surface charge density. As discussed below, control over lateral gradients serves to induce lateral bead transport and facilitates the implementation of such fundamental operations as capturing and channeling of beads to a predetermined destination along conduits in the form of impedance features embedded in the Si/SiOx interface. Photochemical patterning of functionalized chemical overlayers also applies to other types of electrode surfaces including ITO.

III - Light-controlled Modulation of the Interfacial Impedance

The spatial and temporal modulation of the EIS-impedance in accordance with a pattern of external illumination provides the basis to control the electrokinetic forces that mediate bead aggregation. The light-modulated electrokinetic assembly of planar colloidal arrays facilitates remote interactive control over the formation, placement and rearrangement of bead arrays in response to corresponding illumination patterns and thereby offers a wide range of interactive manipulations of colloidal beads and biomolecules.

To understand the principle of this methodology, it will be helpful to briefly review pertinent photoelectric properties of semiconductors, or more specifically, those of the EIS structure formed by the Electrolyte solution (E), the Insulating SiOx layer (I) and the Semiconductor (S). The photoelectric characteristics of this structure are closely related to those of a standard Metal-Insulator-Semiconductor (MIS) or Metal-Oxide-Semiconductor (MOS) devices which are described in S.M. Sze, "The Physics of Semiconductors", 2nd Edition, Chapt. 7 (Wiley Interscience 1981), the contents of which are incorporated herein by reference.

The interface between the semiconductor and the insulating oxide layer deserves special attention. Crucial to the understanding of the electrical response of the MOS structure to light is the concept of a space charge region of small but finite thickness that forms at the Si/SiOx interface in the presence of a bias potential. In the case of the EIS structure, an effective bias, in the form of a junction potential, is present under all but very special conditions. The space charge region forms in response to the distortion of the semiconductor's valence and conduction bands ("band bending") in the vicinity of the interface. This condition in turn reflects the fact that, while there is a bias potential across the interface, there is ideally no charge transfer in the presence of the

insulating oxide. That is, in electrochemical language, the EIS structure eliminates Faradaic effects. Instead, charges of opposite sign accumulate on either side of the insulating oxide layer and generate a finite polarization.

In the presence of a reverse bias, the valence and conduction band edges of an n-doped semiconductor bend upward near the Si/SiOx interface and electrons flow out of the interfacial region in response to the corresponding potential gradient. As a result, a majority carrier depletion layer is formed in the vicinity of the Si/SiOx interface. Light absorption in the semiconductor provides a mechanism to create electron-hole pairs within this region. Provided that they do not instantaneously recombine, electron-hole pairs are split by the locally acting electric field, and a corresponding photocurrent flows. It is this latter effect that affords control over the electrokinetic assembly of beads in the electrolyte solution.

To understand in more detail the pertinent frequency dependence of the light-induced modulation of the EIS impedance, two aspects of the equivalent circuit representing the EIS structure are noteworthy. First, there are close analogies between the detailed electrical characteristics of the electric double layer at the electrolyte-oxide interface, and the depletion layer at the interface between the semiconductor and the insulator. As with the double layer, the depletion layer exhibits electrical characteristics similar to those of a capacitor with a voltage-dependent capacitance. As discussed, illumination serves to lower the impedance of the depletion layer. Second, given its capacitive electrical response, the oxide layer will pass current only above a characteristic ("threshold") frequency. Consequently, provided that the frequency of the applied voltage exceeds the threshold, illumination can lower the effective impedance of the entire EIS structure.

This effective reduction of the EIS impedance also depends on the light intensity which determines the rate of generation of electron-hole pairs. In the absence of significant recombination, the majority of photogenerated electrons flow out of the depletion region and contribute to the photocurrent. The remaining hole charge accumulates near the Si/SiOx interface and screens the electric field acting in the depletion region. As a result, the rate of recombination increases, and the efficiency of electron-hole separation, and hence the photocurrent, decreases. For given values of frequency and amplitude of the applied voltage, one therefore expects that as the

illumination intensity increases, the current initially increases to a maximum level and then decreases. Similarly, the impedance initially decreases to a minimum value (at maximum current) and then decreases.

This intensity dependence may be used to advantage to induce the lateral displacement of beads between fully exposed and partially masked regions of the interface. As the illumination intensity is increased, the fully exposed regions will correspond to the regions of interface of lowest impedance, and hence of highest current, and beads will be drawn into these regions. As the fully exposed regions reach the state of decreasing photocurrent, the effective EIS impedance in those regions may exceed that of partially masked regions, with a resulting inversion of the lateral gradient in current. Beads will then be drawn out of the fully exposed regions. Additionally, time-varying changes in the illumination pattern may be used to effect bead motion.

IV - Integration of Biochemical Analysis in a Miniaturized, Planar Format

The implementation of assays in a planar array format, particularly in the context of biomolecular screening and medical diagnostics, has the advantage of a high degree of parallelity and automation so as to realize high throughput in complex, multi-step analytical protocols. Miniaturization will result in a decrease in pertinent mixing times reflecting the small spatial scale, as well as in a reduction of requisite sample and reagent volumes as well as power requirements. The integration of biochemical analytical techniques into a miniaturized system on the surface of a planar substrate ("chip") would yield substantial improvements in the performance, and reduction in cost, of analytical and diagnostic procedures.

Within the context of DNA manipulation and analysis, initial steps have been taken in this direction (i.e., miniaturization) by combining on a glass substrate, the restriction enzyme treatment of DNA and the subsequent separation of enzyme digests by capillary electrophoresis, see, for example, Ramsey, PCT Publication No. WO 96/04547, the contents of which are incorporated herein by reference, or the amplification of DNA sequences by application of the polymerase chain reaction (PCR) with subsequent electrophoretic separation, see, for example, U.S. Patent Nos. 5,498,392 and 5,587,128 to Wilding et al., the contents of which are incorporated herein by reference.

While these standard laboratory processes have been demonstrated in a miniaturized format, they have not been used to form a complete system. A complete system will require additional manipulation such as front-end sample processing, binding and functional assays and the detection of small signals followed by information processing. The true challenge is that of complete functional integration because it is here that system architecture and design constraints on individual components will manifest themselves. For example, a fluidic process is required to concatenate analytical steps that require the spatial separation, and subsequent transport to new locations, of sets of analyte. Several possibilities have been considered including electroosmotic pumping and transport of droplets by temperature-induced gradients in local surface tension. While feasible in demonstration experiments, these techniques place rather severe requirements on the overall systems lay-out to handle the very considerable DC voltages required for efficient electroosmotic mixing or to restrict substrate heating when generating thermally generated surface tension gradients so as to avoid adverse effects on protein and other samples.

Summary of the Invention

The present invention combines three separate functional elements to provide a method and apparatus facilitating the real-time, interactive spatial manipulation of colloidal particles ("beads") and molecules at an interface between a light sensitive electrode and an electrolyte solution. The three functional elements are: the electric field-induced assembly of planar particle arrays at an interface between an insulating or a conductive electrode and an electrolyte solution; the spatial modulation of the interfacial impedance by means of UV-mediated oxide regrowth or surface- chemical patterning; and, finally, the real-time, interactive control over the state of the interfacial impedance by light. The capabilities of the present invention originate in the fact that the spatial distribution of ionic currents, and thus the fluid flow mediating the array assembly, may be adjusted by external intervention. Of particular interest is the introduction of spatial non-uniformities in the properties of the pertinent EIS structure. As described herein, such inhomogeneities, either permanent or temporary in nature, may be produced by taking advantage of the physical and chemical properties of the EIS structure.

The invention relates to the realization of a complete, functionally integrated system for the implementation of biochemical analysis in a planar, miniaturized format on the surface of a silicon wafer or similar substrate. In addition, the method and apparatus of the present invention may be used to create material surfaces exhibiting desirable topographical relief and chemical functionality, and to fabricate surface-mounted optical elements such as lens arrays.

The combination of three functional elements endows the present invention with a set of operational capabilities to manipulate beads and bead arrays in a planar geometry to allow the implementation of biochemical analytical techniques. These fundamental operations apply to aggregates and arrays of colloidal particles including: beaded polymer resins also referred to as latices, vesicles, whole chromosomes, cells and biomolecules including proteins and DNA, as well as metal or semiconductor colloids and clusters.

Sets of colloidal particles may be captured, and arrays may be formed in designated areas on the electrode surface (Figs. 1a, 1b and Figs. 2a-d). Particles, and the arrays they form in response to the applied field, may be channeled along conduits of any configuration that are either embedded in the Si/SiOx interface by UV-oxide patterning or delineated by an external pattern of illumination. This channeling (Figs. 1c, 1d, 1e, Figs. 3c, 3d), in a direction normal to that of the applied electric field, relies on lateral gradients in the impedance of the EIS structure and hence in the field-induced current. As discussed herein, such gradients may be introduced by appropriate patterns of illumination, and this provides the means to implement a gated version of translocation (Fig. 1e). The electrokinetic flow mediating the array assembly process may also be exploited for the alignment of elongated particles, such as DNA, near the surface of the electrode. In addition, the present invention permits the realization of methods to sort and separate particles.

Arrays of colloidal particles may be placed in designated areas and confined there until released or disassembled. The overall shape of the array may be delineated by UV-oxide patterning or, in real time, by shaping the pattern of illumination. This capability enables the definition of functionally distinct compartments, permanent or temporary, on the electrode surface. Arrays may be subjected to changes of shape imposed in real time, and they may be merged with other arrays (Fig.

1f) or split into two or more subarrays or clusters (Fig. 1g, Figs. 4a, 4b). In addition, the local state of order of the array as well as the lateral particle density may be reversibly adjusted by way of the external electric field or modified by addition of a second, chemically inert bead component.

The present invention also allows for the combination of fundamental operations to develop increasingly complex products and processes. Examples given herein describe the implementation of analytical procedures essential to a wide range of problems in materials science, pharmaceutical drug discovery, genomic mapping and sequencing technology. Important to the integration of these and other functionalities in a planar geometry is the capability, provided by the present invention, to impose temporary or permanent compartmentalization in order to spatially isolate concurrent processes or sequential steps in a protocol and the ability to manipulate sets of particles in a manner permitting the concatenation of analytical procedures that are performed in different designated areas on the substrate surfaces.

This invention is for a method and apparatus to direct the lateral motion and induce the assembly into planar arrays of cells on semiconductor surfaces in response to temporally and spatially varying electric fields and to projected patterns of illumination.

The response of different types of cells to variations in the frequency and voltage of the applied electric field and to specific patterns and changing intensity of illumination enables the on-cue assembly of multiple arrays of cells in any desired position on the substrate, and further enables the repositioning, disassembly and re-assembly and more generally the reconfiguration and segmentation of such arrays. Time-constant and time-varying illumination patterns are generated by an optically programmable illumination pattern generator may be used and can be based on a graphical representation of the pattern created on a personal computer. The novel methodology of image cytometry introduced in the present invention relates to the general fields of molecular and cellular biology and has applications in the fields of diagnostics, genetic analysis, cell biology and drug discovery. The image cytometry facilitates a multiplicity of cell-based biological and biochemical procedures based on the quantitative analysis of images recorded from array of cells. As with Random Encoded Array Detection (READ) or Programmable Array Reconfiguration and Segmentation

(PARSE), these assays are performed in a highly parallel format in a small volume of sample on the surface of a semiconductor wafer.

Brief Description of Drawings

Other objects, features and advantages of the invention discussed in the above brief explanation will be more clearly understood when taken together with the following detailed description of an embodiment which will be understood as being illustrative only, and the accompanying drawings reflecting aspects of that embodiment, in which:

Figs. 1a-h are illustrations of the fundamental operations for bead manipulation;

Figs. 2a and 2b are photographs illustrating the process of capturing particles in designated areas on the substrate surface;

Figs. 2c and 2d are photographs illustrating the process of excluding particles from designated areas on the substrate surface;

Figs. 3a and 3b are illustrations of the oxide profile of an Si/SiOx electrode;

Figs. 3c and 3d are photographs of the channeling of particles along conduits;

Figs. 4a and 4b are photographs of the splitting of an existing aggregate into small clusters;

Fig. 5 is a photograph of the lensing action of individual colloidal beads;

Figs. 6a-c are side view illustrations of a layout-preserving transfer process from a microtiter plate to a planar cell;

Fig. 7 is a photograph of the inclusion of spacer particles within bead clusters;

Fig. 8 is an illustration of binding assay variations;

Figs. 9a and 9b are illustrations of two mechanisms of particle sorting, and Fig. 9c is an illustration of the oxide profile of Fig. 9b;

Fig. 10 is an illustration of a planar array of bead-anchored probe-target complexes;

Figs. 11a-11e are illustrations of DNA stretching in accordance with the

present invention;

Fig. 12 is an illustration of *E.coli* cells (grown to log phase and suspended in 280 mM mannitol solution at the density of 10^8/ml) which form arrays under electric field. (Bright field image. 25x obj, NA 0.45. Electric field: 5V, 200 Hz);

Fig. 13 is an illustration of the same effect as in Figure 12, with *S. cerevisiae* cells (baker's yeast). (Bright field image. 25x obj, NA 0.45; electric field: 5V, 400 Hz);

Fig. 14 is an illustration of *E.coli* forming diamond shapes that are projected to the surface by an LCD panel controlled by appropriate graphics software; (20x objective, NA 0.28; Electric field: 10V, 7 kHz);

Fig. 15, is an illustration of *S. cerevisiae* responding to two circular illumination patterns;(20x objective, NA 0.28; Electric field: 15 V, 15 kHz)

Fig. 16 is an illustration of *S. cerevisiae* responding to illumination, with an arrow indicate the direction cells are moving in or being expelled from the illuminated regions (20x objective, NA 0.28; Electric field: 15 V, 15 kHz);

Fig.17 is an illustration of *E. Coli* cells being repelled from the illumination circle(20x objective, NA 0.28; Electric field: 10 V, 7 kHz; low substrate doping level (1 Ohm cm-10 Ohm cm));

Fig. 18 is an illustration of *E.coli* cells following the expansion of the rectangular illuminated shape (20x objective, NA 0.28; Electric field: 15 V, 8 kHz);

Fig. 19 is an illustration of *E. coli* cells being trapped by the light stripe at the left and being prevented from moving to the circular region on the right (20x objective, NA 0.28; Electric field: 15 V, 8 kHz);

Fig. 20 is an illustration of an illuminated region containing both types of cells being suddenly expanded or "dragged", with *E. coli* cells (left) responding faster than the larger *S. cerevisiae* cells (right), resulting in segregation. *(Bright field image. 25x obj, NA 0.45 Electric field: 5V, 400 Hz)*;

Fig. 21 is an illustration of immunophenotyping according to the present invention;

Fig. 22 is an illustration of affinity cell detection according to the present invention;

Fig. 23 is an illustration of cell sorting and isolation according to the

present invention;

Fig. 24 is an illustration of cell-based screening according to the present invention;

Fig. 25 is an illustration of cell-secretion based screening according to the present invention;

Fig. 26 is an illustration of cell to cell interaction according to the present invention; and

Fig. 27 is an illustration of chemotaxis analysis according to the present invention.

Detailed Description of the Preferred Embodiments

The three functional elements of the present invention may be combined so as to provide a set of fundamental operations for the interactive spatial manipulation of colloidal particles and molecules, assembled into planar aggregates adjacent to an electrode surface. In the following description, fundamental operations in this "toolset" are described in order of increasing complexity. Specifically, it is useful to adopt a classification scheme based on the total number of inputs and outputs, or "terminals", involved in a given operation. For example, the merging of two separate arrays, or sets of particles, into one would be a "three-terminal" operation, involving two inputs and one output. The converse three-terminal operation, involving one input and two outputs, is the splitting of a given array into two subarrays.

Experimental conditions yielding the phenomena depicted in the various photographs included herein are as follows. An electrochemical cell is formed by a pair of planar ITO electrodes, composed of an ITO layer deposited on a glass substrate, or by a Si/SiOx electrode on the bottom and an ITO electrode on the top, separated by a typical gap of 50 microns or less. Given its dependence on the photoelectric properties of the Si/SiOx interface, light control is predicated on the use of a Si/SiOx electrode. Leads, in the form of platinum wires, are attached to the ITO and to the silicon electrode, which is first etched to remove the insulating oxide in the contact region, by means of silver epoxy. The cell is first assembled and then filled, relying on capillary action, with a suspension of colloidal beads, 1 or 2 microns in diameter, at a typical

concentration of 0.1% solids in 0.1mM azide solution, corresponding to approximately 2x10^8 particles per milliliter. The number is chosen so as to yield between ½ and 1 full monolayer of particles on the electrode surface. Anionic (e.g., carboxylated polystyrene, silica), cationic (e.g., aminated polystyrene) or nominally neutral (e.g., polystyrene) have all been used to demonstrate the basic phenomena underlying the three functional elements of the present invention. The silicon electrode was fabricated from a 1 inch-square portion of a Si (100) wafer, typically 200-250 microns thick, n-doped to typically 0.01 Ohm cm resistivity, and capped with a thin oxide of typically 30-40 Angstroms thickness. A thick oxide layer of typically 6000-8000 Angstrom thickness, grown under standard conditions in a furnace at 950 degrees C, may be etched by standard photolithography to define the structures of interest. Alternatively, a thin oxide layer may be regrown on a previously stripped surface of (100)-orientation under UV illumination. Given its ease of implementation and execution, UV-mediated oxide regrowth is the preferable technique: it provides the means to pattern the surface by placing a quartz mask representing the desired pattern in the path of UV illumination and it leaves a chemically homogeneous, topographically flat top surface. To avoid non-specific particle adsorption to the electrode surface, stringent conditions of cleanliness should be followed, such as those set forth in the General Experimental Conditions below.

The fundamental one-terminal operation is a "capture-and-hold" operation (Fig. 1a) which forms an array of particles in a designated area of arbitrary outline on the surface that is delineated by UV-mediated oxide patterning or by a corresponding pattern of illumination projected on an otherwise uniform Si/SiOx substrate surface. Figs. 2a and 2b illustrate bead capture on a surface characterized by a very thin oxide region 22 (approximately 20-30 Angstroms in thickness) and correspondingly low impedance, while the remaining surface is covered with the original, thick oxide with correspondingly high impedance. In Fig. 2a, there is no applied field, and hence, no bead capture. In contrast, in Fig. 2b, an electric field is applied (10Vp-p source, 1 kHz) and bead capture occurs within the thin oxide region 22. Under these conditions, an array starts to grow within less than a second and continues to grow over the next approximately 10 seconds as beads arrive from increasingly larger distances to add to the outward growing perimeter of region 22. Growth stops when the array approaches

the outer limit of the delineated target area, i.e., the area defined by the thin oxide having a low impedance. The internal state of order of the captured aggregate of beads is determined by the strength of the
applied voltage, higher values favoring increasingly denser packing of beads and the eventual formation of ordered arrays displaying a hexagonally crystalline configuration in the form of a bubble raft. The array remains in place as long as the applied voltage is present. Removal of the applied voltage results in the disassembly of the array.

The "capture-and-hold" operation may also be implemented under illumination with visible or infrared light, for example by simply projecting a mask patterned with the desired layout onto the Si/SiOx electrode. A regular 100W quartz microscope illuminator has been used for this purpose on a Zeiss UEM microscope, with apertures or masks inserted in the intermediate image plane to provide the required shape in the plane of the electrode (when focused properly under conditions of Koehler illumination). Alternatively, an IR laser diode with output of 3 mW at 650 - 680nm also has been used. The use of external illumination rather than oxide patterning for the spatial confinement of particles allows the confinement pattern to be easily modified.

Related to "capture-and-hold" is the one-terminal operation of "exclude-and-hold" (Fig. 1b) which clears particles from a designated area on the surface. Increasing the frequency of the applied voltage to approximately 100kHz leads to an inversion in the preference of particles which assemble in the thin-oxide portion of the surface (e.g., region 22, Fig. 2b) and instead form structures decorating the outside of the target area perimeter. Rather than relying on this effect, the exclusion of particles from the desired areas is also accomplished, in analogy to the original "capture-and-hold" operations, by simply embedding the corresponding structure in the Si/SiOx interface by UV- mediated oxide regrowth. In the example of Figs. 2c and 2d, this is achieved, under conditions otherwise identical to those described above, with respect to Figs. 2a and 2b, by applying 20V (pp) at 10kHz. While the oxide thickness in the non designated areas 24 is approximately 30 Angstroms, the value in the designated square areas 26 is approximately 40 Angstroms, implying a correspondingly higher impedance at the applied frequency.

The "capture-and-hold" operation enables the spatial compartmentalization of the substrate surface into functionally distinct regions. For

example, particles of distinct chemical type, introduced into the electrochemical cell at different times or injected in different locations, can be kept in spatially isolated locations by utilizing this operation.

The fundamental two-terminal operation is translocation (Fig. 1c), or the controlled transport of a set of particles from location O to location F on the surface; here, O and F are target areas to which the above-described one-terminal operations may be applied. The one-dimensional, lateral bead transport used in translocation is achieved by imposing a lateral current along a conduit connecting areas O and F, as shown in Figs. 3a and 3b or by projecting a corresponding linear pattern of illumination. In this channeling operation, beads move in the direction of lower impedance in the direction of the arrow shown in Figs. 3a and 3b, in accordance with the underlying electrokinetic flow.

Oxide patterning may be utilized in two ways to create a lateral current along the Si/SiOx interface. The simplest method is depicted in Fig. 3c and shows a large open holding area 32 fed by three narrow conduits 34 defined by etching a thermal oxide. Beads move to the holding area 32 along the narrow conduits 34 to form a bead array. Fig. 3d is a large scale view of the array of Fig. 3c. The principle invoked in creating transport is that of changing the aspect ratio (narrow conduit connected to wide holding area) of the embedded pattern with constant values of thin oxide thickness inside and thick oxide outside, as illustrated in Fig. 3a. In Figs. 3c and 3d, the applied voltage was 10V (pp) at 10kHz. An alternative approach for creating bead transport, enabled by UV-mediated oxide regrowth, is to vary the oxide thickness along the conduit in a controlled fashion. This is readily accomplished by UV exposure through a graduated filter. Differences in the oxide thickness between O and F of as little as 5-10 Angstroms suffice to effect lateral transport. In this situation, the aspect ratio of the conduit and holding areas need not be altered. This is illustrated in Fig. 3b.

The use of external illumination to define conduits, by varying the illumination intensity along the conduit to create the requisite impedance gradient, has the advantage that the conduit is only a temporary structure, and that the direction of motion may be modified or reversed if so desired. The present invention provides for mechanisms of light-mediated active linear transport of planar aggregates of beads under interactive control. This is achieved by adjusting an external pattern of illumination in

real time, either by moving the pattern across the substrate surface in such a way as to entrain the illuminated bead array or by electronically modulating the shape of the pattern to induce motion of particles.

Two modes of light-mediated, active transport are:

i) Direct Translocation ("tractor beam") which is a method of translocating arrays and of delineating their overall shape by adjusting parameters so as to favor particle assembly within illuminated areas of the surface, as described herein. Arrays simply follow the imposed pattern. The rate of motion is limited by the mobility of particles in the fluid and thus depends on particle diameter and fluid viscosity.

ii) Transverse Array Constriction is a bead transport mechanism related to peristaltic pumping of fluids through flexible tubing. The light-control component of the present invention may be used for a simple implementation of this very general concept. A multi-component planar aggregate of beads is confined to a rectangular channel, by UV-patterning if so desired, or simply by light. Beads are free to move along the channel by diffusion (in either direction). An illumination pattern matching the transverse channel dimension is set up and is then varied in time so as to produce a transverse constriction wave that travels in one direction along the channel. Such a constriction wave may be set up in several ways. A conceptually simple method is to project a constricting mask onto the sample and move the projected mask pattern in the desired fashion. This method also may be implemented electronically by controlling the illumination pattern of a suitable array of light sources, thus obviating the need for moving parts in the optical train.

The control of lateral bead transport by changing or moving patterns of illumination has the advantage that it may be applied whenever and wherever (on a given substrate surface) required, without the need to impose gradients in impedance by predefined UV patterning. On the other hand, a predefined impedance pattern can provide additional capabilities in conjunction with light-control. For example, it may be desirable to transport beads against a substrate-embedded impedance gradient to separate beads on the basis of mobility.

Conduits connecting O and F need not be straight: as with tracks directing the motion of trains, conduits may be shaped in any desirable fashion (Fig. 1d). A gated

version of translocation (Fig. 1e) permits the transport of particles from O to F only after the conduit is opened (or formed in real time) by a gating signal. This operation utilizes UV oxide patterning to establish two holding areas, O and F, and also light control to temporarily establish a conduit connecting O and F. An alternative implementation is based on an oxide embedded impedance gradient. A zone along the conduit is illuminated with sufficiently high intensity to keep out particles, thereby blocking the passage. Removal (or reduction in intensity) of the illumination opens the conduit. In the former case, light enables the transport of beads, while in the latter case, light prevents the transport of beads.

The fundamental three-terminal operations are the merging and splitting of sets or arrays of beads (Figs. 1f and 1g). The merging of two arrays (Fig. 1f) involves the previous two fundamental operations of "capture-and-hold", applied to two spatially isolated sets of beads in locations O1 and O2, and their respective channeling along merging conduits into a common target area, and their eventual channeling, subsequent to mixing, or a chemical reaction, into the final destination, a third holding area, F. This is accomplished, under the conditions stated above, by invoking one-terminal and gated two-terminal operations.

The splitting of an array into two subarrays (Fig. 1g) is a special case of a generally more complex sorting operation. Sorting involves the classification of beads in a given set or array into one of two subsets, for example according to their fluorescence intensity. In the simpler special case, a given array, held in area O, is to be split into two subarrays along a demarcation line, and subarrays are to be moved to target areas F1 and F2. Under the conditions stated above, this is accomplished by applying the "capture-and-hold" operation to the array in O. Conduits connect O to F1 and F2. High intensity illumination along a narrowly focused line serves to divide the array in a defined fashion, again relying on gated translocation to control transport along conduits away from the holding area O. An even simpler version, termed indiscriminate splitting, randomly assigns particles into F1 and F2 by gated translocation of the array in O into F1 and F2 after conduits are opened as described above.

Figs. 4a and 4b show a variant in which beads in region O (Fig. 4a) are split into multiple regions F1, F2, ... Fn (Fig. 4b). This reversible splitting of an aggregate or array into n subarrays, or clusters, is accomplished, for carboxylated

polystyrene spheres of 2 micron diameter at a concentration corresponding to an electrode coverage of a small fraction of a monolayer, at a frequency of 500Hz, by raising the applied voltage from typically 5V (pp) to 20V (pp). This fragmentation of an array into smaller clusters reflects the effect of a field-induced particle polarization. The splitting is useful to distribute particles in an array over a wider area of substrate for presentation to possible analytes in solution, and for subsequent scanning of the individual clusters with analytical instruments to make individual readings.

The three functional elements of the present invention described herein may be also combined to yield additional fundamental operations to control the orientation of anisotropic objects embedded in the electroosmotic flow created by the applied electric field at the electrode surface. The direction of the flow, in the plane of the substrate, is controlled by gradients in the impedance that are shaped in the manner described in connection with the channeling operation. This is used to controllably align anisotropic objects as illustrated in Fig. 1h, and may be applied to stretch out and align biomolecules, such as DNA.

An additional fundamental operation that complements the previous set is that of permanently anchoring an array to the substrate. This is best accomplished by invoking anchoring chemistries analogous to those relying on heterobifunctional cross-linking agents invoked to anchor proteins via amide bond formation. Molecular recognition, for example between biotinylated particles and surface-anchored streptavidin, provides another class of coupling chemistries for permanent anchoring.

General Experimental Conditions

The functional elements, namely the electric-field induced assembly of planar particle arrays, the spatial modulation of the interfacial impedance by means of UV-mediated oxide or surface-chemical patterning and finally, the control over the state of the interfacial impedance by light which are used in the present invention, have been demonstrated in experimental studies. These studies employed n-doped silicon wafers (resistivities in the range of 0.01 Ohm cm), capped with either thermally grown oxide layers of several thousand Angstrom thickness, or with thin oxide layers, regrown after removal of the original "native" oxide in HF, under UV illumination from a deuterium source in the presence of oxygen to typical thicknesses between 10 and 50 Angstroms.

Lithographic patterning of thermally grown oxide employed standard procedures implemented on a bench top (rather than a clean room) to produce features in the range of several microns.

Surfaces were carefully cleaned in adherence with industry standard RCA and Piranha cleaning protocols. Substrates were stored in water produced by a Millipore water purification system prior to use. Surfaces were characterized by measuring the contact angle exhibited by a 20 microliter droplet of water placed on the surface and viewed (from the side) through a telescope. The contact angle is defined as the angle subtended by the surface and the tangent to the droplet contour (in side view) at the point of contact with the surface. For example, a perfectly hemispherical droplet shape would correspond to a contact angle of 90 degrees. Surface chemical derivatization with mercapto-propyl-trimethoxysilane (2% in dry toluene) produced surfaces giving typical contact angles of 70 degrees. Oxidation of the terminal thiol functionality under UV irradiation in the presence of oxygen reduced the contact angle to zero in less than 10 min of exposure to UV from the deuterium source. Other silane reagents were used in a similar manner to produce hydrophobic surfaces, characterized by contact angles in excess of 110 degrees.

Simple "sandwich" electrochemical cells were constructed by employing kapton film as a spacer between Si/SiOx and conductive indium tin oxide (ITO), deposited on a thin glass substrate. Contacts to platinum leads were made with silver epoxy directly to the top of the ITO electrode and to the (oxide-stripped) backside of the Si electrode. In this two-electrode configuration, AC fields were produced by a function generator, with applied voltages ranging up to 20V and frequencies varying from DC to 1 MHZ, high frequencies favoring the formation of particle chains connecting the electrodes. Currents were monitored with a potentiostat and displayed on an oscilloscope. For convenience, epi-fluorescence as well as reflection differential interference contrast microscopy employed laser illumination. Light-induced modulations in EIS impedance were also produced with a simple 100W microscope illuminator as well as with a 3mW laser diode emitting light at 650-680 nm.

Colloidal beads, both anionic and cationic as well as nominally neutral, with a diameter in the range from several hundred Angstroms to 20 microns, stored in a NaN_2 solution, were employed.

Close attention was paid to colloidal stability to avoid non-specific interactions between particles and between particles and the electrode surface. Bacterial contamination of colloidal suspensions was scrupulously avoided.

Typical operating conditions producing, unless otherwise indicated, most of the results described herein, were: 0.2 mM NaN_2 (sodium azide) solutions, containing particles at a concentration so as to produce not more than a complete monolayer of particles when deposited on the electrode; applied DC potentials in the range of 1-4V, and AC potentials in the range of 1-10V (peak-to-peak) and 500Hz - 10kHz, with an electrode gap of 50 microns; anionic (carboxylated polystyrene) beads of 2 micron diameter, as well as (nominally neutral) polystyrene beads of 2-20 micron diameter.

The method and apparatus of the present invention may be used in several different areas, examples of which are discussed in detail. Each example includes background information followed by the application of the present invention to that particular application.

Example I - Fabrication of Surfaces and Coatings with Designed Properties

The present invention may be used to fabricate planar surfaces and coatings with designed properties. Specifically, the functional elements of the present invention enable the formation of arrays composed of particles of a wide range of sizes (approximately 100 Angstrom to 10 microns) and chemical composition or surface functionality in response to AC or DC electric fields. These arrays may be placed and delineated in designated areas of the substrate, and the interparticle spacing and internal state of order within the array may be controlled by adjusting the applied field prior to anchoring the array to the substrate. The newly formed surfaces display pre-designed mechanical, optical and chemical characteristics, and they may be subjected to further modification by subsequent treatment such as chemical cross-linking.

The mechanical and/or chemical modification of surfaces and coatings principally determines the interaction between materials in a wide range of applications that depend on low adhesion (e.g., the familiar "non-stick" surfaces important in housewares) or low friction (e.g., to reduce wear in computer hard disks), hydrophobicity (the tendency to repel water, e.g., of certain fabrics), catalytic activity or specific chemical functionality to either suppress molecular interactions with surfaces or

to promote them. The latter area is of particular importance to the development of reliable and durable biosensors and bioelectronic devices. Finally, a large number of applications depend on surfaces of defined topography and/or chemical functionality to act as templates controlling the growth morphology of deposited materials or as "command surfaces" directing the alignment of optically active molecules in deposited thin organic films, as in liquid crystal display applications.

Extensive research has been devoted to the formation of surfaces by adsorption of thin organic films of known composition from the liquid or gas phase by several methods. Notwithstanding their seeming simplicity and wide-spread use, these methods can be difficult to handle in producing reliable and reproducible results. In addition, molecular films are not well suited to produce surfaces displaying a regular topography.

An alternative approach to the problem is the modification of conductive surfaces by electrophoretic deposition of suspended particulates. This is a widely used technique in industrial settings to produce paint coatings of metal parts, and to deposit phosphor for display screens. The active deposition process significantly enhances the kinetics of formation (in contrast to passive adsorption of organic films from solution), an important consideration in practical applications. Electrophoretic deposition requires high DC electric fields and produces layers in which particles are permanently adsorbed to the surface. While particles in so-deposited monolayers are usually randomly distributed, the formation of polycrystalline monolayers of small (150 Angstrom) gold colloids on carbon-coated copper grids is also known. However, the use of carbon-coated copper grids as substrates is not desirable in most applications.

Prior art methods have been described for the formation of ordered planar arrays of particles under certain conditions. For example, the formation of ordered colloidal arrays in response to AC electric fields on conductive indium tin oxide (ITO) electrodes is known. However, the resulting layers were composed of small patches of ordered arrays, randomly distributed over the surface of the otherwise bare ITO substrate. Arrays of monodisperse colloidal beads and globular proteins also have been previously fabricated by using convective flow and capillary forces. However, this latter process has the disadvantage of leaving deposited particle arrays immobilized and exposed to air, making it difficult to modify arrays by subsequent liquid phase

chemistry.

The present invention provides a method of forming planar arrays with precise control over the mechanical, optical and chemical properties of the newly created layer. This method has several distinct advantages over the prior art. These result from the combination of AC electric field-induced array formation on insulating electrodes (Si/SiOx) that are patterned by UV-mediated oxide regrowth. The process of the present invention enables the formation of ordered planar arrays from the liquid phase (in which particles are originally suspended) in designated positions, and in accordance with a given overall outline. This eliminates the above-stated disadvantages of the prior art, i.e., dry state, irregular or no topography, random placement within an aggregate, immobilization of particles and uncontrolled, random placement of ordered patches on the substrate.

An advantage of the present invention is that arrays are maintained by the applied electric field in a liquid environment. The process leaves the array in a state that may be readily disassembled, subjected to further chemical modification such as cross-linking, or made permanent by chemical anchoring to the substrate. Furthermore, the liquid environment is favorable to ensure the proper functioning of many proteins and protein supramolecular assemblies of which arrays may be composed. It also facilitates the subsequent liquid-phase deposition of additional layers of molecules (by chemical binding to beads or proteins in the deposited layer), the cycling of arrays between states of different density and internal order (including complete disassembly of the array) in response to electric fields and the chemical cross-linking of particles into two-dimensionally connected layers, or gels, formed, for example, of chemically functionalized silica spheres. The present invention can be practiced on insulating electrodes such as oxide-capped silicon, to minimize Faradaic processes that might adversely affect chemical reactions involved in the gelation process or in anchoring the array to the substrate. The use of Si/SiOx electrodes also enables the control of array placement by external illumination.

The formation of colloidal arrays composed of small particles in accordance with the present invention provides a route to the fabrication of surfaces with relief structure on the scale of the particle diameter. Aside from their optical properties, such "micro-rough" surfaces are of interest as substrates for the deposition of DNA in

such a way as to alleviate steric constraints and thus to facilitate enzyme access.

Particles to which the invention applies include silica spheres, polymer colloids, lipid vesicles (and related assemblies) containing membrane proteins such as bacteriorhodopsin (bR)- a light-driven proton pump that can be extracted in the form of membrane patches and disks or vesicles. Structured and functionalized surfaces composed of photoactive pigments are of interest in the context of providing elements of planar optical devices for the development of innovative display and memory technology. Other areas of potential impact of topographically structured and chemically functionalized surfaces are the fabrication of template surfaces for the controlled nucleation of deposited layer growth and command surfaces for liquid crystal alignment. The present invention also enables the fabrication of randomly heterogeneous composite surfaces. For example, the formation of arrays composed of a mixture of hydrophobic and hydrophilic beads of the same size creates a surface whose wetting and lubrication characteristics may be controlled by the composition of the deposited mixed bead array. In this way, the location of the individual beads is random, but the relative proportion of each type of bead within the array is controllable.

Example II - Assembly of Lens Arrays and Optical Diffraction Elements

The present invention can be used to fabricate lens arrays and other surface-mounted optical elements such as diffraction gratings. The functional elements of the present invention enable the placement and delineation of these elements on ITO, facilitating integration with existing planar display technology, and on Si/SiOx, facilitating integration with existing silicon-based device technology.

Silica or other oxide particles, polymer latex beads or other objects of high refractive index suspended in an aqueous solution, will refract light. Ordered planar arrays of beads also diffract visible light, generating a characteristic diffraction pattern of sharp spots. This effect forms the basis of holographic techniques in optical information processing applications.

A. - The present invention provides for the use of arrays of refractive colloidal beads as light collection elements in planar array formats in conjunction with low light level detection and CCD imaging. CCD and related area detection schemes

will benefit from the enhanced light collection efficiency in solid-phase fluorescence or luminescence binding assays.

This assay format relies on the detection of a fluorescence signal indicating the binding of probes to bead-anchored targets in the vicinity of the detector. To maximize through-put, it is desirable to monitor simultaneously as many binding events as possible. It is here that array formation by the methods of the present invention is particularly valuable because it facilitates the placement and tight packing of beads in the target area monitored by the CCD detector, while simultaneously providing for the additional benefit of lensing action and the resulting increase in light collection efficiency.

Increased collection efficiency has been demonstrated in experiments employing individual, large (10 micron diameter) polystyrene beads as lensing elements to image small (1 micron diameter) fluorescent polystyrene beads. Under the experimental conditions set forth above an applied voltage of 5V (pp) at 300 Hz induced the collection of small particles under individual large beads within a second. This is shown in Fig. 5, where small beads alone, e.g., 52, appear dim, whereas small beads, e.g., 54, gathered under a large bead 56 appear brighter and magnified. The small beads redisperse when the voltage is turned off.

B. - The use of colloidal bead arrays as diffraction gratings and thus as holographic elements is known. Diffraction gratings have the property of diffracting light over a narrow range of wavelengths so that, for given angle of incidence and wavelength of the illuminating light, the array will pass only a specific wavelength (or a narrow band of wavelengths centered on the nominal value) that is determined by the inter-particle spacing. Widely discussed applications of diffraction gratings range from simple wavelength filtering to the more demanding realization of spatial filters and related holographic elements that are essential in optical information processing.

The present invention provides for a rapid and well controlled process of forming planar arrays in a state of crystalline order which will function as surface-mounted optical diffraction elements. In addition, the resulting surfaces may be designed to display topographical relief to enhance wave-length selective reflectivity. These arrays may be formed in designated areas on a substrate surface. In contrast to the slow and cumbersome prior art method of fabricating such arrays by way of forming

equilibrium crystals in aqueous solutions of low salt content, the present invention provides a novel approach to rapidly and reliably fabricate particle arrays at a solid-liquid interface. This approach relies on field-induced formation of arrays to trigger the process, and on UV-mediated patterning or light control to position and shape the arrays. In addition, the inter-particle distance, and internal state of order, and hence the diffraction characteristics of the array, may be fine-tuned by adjusting the applied electric field. For example, a field-induced, reversible order-disorder transition in the array will alter the diffraction pattern from one composed of sharp spots to one composed of a diffuse ring. The assembly of such arrays on the surface of silicon wafers, as described herein, provides a direct method of integration into existing microelectronic designs. Arrays may be locked in place by chemical coupling to the substrate surface, or by relying on van der Waals attraction between beads and substrate.

Example III - A Novel Mechanism for the Realization of a Particle-Based Display

The present invention provides the elements to implement lateral particle motion as a novel approach to the realization of a particle-based display. The elements of the present invention provide for the control of the lateral motion of small particles in the presence of a pre-formed lens array composed of large, refractive particles.

Colloidal particulates have been previously employed in flat-panel display technology. The operating principle of these designs is based on electrophoretic motion of pigments in a colored fluid confined between two planar electrodes. In the OFF (dark) state, pigments are suspended in the fluid, and the color of the fluid defines the appearance of the display in that state. To attain the ON (bright) state, particles are assembled near the front (transparent) electrode under the action of an electric field. In this latter state, incident light is reflected by the layer of particles assembled near the electrode, and the display appears bright. Prototype displays employing small reflective particles in accordance with this design are known. However, these displays suffered from a number of serious problems including: electrochemical degradation and lack of colloidal stability as a result of prolonged exposure to the high DC electric fields required to achieve acceptable switching speeds; and non-uniformities introduced by particle migration in response to field gradients inherent in the design of the addressing scheme.

The present invention provides a novel mechanism for the design of a particle-based display which takes advantage of electric field-induced array formation as well as controlled, field-induced lateral particle displacements. First, a lens array composed of colloidal beads is formed. This lens array also serves as a spacer array to maintain a well-defined gap between the bottom electrode and the top electrode that may now be placed over the (pre-formed) array. This facilitates fabrication of uniform flat panel displays with a narrow gap that is determined by the particle diameter.

Next, small colloidal particles are added to the electrolyte solution in the gap. These may be fluorescent, or may be reflecting incident white light. Under the action of an AC electric field of appropriate frequency, these small particles can be moved laterally to assemble preferentially within the footprint of a larger bead. When viewed through a larger bead, small fluorescent beads assembled under a large bead appear bright as a result of the increased light collection efficiency provided by the lensing action of the large bead; this is the ON state (Fig. 5). When moved outside the footprint of the larger bead, particles appear dim, and may be made entirely invisible by appropriate masking; this is the OFF state. The requisite lateral particle motion may be induced by a change in the applied voltage or a change in light intensity. Each large or lensing bead introduces a lateral nonuniformity in the current distribution within the electrolyte because the current is perturbed by the presence of each lensing bead.

In contrast to the prior art displays, the present invention employs AC, not DC fields, and insulating (rather than conductive) electrodes, thereby minimizing electrochemical degradation. The lateral non-uniformity introduced by the lens array is desirable because it introduces lateral gradients in the current distribution within the display cell. These gradients mediate the lateral motion of small beads over short characteristic distances set by the diameter of the large lensing beads, to effect a switching between ON and OFF states. Thus, the present invention readily accommodates existing technology for active matrix addressing.

Example IV - Separation and Sorting of Beads and Particles

The present invention can be used to implement several procedures for the separation and sorting of colloidal particles and biomolecules in a planar geometry. Specifically, these include techniques of lateral separation of beads in mixtures.

Individual beads may be removed from an array formed in response to an electric field by the application of optical tweezers.

The separation of components in a given mixture of chemical compounds is a fundamental task of analytical chemistry. Similarly, biochemical analysis frequently calls for the separation of biomolecules, beads or cells according to size and/or surface charge by electrophoretic techniques, while the sorting (most commonly into just two sub-classes) of suspended cells or whole chromosomes according to optical properties such as fluorescence emission is usually performed using field-flow fractionation including flow cytometry and fluorescence-activated cell sorting.

In a planar geometry, bead mixtures undergoing diffusion have been previously separated according to mobility by application of an AC electric field in conjunction with lithographic patterning of the electrode surface designed to promote directional drift. Essentially, as described in U. S. Patent No. 5,593,565 to Ajdari et al., the contents of which are included herein by reference, the AC or pulsing electric field is used to move small beads in a particular direction over a period of time, advancing beads of higher mobility relative to those of lower mobility. Capillary electrophoresis has been implemented in a planar geometry, see e.g., B.B. Haab and R.A. Mathies, Anal. Chem 67, 3253-3260 (1995), the contents of which are incorporated herein by reference.

The methods of the present invention may be applied in several ways to implement the task of separation, sorting or isolation in a planar geometry. In contrast to the prior art approaches, the present invention provides a significant degree of flexibility in selecting from among several available procedures, the one best suited to the particular task at hand. In some cases, more than one separation technique may be applied, and this provides the basis for the implementation of two-dimensional separation. That is, beads may be separated according to two different physical-chemical characteristics. For example, beads may first be separated by size and subsequently, by raising the applied frequency to induce chain formation, by polarizability. This flexibility offers particular advantages in the context of integrating analytical functionalities in a planar geometry. Several techniques will now be described.

i) The present invention may be used to implement "sieving" in

lateral, electric field-induced flow on surfaces patterned by UV-mediated oxide regrowth to sort beads in a mixture by size. The fundamental operations of the invention are invoked to set up directed lateral particle motion along conduits laid out by UV-mediated oxide regrowth. Conduits are designed to contain successively narrower constrictions through which particles must pass. Successively finer stages allow only successively smaller particles to pass in this "sieving" mechanism (Fig. 9a). As shown in Fig. 9a, the primary particle flow is in the direction left to right, while a transverse flow is established in the top to bottom direction utilizing an oxide profile as shown. Additionally, rows of barriers 92 made from thick oxide are positioned along the conduit with the spacing between the barriers in each row decreasing in the transverse direction. As the particles move along the conduit, the rows of barriers act to separate out smaller particles in the transverse direction. In contrast to previous methods based on electrophoretic separation, large DC electric fields, and the attendant potential problem of electrolysis and interference from electroosmotic flow in a direction opposite to the field-directed particle transport, the present invention uses AC electric fields and lateral gradients in interfacial impedance to produce transport. The present method has the advantage of avoiding electrolysis and it takes explicit advantage of electroosmotic flow to produce and control particle transport.

In addition, the use of Si/SiOx electrodes enables the use of the light-control component of the present invention to modify lateral transport of beads in real time. For example, external illumination may be employed to locally neutralize the lateral impedance gradient induced by UV-mediated oxide regrowth. Particles in these neutral "zones" would no longer experience any net force and come to rest. This principle may be used as a basis for the implementation of a scheme to locally concentrate particles into sharp bands and thereby to improve resolution in subsequent separation.

ii) The present invention may be used to implement “zone refining”, a process of excluding minority components of a mixture by size or shape from a growing crystalline array of majority component. This process explicitly depends on the capabilities of the present invention to induce directional crystallization.

The process of zone refining is employed with great success in producing large single crystals of silicon of very high purity by excluding impurities from the host

lattice. The concept is familiar from the standard chemical procedure of purification by re-crystallization in which atoms or molecules that are sufficiently different in size, shape or charge from the host species so as not to fit into the forming host crystal lattice as a substitutional impurity, are ejected into solution.

By enabling the growth of planar arrays, in a given direction and at a controlled rate, the present invention facilitates the implementation of an analogous zone refining process for planar arrays. The most basic geometry is the linear geometry. A multi-component mixture of beads of different sizes and/or shapes is first captured in a rectangular holding area on the surface, laid out by UV-patterning. Next, crystallization is initiated at one end of the holding area by illumination and allowed to slowly advance across the entire holding area in response to an advancing pattern of illumination. In general, differences of approximately 10% in bead radius trigger ejection.

iii) The present invention may be used to implement fractionation in a transverse flow in a manner that separates particles according to mobility.

Field-flow fractionation refers to an entire class of techniques that are in wide use for the separation of molecules or suspended particles. The principle is to separate particles subjected to fluid flow in a field acting transverse to the flow. A category of such techniques is subsumed under the heading of electric-field flow fractionation of which free-flow electrophoresis is a pertinent example because it is compatible with a planar geometry. Free-flow electrophoresis employs the continuous flow of a replenished buffer between two narrowly spaced plates in the presence of a DC electric field that is applied in the plane of the bounding plates transverse to the direction of fluid flow. As they traverse the electric field, charged particles are deflected in proportion to their electrophoretic mobility and collected in separate outlets for subsequent analysis. In contrast to conventional electrophoresis, free-flow electrophoresis is a continuous process with high throughput and it requires no supporting medium such as a gel.

The present invention enables the implementation of field-flow fractionation in a planar geometry. As previously discussed herein, impedance gradients imposed by UV-oxide profiling serve to mediate particle motion along the electrode surface in response to the external electric field. In a cell with a narrow gap, the resulting electrokinetic flow has a "plug" profile and this has the advantage of exposing

all particles to identical values of the flow velocity field, thereby minimizing band distortions introduced by the parabolic velocity profile of the laminar flow typically employed in free-flow electrophoresis.

A second flow field, transverse to the primary flow direction, may be employed to mediate particle separation. This deflecting flow may be generated in response to a second impedance gradient. A convenient method of imposing this second gradient is to take advantage of UV-oxide patterning to design appropriate flow fields. Both longitudinal and transverse flow would be recirculating and thus permit continuous operation even in a closed cell, in contrast to any related prior art technique.

Additional flexibility is afforded by invoking the light-control component of the present invention to illuminate the substrate with a stationary pattern whose intensity profile in the direction transverse to the primary fluid flow is designed to induce the desired impedance gradient and hence produce a transverse fluid flow. (Fig. 9b). This has the significant advantage of permitting selective activation of the transverse flow in response to the detection of a fluorescent bead crossing a monitoring window upstream. Non-fluorescent beads would not activate the transverse flow and would not be deflected. This procedure represents a planar analog of flow cytometry, or fluorescence-activated cell sorting.

iv) The invention may be used to induce the formation of particle chains in the direction normal to the plane of the electrode. The chains represent conduits for current transport between the electrodes and their formation may reflect a field-induced polarization. Chains are much less mobile in transverse flow than are individual particles so that this effect may be used to separate particles according to the surface properties that contribute to the net polarization. The effect of reversible chain formation has been demonstrated under the experimental conditions stated herein. For example, the reversible formation of chains occurs, for carboxylated polystyrene beads of 1 micron diameter, at a voltage of 15 V (pp) at frequencies in excess of 1MHz.

v) The invention may be used to isolate individual beads from a planar array.

Fluorescence binding assays in a planar array format, as described herein, may produce singular, bright beads within a large array, indicating particularly strong binding. To isolate and retrieve the corresponding beads, optical tweezers in the form

of a sharply focused laser spot, may be employed to lock onto an individual bead of interest. The light-control component of the present invention may be used in conjunction with the optical tweezers to retrieve such an individual bead by moving the array relative to the bead, or vice versa, or by disassembling the array and retaining only the marked bead. This is a rather unique capability that will be particularly useful in the context of isolating beads in certain binding assays.

Commercial instrumentation is available to position optical tweezers in the field of a microscope. Larger scale motion is facilitated by translocating the array in-situ or simply by moving the external sample fixture. This process lends itself to automation in conjunction with the use of peak-finding image analysis software and feedback control.

vi) The invention may be used to implement a light-induced array sectioning ("shearing") operation to separate fluorescent, or otherwise delineated portions of an array from the remainder. This operation makes it possible to segment a given array and to isolate the corresponding beads for downstream analysis.

The basis for the implementation of this array segmentation is the light-control component of the present invention, in the mode of driving particles from an area of a Si/SiOx interface that is illuminated with high intensity. It is emphasized here that this effect is completely unrelated to the light-induced force on beads that underlies the action of optical tweezers. The present effect which operates on large sets of particles, was demonstrated under the experimental conditions stated herein using a 100W illuminator on a Zeiss UEM microscope operated in epi-illumination. A simple implementation is to superimpose, on the uniform illumination pattern applied to the entire array, a line-focussed beam that is positioned by manipulation of beam steering elements external to the microscope. Beads are driven out of the illuminated linear portion. Other implementations take advantage of two separately controlled beams that are partially superimposed. The linear sectioning can be repeated in different relative orientations of shear and array.

Example V - Fabrication of Spatially Encoded Bead Arrays

The present invention provides a method to transfer suspensions of beads or biomolecules to the electrode surface in such a way as to preserve the spatial

encoding in the original arrangement of reservoirs, most commonly the conventional 8x12 arrangement of wells in a microtiter plate. Such a fluid transfer scheme is of significant practical importance given that compound libraries are commonly handled and shipped in 8x12 (or equivalent) wells.

The present invention utilizes chemical patterning to define individual compartments for each of MxN sets of beads and confine them accordingly. In the present instance, patterning is achieved by UV-mediated photochemical oxidation of a monolayer of thiol-terminated alkylsilane that is chemisorbed to the Si/SiOx substrate. Partial oxidation of thiol moieties produces sulfonate moities and renders the exposed surface charged and hydrophilic. The hydrophilic portions of the surface, in the form of a grid of squares or circles, will serve as holding areas.

In accordance with the present invention, the first function of surface-chemical patterning into hydrophilic sections surrounded by hydrophobic portions is to ensure that droplets, dispensed from different wells, will not fuse once they are in contact with the substrate. Consequently, respective bead suspensions will remain spatially isolated and preserve the lay-out of the original MxN well plate. The second role of the surface chemical patterning of the present invention is to impose a surface charge distribution, in the form of the MxN grid pattern, which ensures that individual bead arrays will remain confined to their respective holding areas even as the liquid phase becomes contiguous.

The layout-preserving transfer procedure involves the steps illustrated in Figs. 6a-c. First, as shown in side view in Fig. 6a, the MxN plate of wells 62 is registered with the pattern 64 on the planar substrate surface. Well bottoms 62, are pierced to allow for the formation of pendant drops of suspension or, preferably, the process is facilitated by a fixture (not shown) providing MxN effective funnels to match the geometric dimensions of the MxN plate on the top and reduce the size of the dispensing end. Such a dispensing fixture will also ensure the precise control of droplet volumes, adjusted so as to slightly overfill the target holding area on the patterned substrate surface. The set of MxN drops is then deposited by bringing them in contact with the hydrophilic holding areas of the pre-patterned substrate and relying on capillary action.

Next, the plate is retracted, and the top electrode is carefully lowered to

form the electrochemical cell, first making contact as shown in Fig. 6b, with individual liquid-filled holding areas on the substrate to which suspensions are confined. Overfilling ensures that contact is made with individual suspensions. The electric field is now turned on to induce array formation in the MxN holding areas and to ensure the preservation of the overall configuration of the MxN sets of beads while the gap is closed further (or filled with additional buffer) to eventually fuse individual droplets of suspension into a contiguous liquid phase as shown in Fig. 6c. In the fully assembled cell of Fig. 6c, while the droplets are fused together, the beads from each droplet are maintained in and isolated in their respective positions, reflecting the original MxN arrangement of wells. The present invention thus provides for the operations required in this implementation of a layout-preserving transfer procedure to load planar electrochemical cells.

Example VI - Fabrication of Dynamic Planar Bead Arrays for Parallel Assays

The present invention provides a method to produce a heterogeneous panel of beads and potentially of biomolecules for presentation to analytes in an adjacent liquid. A heterogeneous panel contains particles or biomolecules which differ in the nature of the chemical or biochemical binding sites they offer to analytes in solution. The present method relies on the functional elements of the invention to assemble a planar array of a multi-component mixture of beads which carry chemical labels in the form of tag molecules and may be so identified subsequent to performing the assay. In the event of binding, the analyte is identified by examination of the bead, or cluster of beads, scoring positive.

Diagnostic assays are frequently implemented in a planar format of a heterogeneous panel, composed of simple ligands, proteins and other biomolecular targets. For example, in a diagnostic test kit, a heterogeneous panel facilitates the rapid testing of a given analyte, added in solution, against an entire set of targets. Heterogeneous panels of proteins are of great current interest in connection with the emerging field of proteome research. The objective of this research is to identify, by scanning the panel with sensitive analytical techniques such as mass spectrometry, each protein in a multi-component mixture extracted from a cell and separated by two-dimensional gel electrophoresis. Ideally, the location of each spot uniquely corresponds to one particular

protein. This analysis would permit, for example, the direct monitoring of gene expression levels in a cell during a particular point in its cycle or at a given stage during embryonic development.

The fabrication of an array of heterogeneous targets is central to recently proposed strategies of drug screening and DNA mutation analysis in a planar format. The placement of ligands in a specific configuration on the surface of a planar substrate serves to maintain a key to the identity of any one in a large set of targets presented simultaneously to an analyte in solution for binding or hybridization. In an assay relying on fluorescence, binding to a specific target will create bright spots on the substrate whose spatial coordinates directly indicate the identity of the target.

Three principal strategies have been previously employed to fabricate heterogeneous panels. First, protein panels may be created by two-dimensional gel electrophoresis, relying on a DC electric field to separate proteins first by charge and then by size (or molecular weight). Even after many years of refinement, this technique yields results of poor reproducibility which are generally attributed to the poorly defined properties of the gel matrix.

Second, individual droplets, drawn from a set of reservoirs containing solutions of the different targets, may be dispensed either by hand or by employing one of several methods of automated dispensing (or "printing"; see e.g., Schena et al., Science 270, 467-470 (1995), the contents of which are incorporated herein by reference). Printing has been applied to create panels of oligonucleotides intended for screening assays based on hybridization. Printing leaves a dried sample and may thus not be suitable for proteins that would denature under such conditions. In addition, the attendant fluid handling problems inherent in maintaining, and drawing samples from a large number of reservoirs are formidable.

Third, target ligands may be created by invoking a variant of solid phase synthesis based on a combinatorial strategy of photochemically activated elongation reactions. This approach has been limited by very formidable technical problems in the chemical synthesis of even the simplest, linear oligomers. The synthesis of non-linear compounds in this planar geometry is extremely difficult.

The present invention of forming heterogeneous panels requires the chemical attachment of target ligands to beads. Ligands may be coupled to beads

"off-line" by a variety of well established coupling reactions. For present purposes, the bead identity must be chemically encoded so it may be determined as needed. Several methods of encoding, including binary encoding, of beads are available. For example, short oligonucleotides may serve the purpose of identifying a bead via their sequence which may be determined by microscale sequencing techniques. Alternatively, chemically inert molecular tags may be employed that are readily identified by standard analytical techniques.

In contrast to all prior art methods, the present invention provides a novel method to create heterogeneous panels by in-situ, reversible formation of a planar array of chemically encoded beads in solution adjacent to an electrode. The array may be random with respect to chemical identity but is spatially ordered. This procedure offers several advantages. First, it is reversible so that the panel may be disassembled following the binding assay to discard beads scoring negative. Positive beads may be subjected to additional analysis without the need for intermediate steps of sample retrieval, purification or transfer between containers. Second, the panel is formed when needed, that is, either prior to performing the actual binding assay, or subsequent to performing the assay on the surface of individual beads in suspension. The latter mode minimizes potential adverse effects that can arise when probes bind to planar target surfaces with a high concentration of target sites. Third, to accommodate optical analysis of individual beads, interparticle distances within the array may be adjusted by field-induced polarization or by the addition of inert spacer particles that differ in size from the encoded beads. Fig. 7 shows the use of small spacer beads 72 for separating encoded beads 74. As shown, the spacing of beads 74 is greater than the spacing of comparable beads in Fig. 4b. Finally, UV-mediated oxide regrowth, as provided by the present invention, readily facilitates the embedding of a grid pattern of selected dimension into the substrate to ensure the formation of small, layout-preserving subarrays in the low-impedance fields of the grid.

To create the panel, a multi-component mixture of beads carrying, for example, compounds produced by bead-based combinatorial chemistry, is placed between electrodes. Each type of bead may be present in multiple copies. Arrays are formed in response to an external field in a designated area of the electrode surface. This novel approach of in-situ assembly of panels relies on beads that carry a unique

chemical label, or code, to permit their identification subsequent to the completion of a binding assay. Alternatively, beads may be marked ("painted") on-line by way of a photochemical bead-coloring method. Selected beads in an array are individually illuminated by a focused light source to trigger a coloring reaction on the bead surface or in the bead interior to indicate a positive assay score. Beads so marked can be subsequently separated from unmarked beads by a light-activated sorting method described herein. Numerous UV-activated reactions are available to implement this bead-coloring method.

The present invention provides for several methods of discarding beads with negative scores, typically the vast majority, while retaining those with positive scores. This method take advantage of the fact that, in contrast to all prior art methods, the array represents a temporary configuration of particles that is maintained by the applied electric field and may be rearranged or disassembled at will. This capability, along with the fact that biomolecules are never exposed to air (as in the prior art method of printing) facilitates the in-situ concatenation of analytical procedures that require the heterogeneous panel in conjunction with subsequent, "downstream" analysis.

First, if positive beads are clustered in a subsection of the array, the light-controlled array splitting operation of the present invention may be invoked to dissect the array so as to discard negative portions of the array (or recycle them for subsequent use). Second, if positive and negative beads are randomly interspersed, a fluorescence-activated sorting method, implemented on the basis of the present invention in a planar format, as described herein, may be invoked. In the case of fluorescence-activated sorting, positive and negative beads may be identified as bright and dark objects, respectively. In the special case that only a few positive beads stand out, these may be removed from the array by locking onto them with optical tweezers, a tool to trap and/or manipulate individual refractive particles under illumination, and disassembling the array by removing the field, or subjecting the entire array to lateral displacement by the fundamental operations of the present invention.

The typical task in screening a large set of compounds is one of looking for a very small number of positive events in a vast number of tests. The set of discarded beads will typically involve the majority at each stage in the assay. The procedure of the present invention therefore minimizes the effort invested in negative

events, such as the challenging in-situ synthesis of target ligands irrespective of whether or not they will prove to be of interest by binding a probe offered in solution.

The method of forming a heterogeneous panel according to the present invention contains beads of each type in generally random assembly. The creation of a heterogeneous panel with each position in the panel containing a cluster of beads of the same type, that is, beads originating in the same reservoir (Fig. 6a), may be desirable so as to ensure a sufficiently large number of positive events to facilitate detection. A practical solution follows from the application of the layout-preserving fluidic transfer scheme described herein. In this procedure, beads from an MxN well plate are transferred layout-preservingly onto a chemically patterned substrate in such a way as to preserve the spatial encoding of bead identities.

Example VII - Binding and Functional Assays in Planar Bead Array Format

The present invention can be used to implement mixed-phase binding assays as well as certain functional assays in a planar array format. Several combinations are possible reflecting the presence of probe or target in solution, on the surface of colloidal beads, or on the electrode surface. The methods of the present invention facilitate the formation of a planar array to present targets to probes in solution prior to performing the binding assay ("pre-formed" array; Fig. 8). Alternatively, a planar array of beads may be formed in front of a detector surface subsequent to performing the binding assay in suspension ("post-formed" array; Fig. 8). The present invention also provides the methods to implement functional assays by enabling the assembly of certain cell types adjacent to a planar detector or sensor surface to monitor the effects of exposure of the cells to small molecule drugs in solution.

Binding assays, particularly those involving proteins such as enzymes and antibodies, represent a principal tool of medical diagnostics. They are based on the specific biochemical interaction between a probe, such as a small molecule, and a target, such as a protein. Assays facilitate the rapid detection of small quantities of an analyte in solution with high molecular specificity. Many procedures have been designed to produce signals to indicate binding, either yielding a qualitative answer (binding or no binding) or quantitative results in the form of binding or association constants. For example, when an enzyme binds an analyte, the resulting catalytic reaction may be used

to generate a simple color change to indicate binding, or it may be coupled to other processes to produce chemical or electrical signals from which binding constants are determined. Monoclonal antibodies, raised from a single common precursor, may be prepared to recognize virtually any given target, and immunoassays, based on antibody-antigen recognition and binding, have developed into an important diagnostic tool. As with enzyme binding, antibody binding of an antigenic analyte may be detected by a variety of techniques including the classic method of enzyme-linked immunoassays (ELISA) in which the reaction of an antibody-coupled enzyme is exploited as an indicator. A common and conceptually simple scheme ensures the detection of antibody binding to a target analyte by supplying a fluorescently labeled second antibody that recognizes the first (or primary) antibody.

Binding assays involving soluble globular proteins are often performed in solution to ensure unbiased interactions between protein and target. Such liquid phase assays, especially when performed at low concentrations of target or probe, minimize potential difficulties that may arise when either target or probe are present in abundance or in close proximity. By the same token, the kinetics tend to be slow. Cooperative effects, such as crowding, arising from the close proximity of probes must be carefully controlled when either probe or target is chemically anchored to a solid substrate.

Nonetheless, this latter solid phase format of binding assays is also very commonly employed whenever the situation demands it. For example, the presence of a protein on the surface of a cell may be exploited in "panning" for the cells that express this protein in the presence of many other cells in a culture that do not: desired cells attach themselves to the surface of a container that is pre-coated with a layer of a secondary antibody directed against a primary antibody decorating the desired cell-surface protein. Similarly, certain phages may be genetically manipulated to display proteins on their surface, and these may be identified by a binding assay involving a small molecule probe such as an antigen if the protein displayed is an antibody (Watson et al., "Recombinant DNA", 2nd Edition (Scientific American Books, W.H. Freeman and Co., New York , NY, 1983), the contents of which are incorporated herein by reference). In addition, the planar geometry accommodates a variety of optical and electrical detection schemes implemented in transducers and sensors.

A combination of liquid phase and solid phase assay may be developed by

using beads that are decorated with either probe or target, as in procedures that employ decorated magnetic beads for sample preparation or purification by isolating binding from non-binding molecules in a given multi-component mixture. Recent examples of the use of these beads include the purification of templates for DNA sequencing applications or the extraction of mRNAs from (lysed) cells by hybridization to beads that are decorated with poly-adenine (polyA) residues.

Functional assays involving suitable types of cells are employed to monitor extracellular effects of small molecule drugs on cell metabolism. Cells are placed in the immediate vicinity of a planar sensor to maximize the local concentration of agents released by the cell or to monitor the local pH.

The present invention provides the means to implement mixed phase binding assays in a planar geometry with a degree of flexibility and control that is not available by prior art methods. Thus, it offers the flexibility of forming, in-situ, reversibly and under external spatial control, either a planar panel of target sites for binding of analyte present in an adjacent liquid phase, or a planar array of probe-target complexes subsequent to performing a binding assay in solution. Binding may take place at the surface of individual beads suspended in solution, at the surface of beads pre-assembled into arrays adjacent to the electrode surface, or at the electrode surface itself. Either the target or probe molecule must be located on a bead to allow for a bead-based assay according to the present invention. As shown in Fig. 8, if the probe molecule P is located on a bead, then the target molecule T may be either in solution, on a bead or on the electrode surface. The converse is also true.

For example, the methods of the present invention may be used to implement panning, practiced to clone cell surface receptors, in a far more expeditious and controlled manner than is possible by the prior art method. Given a substrate that has been coated with a layer of antibody directed against the sought-after cell surface protein, the present invention facilitates the rapid assembly of a planar array of cells or decorated beads in proximity to the layer of antibodies and the subsequent disassembly of the array to leave behind only those cells or beads capable of forming a complex with the surface-bound antibody.

A further example of interest in this category pertains to phage displays. This technique may be employed to present a layer of protein targets to bead-anchored

probes. Bead arrays may now be employed to identify a protein of interest. That is, beads are decorated with small molecule probes and an array is formed adjacent to the phage display. Binding will result in a probe-target complex that retains beads while others are removed when the electric field is turned off, or when light-control is applied to remove beads from the phage display. If beads are encoded, many binding tests may be carried out in parallel because retained beads may be individually identified subsequent to binding.

The methods of the present invention readily facilitate competitive binding assays. For example, subsequent to binding of a fluorescent probe to a target-decorated bead in solution and the formation of a planar bead array adjacent to the electrode, fluorescent areas within the array indicate the position of positive targets, and these may be further probed by subjecting them to competitive binding. That is, while monitoring the fluorescence of a selected section of the planar array, an inhibitor (for enzyme assays) or other antagonist (of known binding constant) is added to the electrochemical cell, and the decrease in fluorescence originating from the region of interest is measured as a function of antagonist concentration to determine a binding constant for the original probe. This is an example of a concatenation of analytical steps that is enabled by the methods of the present invention.

The fact that a probe-target complex is fixed to a colloidal bead, as in the methods of the present invention, conveys practical advantages because this facilitates separation of positive from negative events. Particularly when solid phase assays are performed on a planar substrate, an additional advantage of planar bead arrays is the enhancement of light collection efficiency provided by the beads, as discussed herein.

If desired, beads may serve strictly as delivery vehicles for small molecule probes. That is, an array of probe-decorated beads is formed adjacent to a target-decorated surface in accordance with the methods of the present invention. UV-activated cleavage of the probe from the bead support will ensure that the probe is released in close proximity to the target layer, thereby enhancing speed and efficiency of the assay. The identity of the particular probe interacting with the target may be ascertained from the positional location of the bead delivering the probe.

The methods of the present invention apply not only to colloidal beads of a wide variety (that need no special preparative procedures to make them magnetic, for

example), but also to lipid vesicles and cells that are decorated with, or contain embedded in their outer wall, either probe or target. The methods of the present invention may therefore be applied not only to bead-anchored soluble proteins but potentially to integral membrane receptors or to cell surface receptors.

In particular, the rapid assembly of cells in a designated area of the substrate surface facilitates the implementation of highly parallel cell-based functional assays. The present invention makes it possible to expose cells to small molecule drug candidates in solution and rapidly assemble them in the vicinity of a sensor embedded in the electrode surface, or to expose pre-assembled cells to such agents that are released into the adjacent liquid phase. In the simplest case, all cells will be of the same type, and agents will be administered sequentially. Even in this sequential version, electrokinetic mixing will enhance through-put. However, as described herein, the methods of the present invention also enable the parallel version of binding assays and thus of functional assays in a planar format by encoding the identity of different cells by a "Layout-Preserving Transfer" process from an 8x12 well plate, as discussed herein, and to isolate cells scoring positive by providing feed-back from a spatially resolved imaging or sensing process to target a specific location in the array of cells.

Example VIII - Screening for Drug Discovery in Planar Geometry

The functional elements of the present invention may be combined to implement procedures for handling and screening of compound and combinatorial libraries in a planar format. The principal requisite elements of this task are: sample and reagent delivery from the set of original sample reservoirs, commonly in a format of 8x12 wells in a microtiter plate, into a planar cell; fabrication of planar arrays of targets or of probe-target complexes adjacent to the planar electrode surface prior to or subsequent to performing a binding assay; evaluation of the binding assay by imaging the spatial distribution of marker fluorescence or radioactivity, optionally followed by quantitative pharmacokinetic measurements of affinity or binding constants; isolation of beads scoring positive, and removal from further processing of other beads; and collection of specific beads for additional downstream analysis. The present invention relates to all of these elements, and the fundamental operations of the invention provide the means to concatenate these procedures in a planar format.

A central issue in the implementation of cost-effective strategies for modern therapeutic drug discovery is the design and implementation of screening assays in a manner facilitating high throughput while providing pharmacokinetic data as a basis to select promising drug leads from a typically vast library of compounds. That is, molecular specificity for the target, characterized by a binding constant, is an important factor in the evaluation of a new compound as a potential therapeutic agent. Common targets include enzymes and receptors as well as nucleic acid ligands displaying characteristic secondary structure.

The emerging paradigm for lead discovery in pharmaceutical and related industries such as agricultural biotechnology, is the assembly of novel synthetic compound libraries by a broad variety of new methods of solid state "combinatorial" synthesis. Combinatorial chemistry refers to a category of strategies for the parallel synthesis and testing of multiple compounds or compound mixtures in solution or on solid supports. For example, a combinatorial synthesis of a linear oligopeptide containing n amino acids would simultaneously create all compounds representing the possible sequence permutations of n amino acids. The most commonly employed implementation of combinatorial synthesis relies on colloidal bead supports to encode reaction steps and thus the identity of each compound. Beads preferred in current practice tend to be large (up to 500 microns in diameter) and porous to maximize their compound storage capacity, and they must be encoded to preserve the identity of the compound they carry.

Several methods of encoding, or binary encoding, of beads are available. Two examples are as follows. First, beads may be labeled with short oligonucleotides such as the 17-mers typically employed in hybridization experiments. The sequence of such short probes may be determined by microscale sequencing techniques such as direct Maxam-Gilbert sequencing or mass spectrometry. This encoding scheme is suitable when the task calls for screening of libraries of nucleic acid ligands or oligopeptides. Second, members of a combinatorial library may be associated with chemically inert molecular tags. In contrast to the previous case, these tag molecules are not sequentially linked. Instead, the sequence of reaction steps is encoded by the formal assignment of a binary code to individual tag molecules and their mixtures that are attached to the bead in each successive reaction step. The tags are readily identified by standard analytical

techniques such as gas chromatography. This general encoding strategy is currently employed in the synthesis of combinatorial libraries on colloidal beads.

Commercial compound libraries are large, given that even for the aforementioned 17-mer, the number of sequence permutations is 4^17, or approximately 10^10. However, the high specificity of typical biological substrate-target interactions implies that the vast majority of compounds in the collection will be inactive for any one particular target. The task of screening is to select from this large set the few potential lead compounds displaying activity in binding or in functional assays. The principal drug discovery strategy widely applied to natural compound libraries in the pharmaceutical industry is to select individual compounds from the library at random and subject them to a series of tests. Systematic screening procedures are thus required to implement the rapid screening and scoring of an entire library of synthetic compounds, in practice often containing on the order of 10^7 items.

In current practice, compounds are first cleaved and eluted from their solid supports and are stored in microtiter plates. Further sample handling in the course of screening relies primarily on robotic pipetting and transfer between different containers, typically wells in microtiter plates. While robotic workstations represent a step in the direction of automating the process, they rely on the traditional format of microtiter plates containing 8x12 wells and sample handling by pipetting and thus represent merely an incremental operational improvement. A significant additional consideration is the need to conserve reagent and sample by reducing the spatial scale of the analytical procedures.

The present invention provides a set of operations to realize integrated sample handling and screening procedures for bead-based compound libraries in a planar format. This will significantly reduce time and cost due to reagent and sample volumes. The principal advantage of the methods of the present invention is that they provide a large set of fundamental operations to manipulate sets of beads in a planar format, permitting the handling of beads between stations in a multi-step analytical procedure.

In particular, as previously described herein, the methods of the present invention facilitate the implementation of the following pertinent procedures: transfer of samples from microtiter plates to a planar electrochemical cell; formation of heterogeneous panels of target sites adjacent to the substrate surface; solid phase binding

assays; and isolation of specific beads from an array. In addition, the fundamental operations of the present invention provide the means to concatenate these procedures on the surface of a planar electrode.

As described herein for hybridization assays, several variants are possible. That is, binding assays may be performed by allowing protein targets such as enzymes to bind to compounds on the surface of a bead, either in suspension or arranged in a planar array. The common practice of combinatorial chemistry based on large porous carrier beads accommodates the concurrent handling of smaller beads to whose outer surface compounds are anchored via inert chemical spacers. Such small beads (up to 10 microns in diameter) are readily manipulated by the methods of the present invention. Large beads are used as labeled compound storage containers.

Alternatively, binding between target and a radioactively or otherwise labeled probe may occur in solution, within microtiter plate wells, if compounds have already been cleaved from their synthesis support. In that case, probe-target complexes may be captured by complexation to encoded beads in each well, for example via the secondary antibody method of coupling the protein target to a bead-anchored antibody. Bead-captured probe-target complexes are then transferred to the planar cell for proximity analysis and further processing as illustrated in Fig. 10. As shown in Fig. 10, probe-target complexes 102 are allowed to form in solution. Antibody coated beads 104 are added to the solution, resulting in a bead anchored complex 106. The bead anchored complexes 106 are deposited onto electrode 108 from wells 110, and a planar array of bead anchored complexes is formed. When fluorescent probes 114 are used, these impart fluorescence to the bead anchored complex, facilitating detection.

The methods and apparatus of the present invention are well suited to the task of identifying a small number of positive events in a large set. The imaging of an entire array of probe-target complexes is further enhanced by proximity to an area detector, and by bead lensing action. The isolation of a small number of positive scores from the array is readily achieved, for example by applying optical tweezers, as described herein. The large remainder of the array may then be discarded. This in turn considerably reduces the complexity of applying more stringent tests, such as the determination of binding constants, because these may be restricted to the few retained beads. These tests may be directly applied, without the need for additional sample

transfer to new containers, to the samples surviving the first screening pass.

Example IX - Hybridization Assays in Planar Array Format

The present invention can be used to implement solid phase hybridization assays in a planar array format in a configuration related to that of a protein binding assay in which target molecules are chemically attached to colloidal beads. The methods of the present invention facilitate the formation of a planar array of different target oligonucleotides for presentation to a mixture of strands in solution. Alternatively, the array may be formed subsequent to hybridization in solution to facilitate detection and analysis of the spatial distribution of fluorescence or radioactivity in the array.

Considerable research and development is presently being invested in an effort to develop miniaturized instrumentation for DNA sample extraction and preparation including amplification, transcription, labeling and fragmentation, with subsequent analysis based on hybridization assays as well as electrophoretic separation. Hybridization assays in planar array format are being developed as a diagnostic tool for the rapid detection of specific single base pair mutations in a known segment of DNA, and for the determination of expression levels of cellular genes via analysis of the levels of corresponding mRNAs or cDNAs. Hybridization of two complementary single strands of DNA involves molecular recognition and subsequent hydrogen bond formation between corresponding nucleobases in the two opposing strands according to the rules A-T and G-C; here A, T, G and C respectively represent the four nucleobases Adenine, Thymine, Guanosine and Cytosine found in DNA; in RNA, Thymine is replaced by Uracil. The formation of double-strand, or duplex, DNA requires the pairing of two highly negatively charged strands of DNA, and the ionic strength of the buffer, along with temperature, plays a decisive role.

As previously discussed herein, two principal methods to prepare heterogeneous arrays of target strands on the surface of a planar substrate are micro-dispensing ("printing") and in-situ, spatially encoded synthesis of oligonucleotides representing all possible sequence permutations for a given total length of strand. In this context, hybridization must necessarily occur in close proximity to a planar substrate surface and this condition requires care if complications from steric hindrance and from non-specific binding of strands to the substrate are to be avoided. Non-specific

adsorption can be a serious problem, especially in the presence of DC electric fields employed in current commercial designs that rely on electrophoretic deposition to accelerate the kinetics of hybridization on the surface. In addition, there are the technical difficulties, previously discussed herein, resulting from steric hindrance and from collective effects reflecting the crowding of probe strands near the surface.

In the context of DNA analysis, colloidal (magnetic) beads are commonly used. For example, they are employed to capture DNA in a widely used screening procedure to select cDNAs from clone libraries. Specifically, cDNAs are allowed to hybridize to sequences within long genomic DNA that is subsequently anchored to magnetic beads to extract the hybridized cDNA from the mixture.

The present invention facilitates the formation of planar arrays of oligonucleotide-decorated colloidal beads, either prior to or subsequent to hybridization of a fluorescence probe strand to the bead-anchored target strand or subsequent to hybridization in free solution and bead capture of the end-functionalized target strand. In contrast to prior art methods, the present invention does not require hybridization to occur in the vicinity of planar substrate surface, although this is an option if bead-anchored probe strands are to be delivered to substrate-anchored target strands.

The ability to perform hybridization either in solution, on the surface of individual beads, or at the substrate surface provides an unprecedented degree of flexibility. In addition, the advantages of bead arrays, as described herein, make it feasible to select and isolate individual beads, or groups of beads, from a larger array on the basis of the score in a hybridization assay. This isolation facilitates the implementation of subsequent assays on the strands of interest. The fact that beads remain mobile also means that beads of interest may be collected in designated holding areas for micro-sequencing, or may be moved to an area of substrate designated for PCR amplification.

The methods of the present invention may be used to implement a hybridization assay in a planar array format in one of two principal variations. All involve the presence of the entire repertoire of beads in the planar array or panel formed adjacent to the electrode surface for parallel read-out. As with heterogeneous panels in general, the arrangement of beads within the array is either random (with respect to chemical identity), and the identity of beads scoring high in the binding assay must be

determined subsequently, or it is spatially encoded by invoking the "Layout-Preserving Transfer" method of sample loading described herein.

The former variant is readily implemented and accommodates array formation either prior to or subsequent to performing the binding assay. For example, binding may be performed in suspension before beads are assembled into the array. As with the aforementioned cDNA selection procedure, the method of the present invention also accommodates the use of beads as capture elements for end-functionalized target DNA, for example, via biotin-streptavidin complexation. In this latter case, beads serve as a delivery vehicle to collect all probe-target complexes to the electrode surface where they are assembled into an array for ease of analysis. In particular, proximity CCD detection of beads on electrodes will benefit from the lensing action of the beads in the array. This version of the assay is preferably used if only a small number of positive scores are expected.

Hybridization to a pre-formed bead array can take advantage of a variant of the assay which preserves spatial encoding. An array of bead clusters is formed by the "Layout-Preserving Transfer" method previously described herein, and exposed to a mixture of cDNAs. The resulting spatial distribution of fluorescence intensity or radioactivity reflects the relative abundance of cDNAs in the mixture. This procedure relies on the detection of a characteristic fluorescence or other signal from the probe-target complex on the surface of a single bead. Given the fact that the array is readily held stationary by the methods of the present invention, image acquisition may be extended to attain robust signal-to-noise for detection of low level signals. For example, a signal generated by a bead of 10 micron diameter with at most 10^8 probe-target complexes on the surface of the bead may be detected. Bead lensing action also aids in detection.

As with the implementation of drug screening, the functional elements of the present invention may be combined to perform multiple preparative and analytical procedures on DNA.

Example X - Alignment and Stretching of DNA in Electric Field-Induced Flow

The present invention can be used to position high-molecular weight DNA in its coiled configuration by invoking the fundamental operations as they apply to other

colloidal particles. However, in addition, the electrokinetic flow induced by an electric field at a patterned electrode surface may be employed to stretch out the DNA into a linear configuration in the direction of the flow.

Procedures have been recently introduced which rely on optical imaging to construct a map of cleavage sites for restriction enzymes along the contour of an elongated DNA molecule. This is generally known as a "restriction map". These procedures, which facilitate the study of the interaction of these and other proteins with DNA and may also lead to the development of techniques of DNA sequencing, depend on the ability to stretch and align DNA on a planar substrate.

For individual DNA molecules, this has been previously achieved by subjecting the molecule to elongational forces such as those exerted by fluid flow, magnetic fields acting on DNA-anchored magnetic beads or capillary forces. For example, DNA "combs" have been produced by simply placing DNA molecules into an evaporating droplet of electrolyte. If provisions are made to promote the chemical attachment of one end of the molecule to the surface, the DNA chain is stretched out as the receding line of contact between the shrinking droplet and the surface passes over the tethered molecules. This leaves behind dry DNA molecules that are attached in random positions within the substrate area initially covered by the droplet, stretched out to varying degrees and generally aligned in a pattern of radial symmetry reflecting the droplet shape. Linear "brushes", composed of a set of DNA molecules chemically tethered by one end to a common line of anchoring points, have also been previously made by aligning and stretching DNA molecules by dielectrophoresis in AC electric fields applied between two metal electrodes previously evaporated onto the substrate.

The present invention invokes electrokinetic flow adjacent to an electrode patterned by UV-mediated regrowth of oxide to provide a novel approach to the placement of DNA molecules in a predetermined arrangement on a planar electrode surface, and to the stretching of the molecules from their native coil configuration into a stretched, linear configuration that is aligned in a pre-determined direction. This process is shown in Fig. 11 and is accomplished by creating controlled gradients in the flow vicinity across the dimension of the DNA coil. The velocity gradient causes different portions of the coil to move at different velocities thereby stretching out the coil. By maintaining a stagnation point at zero velocity, the stretched coil will be fixed

in position. This method has several advantages over the prior art approaches. First, DNA molecules in their coiled state are subjected to light control to form arrays of desired shape in any position on the surface. This is possible because large DNA from cosmids or YACs forms coils with a radius in the range of one micron, and thus acts in a manner analogous to colloidal beads. A set of DNA molecules may thus be steered into a desired initial arrangement. Second, UV-patterning ensures that the elongational force created by the electrokinetic flow is directed in a predetermined direction. The presence of metal electrodes in contact with the sample, a disadvantage of the dielectrophoretic prior art method, is avoided by eliminating this source of contamination that is difficult to control especially in the presence of an electric field. On patterned Si/SiOx electrodes, flow velocities in the range of several microns/second have been generated, as required for the elongation of single DNA molecules in flow. Thus, gradients in the flow field determines both the fractional elongation and the orientation of the emerging linear configuration. Third, the present invention facilitates direct, real-time control of the velocity of the electric field-induced flow, and this in turn conveys explicit control over the fractional elongation.

This invention is also for a method and apparatus to direct the lateral motion and induce the assembly into planar arrays of cells on semiconductor surfaces in response to temporally and spatially varying electric fields and to projected patterns of illumination.

The response of different types of cells to variations in the frequency and voltage of the applied electric field and to specific patterns and changing intensity of illumination enables the on-cue assembly of multiple arrays of cells in any desired position on the substrate, and further enables the repositioning, disassembly and re-assembly and more generally the reconfiguration and segmentation of such arrays. Time-constant and time-varying illumination patterns are generated by an optically programmable illumination pattern generator may be used and can be based on a graphical representation of the pattern created on a personal computer. The novel methodology of image cytometry introduced in the present invention relates to the general fields of molecular and cellular biology and has applications in the fields of diagnostics, genetic analysis, cell biology and drug discovery. The image cytometry facilitates a multiplicity of cell-based biological and biochemical procedures based on the

quantitative analysis of images recorded from array of cells. As with Random Encoded Array Detection (READ) or Programmable Array Reconfiguration and Segmentation (PARSE), these assays are performed in a highly parallel format in a small volume of sample on the surface of a semiconductor wafer.

The novel method of image cytometry according to the present invention may be used in a number of different applications, including:

- cell-typing analysis based on the decoration of cell-surface markers with tagged probes such as fluorescence labeled antibodies or antibodies that are coupled to flourescent beads. Specific applications include bacterial strain typing; HLA typing (transplantation donor-recipient matching); immunophenotyping;

- diagnostic assays based on the identification of cell types and quantitative evaluation of relative counts of different cell types in a given sample, for example, the determination of the fraction of CD-4 cells in a population to assess levels of HIV infection;

- on-chip fractionation of a mixture of cells into sub-populations based on:

- selective affinity of cell phenotypes using antibodies or lectin (Nordon et al, "An Experimental Model of Affinity Cell Separation", Cytometry 16, 25-33 (1994), the contents of which are incorporated herein by reference; see also: "panning" for clones expressing specific antibodies). Using beads displaying such antibodies or lectins, cross-linking between cells may be induced to form clusters that are readily distinguished and separated from individual cells;

- differential mobility in response to time-varying or to time-constant but spatially varying (inhomogeneous) electric fields or to gradients in electrochemical environment created by LEAPS near an electrolyte-insulator-semiconductor (EIS) interface; - combinations with controlled flow configurations such as those in various realizations of field flow fractionation (for example, see: Bigelow et al, "Separation of B and T lymphocytes by a hybrid field-flow fractionation/adhesion chromatography technique", J. Immunological Methods 117, 289-293 (1989); Markx et al, "Dielectrophoretic characterization and separation of microorganisms", Microbiology 140, 585-591 (1994); Huang et al, "Introducing Dielectrophoresis as a New Force Field for Field Flow Fractionation", Biophysical J. 73, 1118-1129 (1997); the contents of all of which are incorporated herein by reference);

- cell-based functional assays monitoring the levels of expressed cell-surface markers or the activation or regulation of pathways, such as signal transduction pathways and metabolic pathways. The expression of the surface markers may be detected by affinity binding between the markers and fluorescence or enzyme labeled probes. The endpoints of the pathways may be analyzed through intracellular fluorescent signal generated by enzymes such as luciferin (PE Corporation/Tropix) or β-lactamase (Aurora);

- cell secretion-based analysis that could be used in drug and ligand screening;

- cell-cell interactions revealed by adhesion markers on cell surfaces;

- cytotoxicity assays based on the labeling of indicators such as phosphatidylserine which appears on the outer membrane of cells in the early stages of apoptosis; and

- chemotaxis analysis based on the tracking of cells positioned in controlled configurations in known spatial relation to gradients in the concentration of chemotactic agents.

The image cytometry according to the present invention permits the efficient realization of a wide variety of cell-based functional as well as diagnostic assays based on the analysis of multi-color fluorescence or other optical markers imaged on the surfaces of individual cells. In addition, the methodology also permits the separation and fractionation of multiple cell types in a mixture based on differential response to electro-optic stimuli as disclosed herein. Imaging and profiling applications may invoke a variation of the READ format while fractionation and sorting into multiple sub-populations may invoke the PARSE format.

The integration of this novel methodology of cellular analysis with sample preparation (electric-field induced lysis of cells) along with the technology of optically programmable bead arrays as disclosed herein creates a novel platform that combines an "on-chip" format of quantitative molecular interaction analysis and cellular analysis with microfluidics and optical imaging into a novel, multi-purpose assay platform.

Imaging of cells - In its simple form, image cytometry is practiced in the doctor's office when blood or other liquid samples are spotted onto segmented glass slides and populations of cells are "manually" counted. These and related simple implementations of cellular analysis invoke the simple spotting of cells onto solid

supports where cells settle under gravity, typically in random positions at low lateral density. However, even in this simple format, certain advantages of the approach have been recognized (Galbraith et al, "Imaging Cytometry by Multiparameter Fluorescence", Cytometry 12, 579-596 (1991), the contents of which are incorporated herein by reference): the power of image analysis can be brought to bear to facilitate the analysis of multi-color fluorescence images. A major drawback is the very low "throughput" of this approach, given the low density of cells which must be maintained to avoid the risk of forming three-dimensional clusters in which the view of some of the cells would be obstructed. A further drawback is that methods of cellular analysis invoking differential mobility of cells in response to external stimuli are not available.

Flow cytometry - Flow cytometry and its application to fluorescence-activated cell sorting rely on a serial mode of analysis that requires the capture of individual cells into droplets which are passed, one at a time, through a small observation window. During a short time, e.g., 1-10μs, all measurements on a given cell must be completed. For example, the simultaneous analysis of a small number of cell surface markers represents the state of the art in flow cytometry (Beavis and Pennline, "Detection of Cell-Surface Antigens Using Antibody-conjugated Fluorospheres", BioTechniques 21, 498-503 (1996), the contents of which are incorporated herein by reference). Flow cytometric analysis represents just the last step in an assay: the instrumentation does not lend itself to the integration of liquid handling and automation of bioanalytical assays. This is in significant contrast to the present invention which integrates the optical manipulation of bead arrays and cellular assembly technology with microfluidics and optical imaging into a miniaturized platform for quantitative molecular and cellular analysis.

Recent developments of smaller and more specialized instruments notwithstanding, the intrinsic complexity of flow cytometry requires specialized and expensive instrumentation including multiple laser sources of high power, matching detectors and fast electronics to handle the acquisition and processing of signals during the short available interrogation interval during which each cell traverses the detection window. Sorting requires additional steps and equipment. Even after many years of commercial development, the operation of flow cytometers requires trained personnel to handle calibration and ensure reliable performance. The cost of commercial instruments,

for example, those offered by Becton-Dickinson, Beckman Coulter or Cytomation is prohibitive for all but major medical centers or research institutions.

Sorting and separation of cells by affinity and/or flow - Several realizations have been described for the fractionation of cells by selective affinity for, and hence differential retention by a nearby solid phase or by differential mobility in superimposed force fields such as a fluid flow with superimposed transverse electric field (field flow fractionation or FFF) or by combinations thereof. Differential dielectrophoretic levitation of cells above a set of interdigitated electrodes and fractionation in a transverse flow represents a recent addition to this set of techniques which offers the potential for differential blood cell analysis. As with other FFF techniques, the latter method relies on downstream detection of individual cells and thus resembles flow cytometry with respect to the implied serial processing to establish the identity of sub-populations. A further disadvantage is the requirement for an elongated channel to permit the spatial separation between sub-populations.

Cytometry According to the Present Invention

Cytometry according to the present invention subsumes the entire set of capabilities for cellular analysis illustrated above into a common platform which is miniaturized and integrated with microfluidics and optical imaging and compatible with quantitative bear array format of molecular interaction analysis whose capabilities may be used in "multiplexed" fluorescence immunoassays and DNA hybridization assays.

The specific advantages of the cytometry according to the present invention include:

- high throughput resulting from the high densities of the cellular assemblies (for example, 10^4 - 10^6/mm^2) and the parallel mode of signal acquisition afforded by imaging;

- integration of sample preparation, fluid handling and assay readout, for example, integration with on-chip assay formats;

- parallel processing based on imaging of multiple colors and morphological parameters;

- light-directed transport and assembly of cells;

- light-guided separation and fractionation of cells into sub-populations based on optically programmable illumination patterns and frequency dispersion in the

dielectrophoretic response; and

- significant reduction in expense of instrumentation.

The method and apparatus of the present invention enables the parallel processing of large numbers of cells assembled into dense planar assemblies; typically, 10^4 cells of 10μm diameter will be accommodated in an area as small as 1mm^2. Smaller bacterial cells such as E.coli can be accommodated in even higher density. In significant contrast to flow cytometry, signals are collected simultaneously from a large number of cells by way of imaging, and the rate of acquiring signal is very high even when cells are held stationary.

The present invention presents a very significant advantage, especially when multiple colors are to be discriminated and are to be quantitatively determined by analysis of multi-color images of individual cell surfaces. Thus, the cytometry according to the present invention will readily accommodate multiple cell surface markers and additional morphological indicators for individual cells (size, shape) or mutually interacting clusters of cells. This is generally possible even without specialized equipment. In fluorescence immunoassays using the READ bead array format, the capability has been demonstrated using an inexpensive consumer-grade video CCD camera, of detecting signals from individual beads displaying only 1000 labeled antibody molecules. Higher levels of sensitivity are possible for detection of markers on cell surfaces because fluorescent beads may be used as the label to decorate the cells.

A. Cells form arrays in response to AC electric fields

Referring now to Figure 12, therein is illustrated *E.coli* cells (grown to log phase and suspended in 280 mM mannitol solution at the density of 10^8/ml) which form arrays under electric field. On the left side of the field, the cells are assembled at very high density. Some singlet cells at the right side are moving to the left to join the array. (Bright field image. 25x obj, NA 0.45. Electric field: 5V, 200 Hz).

Referring now to Figure 13, therein is illustrated the same effect as in Figure 12, with *S. cerevisiae* cells (baker's yeast). (Bright field image. 25x obj, NA 0.45; electric field: 5V, 400 Hz).

B. Cells assemble in accordance with projected illumination patterns.

Referring now to Figure 14, therein is illustrated *E.coli* forming diamond shapes that are projected to the surface by an LCD panel controlled by appropriate graphics software; (20x objective, NA 0.28; Electric field: 10V, 7 kHz).

Referring now to Figure 15, therein is illustrated *S. cerevisiae* responding to two circular illumination patterns;(20x objective, NA 0.28; Electric field: 15 V, 15 kHz).

C. Light-directed cell transport and reconfiguration of cellular assemblies

By enabling the controlled transport of cells and cellular assemblies in accordance with the PARSE format, the cytometry according to the present invention provides an entire additional level of cellular analysis which has the inherent capability to fractionate cell populations "in-situ".

The principal approach according to the present invention is to generate gradients in electrochemial properties of the interface to which sub-populations of cells exhibit a differential response. Such gradients may be produced by invoking patterns of illumination in geometry, as disclosed herein. Examples in the figures illustrate that cells respond to these patterns in a variety of ways.

For example, cells may be collected into illuminated regions (Fig.16; *S. cerevisiae* responding to the illumination. Arrow indicate the direction cells are moving in. (20x objective, NA 0.28; Electric field: 15 V, 15 kHz)) or expelled from illuminated regions (Fig.17; *E. Coli* cells being repelled from the illumination circle.(20x objective, NA 0.28; Electric field: 10 V, 7 kHz; low substrate doping level (1 Ohm cm-10 Ohm cm)); "channeled" by reconfiguring illuminated regions or "dragged" and "dropped" by displacing illuminated regions (Fig. 18; *E.coli* cells following the expansion of the rectangular illuminated shape (20x objective, NA 0.28; Electric field: 15 V, 8 kHz)).

Cells also may be "trapped" in well defined regions in response to light-induced spatial gradients in the electrical properties of the EIS structure (Fig.19; *E. coli* cells being trapped by the light stripe at the left and being prevented from moving to the circular region on the right. (20x objective, NA 0.28; Electric field: 15 V, 8 kHz)).

This set of fundamental operations permits the creation of spatially varying electrochemical potentials and corresponding lateral, highly inhomogeneous electric fields. In addition, the cytometry according to the present invention permits the

programmable lateral displacement and reconfiguration of illumination patterns so that the location and the strength of the spatial inhomogeneities in electric fields can be adjusted at will.

In addition, the cytometry according to the present invention also provides frequency-dependent forces to which different constituents within mixtures of cells will generally exhibit differential responses. An example of segregation between different types of cells is shown in Fig.20 As shown in Figure 20, when an illuminated region containing both types of cells is suddenly expanded or "dragged", *E. coli* cells (left) respond faster than the larger *S. cerevisiae* cells (right), resulting in segregation. *(Bright field image. 25x obj, NA 0.45 Electric field: 5V, 400 Hz).*

Specifically, on the basis of the known dielectrophoretic properties of various cells, fractionation will be possible on the basis of characteristic frequency dispersion . For example, the characteristic transition of cells from positive to negative dielectrophoresis (as evidenced in electrorotation; see e.g.: Huang et al, op. cit.) will be the basis for the fractionation of cell mixtures in light-induced gradients in electrochemical EIS potential (and the resulting inhomogeneous fields): different cells will be expelled from confined regions on the substrate at characteristic frequencies and light intensities.

D. Non-viable cells behave differently than the viable cells.

The present inventors have observed that heat treated *S. cerevisiae* cells stop responding to the electric field and light after the first ten to twenty seconds and adhere to the substrate.

Example: Immunophenotyping - In a manner analogous to READ, a cellular array format will provide a sensitive and highly parallel mode of monitoring the presence of characteristic cell surface markers as illustrated in Figure 21. For example, groups of lymphocyte cells can be put into array format on the substrate. Antibodies for the cell surface markers (CD3, CD4, CD8, CD56, etc.) are coupled to small microspheres that are differentially coded. The mixture of beads with different antibodies are directed to the cells arrays using the optical arrangement principles of the present invention. Antibodies on the bead surface will bind to the cell surface markers, thus, the correspondent beads will link to the cells. The number of surface markers which can be detected simultaneously depends on chemical coding capacity of the bead. Hundreds of

chemical codings may possibly be used. Signal integration capability according to the present invention enables more accurate and quantitative analysis of the cells.

Example: Multiplexed Affinity Detection - A novel format enabling the simultaneous affinity identification of multiple types of cells is shown in Fig. 22. A random encoded bead array of antibodies will provide a novel, highly diverse set of anchoring points for lymphocytes displaying cognate cell surface antigens. Under conditions of controlled stringency, flow-induced shear will remove cells that are not specifically bound while leaving "positives" attached to individual beads. Focusing on this downstream step takes advantage of the intrinsic "concentration" of sample on the cell surface and the ability to capture, with a single antibody probe, multiple pathogens sharing common processed antigenic determinants.

Example: Light-guided Sorting and Isolation - Combination of AC electric fields with light-controlled modulation of surface-electrochemical properties is particularly powerful. The illuminated areas mimics the behavior of patterned electrodes and can be repositioned in real time to isolate designated cells. Cells can be sorted and isolated due to their natural properties or due to the property change caused by specific binding to probe labeled beads. The model is illustrated in Fig.23. Cells can be recovered after the isolation and further culturing or analysis may be performed. Specific lymphocyte cells and cancer cells may be of particular interest for such cell sorting and analysis.

Example: Cell-based Functional Assays - Cell-based functional assays may be performed in a highly efficient manner by assembling cellular arrays and imaging reaction of the cells as shown in Fig. 24. The reaction may be the direct change of extracellular markers, or the endpoints of some metabolic pathways. The endpoints may be intracellular fluorescence signal generated by enzymes such as β-lactamase whose substrate can penetrate the cells, thus, the signal may be directly quantitatively detected by imaging (Zlokarnik et al., "Quantitation of Transcription and Clonal Selection of Single Living Cells with β-lactamase as Reporter", Science 279, 84-88 (1998); see also: Bronstein et al, "Chemiluminescent and Bioluminescent Reporter Gene Assays", Anal. Biochemistry 219, 169-181 (1994); the contents of all of which are incorporated herein

by reference). If combined with microfludic control, such cell-based functional assays have tremendous applications in drug screening and ligand analysis. Cells may be placed at the same location as the chemical carrier beads. Under temporal and spacial controlled release of the compounds or ligands, response of the cells may be recorded. After image analysis, compounds or ligands that generate the desired cell response may be easily identified. Because the cells may be directly imaged, cell lysis and the following assays will generally not be necessary as part of the traditional cell-based analysis. Furthermore, the controlled assembly of the cells and chemical carrying beads, according to the present invention, may eliminate the use of microtiter plates in the cell-based screening.

Example: Integrated Format Profiling of Secreted Cellular Markers - Control of cell assembly and microfluidic pathway enables direct in situ screening of secreted cellular markers. This model is illustrated in Fig. 25. Cellular arrays are assembled in one compartment of a two-compartment liquid reactor and are exposed to candidate drugs; the response is measured by monitoring a panel of secreted markers such as cytokines. Cell secretion are collected and introduced into a second compartment where they are profiled based on the response of a pre-deposited bead array using the READ format. Such a two-compartment reactor may be operated in parallel for multiple drug screening or kinetic analysis.

Example: Cell-to-Cell Interaction - Cell-to-cell in vivo interaction was seldom studied directly due to the lack of control in cell manipulation. The present invention provides a powerful tool in this area. As illustrated in Fig. 26, cells may be temporally and spatially controlled to favor the binding between the cells. Alternatively, cells may be directed to preassembled cell arrays for cell binding analysis. It may also be possible to study the interaction between cells such as T cells, B cells and polymorphonulear leukocytes, in the activation of immune responses.

Example: Chemotaxis Analysis - Cells can be positioned in a controlled configuration to test their chemotaxis properties. This model is illustrated in Fig. 27. For example, *D. discoideum* cells may be tested under cAMP or specific polypeptide gradient. (Parent &

Devreotes, "A cell's sense of direction", Science 284, 765-770 (1999); the contents of which are incorporated herein by reference). Response of the receptors, G proteins, actin or actin-binding proteins may be detected through fluorescence signal generated by their correspondent fusion protein with GFP (Green Fluorescent Protein). Movement of the cells may also be recorded. More importantly, the force of the movement may be measured by comparing to the energy used in the light control of the cells.

The present invention may be used for cell identification and detection, for example, in immunophenotyping and affinity cell detection. In this application, affinity probes are coupled to beads on a solid surface. The beads are chemically and spatially coded. The cells are transported to the desired position using light control. Cells with the surface markers will bind at the correspondent positions. The advantages of this approach include a simpler and less expensive system, a higher degree of throughput for parallel immunophenotyping, the manipulation of a large number of cells, thus the detection of rare events, and better integration between the reaction and the detection steps, thus allowing the entire process to be performed on-chip.

The present invention may also be used in cell sorting, fractionation and isolation. In this case, cells can be separated according to their intrinsic properties, such as their dielectrophoretic mobility, viability, size, etc. Cells can also be sorted by their association with antibody coupled beads. The advantages of this approach include multiple cell sorting, simpler substrate fabrication, and flexible and programmable control.

The present invention may also be used for screening, for example, cell mediated drug and ligand screening, or cell secretion mediated screening. In this approach, cells may be transported and assembled a the desired position. Chemical compounds or ligands are released at the corresponding position. The response of the cells may then be observed through their reporter genes. The advantages of this approach include higher throughput, a simpler system without the complication of automated pipetting machines, and the possible combination with microscopy for simpler and more direct assays.

The present invention may also be used for cell to cell interaction, for example, in suspension or on surface interaction. In this approach, cells may link with each other through their surface interaction. Movement of the cells may then be guided

by light. The advantages of this approach include the ability to study the interaction between cells.

While the invention has been particularly shown and described with reference to a preferred embodiment thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention.

WHAT IS CLAIMED IS:

1. A method of manipulating cells suspended at an interface between an electrode and an electrolyte solution, the method comprising the following steps:

providing a plurality of cells suspended at an interface between an electrode and an electrolyte solution, the electrode being a light-sensitive electrode;

generating an electric field at the interface; and

illuminating the interface with a predetermined light pattern to control the movement of the cells in accordance with the predetermined light pattern and the properties of the electrode.

2. The method of claim 1, in which the movement of the cells is controlled by the combination of illumination and the electric field.

3. The method of claim 1, wherein the electric field is generated by application of an AC voltage.

4. The method of claim 1, further comprising the step of varying the configuration or the intensity of the light pattern.

5. The method of claim 1, wherein the properties of the electrode comprise the properties affecting the local distribution of the interfacial electric field.

6. The method of claim 5, wherein the properties of the electrode comprise impedance or surface charge depositing.

7. The method of claim 1, wherein the properties of the electrode comprise impedance, the modification in the impedance causing the cells to move to the area of low impedance in response to the electric field.

8. The method of claim 1, wherein the electrode comprises a silicon electrode which is coated with a dielectric layer.

9. The method of claim 1, which provides an additional electrode, the additional electrode and the light-sensitive electrode being substantially planar and aligned parallel to one another and separated by a gap, with the electrolyte solution containing the cells being located in the gap.

10. The method of claim 9, wherein the additional electrode is optically transparent electrode.

11. The method of claim 1, in which the electrode is patterned by spatially modulated oxide growth, surface chemical patterning or surface profiling, wherein said patterning modifies the spatial distribution of the interfacial electric field.

12. The method of claim 11, wherein the patterning step is used to create a plurality of areas of low impedance on the electrode, and the illumination step is used to selectively connect one or more of the areas of low impedance, thereby causing the cells to move therebetween in accordance with the patterning and the predetermined light pattern.

13. The method of claim 1, wherein the illumination step comprises illuminating a selected area of the electrode, which in combination with the electric field, causes the cells to move into the selected area or out of the selected area.

14. The method of claim 1, wherein the illuminating step comprises illuminating the electrode with a predetermined light pattern to form a lateral gradient in the electrochemical properties of the electrode to control the movement of the cells in accordance with the illumination pattern in a direction orthogonal to the direction of the applied electric field.

15. The method of claim 1, wherein the cells comprise eukaryotic or prokaryotic cells.

16. A method of manipulating cells suspended at an interface between an electrode and

an electrolyte solution, the method comprising the following steps:

providing a first electrode positioned in a first plane, and a second electrode positioned in a second plane different from the first plane, an electrolyte solution located therebetween and a plurality of cells suspended at an interface between the electrolyte solution and the second electrode, wherein the second electrode comprises a planar electrode having a surface and an interior, the surface or interior having been modified to produce spatial modulations affecting the local distribution of the electric field at the surface of the electrode; and

generating an electric field between the first and the second electrode by applying a voltage between the two electrodes to control the movement of the cells in accordance with the spatial modulations of the properties of the second electrode.

17. The method of claim 16, wherein the properties of the electrode comprise impedance or surface charge depositing.

18. The method of claim 16, wherein the spatial modulation of the properties of the second electrode is carried out by modifying the surface or the interior of the second electrode by spatially modulated oxide growth, surface chemical patterning or surface profiling.

19. The method of claim 16, wherein the second electrode comprises a silicon electrode coated with a dielectric layer.

20. The method of claim 16, wherein the first electrode comprises a planar electrode, with the first electrode and the second electrode being substantially planar and parallel to one another and separated by a gap, with the electrolyte solution containing the cells being located in the gap.

21. The method of claim 21, wherein the first electrode comprises an optically transparent electrode.

22. The method of claim 16, wherein the properties of the second electrode being

modulated comprise impedance, with one or more areas of the surface or the interior of the second electrode being modified to exhibit low impedance, and wherein the movement of the cells is to areas of low impedance.

23. The method of claim 16, wherein the electric field is generated by applying an AC voltage between the first and the second electrode.

24. An apparatus for manipulating cells suspended at an interface between an electrode and an electrolyte solution, comprising:

a light-sensitive electrode and an electrolyte solution;

a means for generating an electric field at an interface between the electrode and the electrolyte solution; and

a means for illuminating the electrode.

25. An apparatus for manipulating cells suspended at an interface between an electrode and an electrolyte solution, comprising

a first electrode positioned in a first plane, and a second electrode positioned in a second plane different from the first plane, the first and the second electrode each comprising a planar electrode and said electrodes being in a substantially parallel alignment,

a gap between the first and the second electrode, the gap being capable of containing an electrolyte solution in which a plurality of cells is suspended,

wherein the second electrode comprises a light-sensitive electrode,

and wherein the first and the second electrode are configured so that when a voltage is applied between the electrodes, with the electrolyte solution containing the cells located in the gap, an electric field is generated at an interface between the second electrode and the electrolyte solution,

the apparatus being capable of controlling movement of the cells at the interface between the second electrode and the electrolyte solution when the electric field is generated at the interface and the interface is illuminated with a predetermined light pattern.

26. The apparatus of claim 24 or 25, wherein the light-sensitive electrode comprises a silicon electrode which is coated with a dielectric layer.

27. The apparatus of claim 24 or 25, wherein the light-scnsitive electrode is patterned by spatially modulated oxide growth, surface chemical patterning or surface profiling, wherein said patterning modifies the spatial distribution of the electric field at the interface.

28 The apparatus of claim 24 or 25, wherein the properties of the light-sensitive electrode comprise impedance.

29. The apparatus of claim 25, further comprising an electric field generator which generates an electric field the interface, and an illumination source which illuminates the interface.

30. The apparatus of claim 24 or 25, further comprising an electrolyte solution and a plurality of cells located at the interface.

31. An apparatus for manipulating the movement of cells suspended at an interface between an electrodc and an electrolyte solution, comprising:

a first electrode positioned in a first plane, and a second electrode positioned in a second plane different from the first plane, the first and the second electrode being substantially planar electrodes in parallel alignment;

a gap between the first and the second electrode, the gap being capable of containing an electrolyte solution in which a plurality of cells are suspended,

wherein the second electrode is patterned to modify its properties and wherein the first and the second electrode are configured so that an electric field is generated when a voltage is applied between the first and the second electrode, with the electrolyte solution containing the cells located in the gap, and wherein the patterning of the second electrode modifies its properties affecting the local distribution of the interfacial

electric field.

32. The apparatus of claim 31, wherein the second electrode is modified by spatially modulated oxide growth, surface chemical patterning or surface profiling.

33. The apparatus of claim 31, wherein the properties comprise impedance.

34. The apparatus of claim 31, wherein the second electrode comprises a silicon electrode which is coated with a dielectric layer.

35. A method of forming an assembly of cells into a designated area on a substrate, comprising:

providing a plurality of cells suspended at an interface between a light-sensitive electrode and an electrolyte solution;

generating an electric field at the interface; and

illuminating the interface with a predetermined light pattern to control the formation of a planar assembly of substantially one layer of cells in a designated area on the electrode, wherein the designated area is defined by the pattern of illumination.

36. The method of claim 35, in which the movement of the cells is controlled by the combination of the illumination and the electric field.

37. The method of claim 35, which provides an additional electrode, such that the additional electrode and the light-sensitive electrode being substantially planar and aligned to one another and separated by a gap, wherein the additional electrode comprises an optically transparent electrode.

38. The method of claim 35, further comprising modifying the electrode by spatially modulated oxide growth, surface chemical patterning or surface profiling.

39. The method of claim 35, further comprising the step of spatially or temporally varying the light pattern to cause the alteration of the assembly, said alteration

being selected from the group consisting of disassembly, disassembly followed by reassembly, repositioning of the assembly, reconfiguration of the assembly, and segmentation of the assembly.

40. The method of claim 35, further comprising the step of varying the frequency or the voltage of the electric field to cause the alteration of the assembly, said alteration being selected from the group consisting of disassembly, disassembly followed by reassembly, repositioning of the assembly, reconfiguration of the assembly, and segmentation of the assembly.

41. The method of claim 35, in which the plurality of cells comprises more than one type of cells, the method further comprising the step of fractionating one type of cells from another to induce a displacement of the plurality of cells within the assembly, fractionation arising as a result of differences in mobility of said types of cells.

42. The method of claim 35, further comprising the step of maintaining the cells in the assembly, wherein the maintenance step comprises maintaining the electric field and the predetermined light pattern, or immobilizing the cells by chemical or physical means.

43. The method of claim 42, in which the cellular assembly is immobilized on the electrode by chemically linking the cells or confining the cells.

44. A method of forming an assembly of cells in a designated area on a substrate, comprising:

providing a first electrode positioned in a first plane, and a second electrode positioned in a second plane different from the first plane, an electrolyte solution located therebetween and a plurality of cells suspended at an interface between the electrolyte solution and the second electrode, wherein the second electrode comprises a planar electrode having a surface and an interior, the surface or interior having been modified to produce spatial modulations in properties of the second electrode; and

generating an electric field between the first and the second electrode by applying a voltage between the two electrodes, thereby forming a planar assembly of substantially one layer of cells in a designated area on the second electrode, wherein the designated area is defined by the spatial modulation affecting the local distribution of the electric field at the surface of the second electrode and the properties of the second electrode are those affecting the local distribution of the interfacial electric field.

45. The method of claim 44, wherein the properties being modulated comprise impedance or surface charge density.

46. The method of claim 44, wherein the properties being modulated comprise impedance, and the designated area comprising the area of lower impedance.

47. The method of claim 44, further comprising the step of maintaining the assembly of cells by maintaining the electric field or by immobilizing the cells by chemical or physical means.

48. The method of claim 44, in which the cellular assembly is immobilized on the electrode by chemically linking the cells or confining the cells.

49. The method of claim 44, in which the first electrode comprises a substantially planar, optically-transparent electrode.

50. The method of claim 49, further comprising the step of optically imaging the cells in the assembly.

51. The method of claim 50, in which the optical imaging step is performed to determine the morphology of the cells in the assembly.

52. The method of claim 50, in which the optical imaging step is performed to determine the size, shape or granularity of the cells in the assembly.

53. An assembly of cells formed in a designated area on a substrate, comprising
a substrate; and
a dynamically formed planar assembly of cells comprising substantially one layer of cells in a designated area on the substrate.

54. The assembly of claim 1, in which the designated area is defined by the electrochemical properties of the substrate.

55. An assembly of cells formed in a designated area on a substrate, comprising
a substrate; and
a planar assembly of cells comprising substantially one layer of cells in a designated area on the substrate, wherein the assembly is formed according to the method of claim 33.

56. An assembly of cells formed in a designated area on a substrate, comprising
a substrate; and
a planar assembly of cells comprising substantially one layer of cells in a designated area on the substrate, wherein the assembly is formed according to the method of claim 35 or 44.

57. The assembly of cells according to any one of claims 53 to 56, wherein the substrate comprises a silicon chip which is coated with a dielectric layer.

58. The assembly of any one of claims 53 to 56, wherein the assembly of cells, subsequent to being formed, is immobilized by chemical or physical means.

59. The assembly of claim 58, in which the cellular assembly is immobilized on the electrode by chemically linking the cells or confining the cells.

60. A method of conducting a bioassay involving an assembly of cells, comprising:
providing a dynamically formed planar assembly of cells comprising substantially one layer of cells in a designated area on the substrate,

contacting the cells with an analyte; and

detecting the binding of the analyte to the cells.

61. A method of conducting a bioassay involving an assembly of cells, comprising:

providing a planar assembly of cells comprising substantially one layer of cells in a designated area on the substrate, wherein the assembly of the cells is formed according to the method of claim 35 or 44;

contacting the cells with an analyte; and

detecting the binding of the analyte to the cells.

62. The method of claim 60 or 61, in which the cells in the assembly of the cells are immobilized prior to contacting the cells with the analyte.

63. The method of claim 62, in which the cells are immobilized by chemically linking the cells or confining the cells.

64. The method of claim 60 or 61, in which the analyte is directed to a specific cellular marker, and the bioassay is for determining the presence of the cellular marker on the surface of the cells.

65. The method of claim 64, in which the bioassay is directed to cell typing, with the presence of the marker on the cell surface indicating the cell type.

66. The method of claim 64, wherein the analyte is attached to a label and the detection of the binding of the analyte to the cells is carried out by detecting the presence of the label.

67. The method of claim 64, further comprising the step of removing the analyte that is not bound to the cell before the detection step is carried out.

68. The method of claim 64, wherein the analyte is attached to a label, said label comprising a flourescent tag.

69. The method of claim 64, wherein the analyte is attached to a label, said label comprising a bead which is distinguishable by chemical or physical characteristics.

70. The method of claim 64, in which more than one analyte is tested simultaneously for binding with the cells, with each analyte being attached to an encoded bead that is distinguishable by chemical or physical characteristics, and the detecting step comprising decoding the beads bound to the cells to determine the respective identities of analytes bound to the cells.

71. The method of claim 64, wherein the analyte is a ligand directed to a specific cellular receptor and the bioassay is for determining the presence of the receptor on the surface of the cells.

72. The method of claim 64, wherein the analyte is an antibody directed to a specific cellular antigen, and the bioassay is for determining the presence of the antigen on the surface of the cells.

73. The method of claim 64, in which the presence of more than one antigen is determined using more than one antibody, wherein each antibody is attached to a fluorophore tag that is chemically distinguishable, wherein the detection step comprises multicolor imaging of the cells.

74. A method of assaying the binding of cells with a ligand or an antibody, said method involving a planar assembly of cells and comprising:

providing an assembly of cells prepared according to claim 35 or 44, wherein the assembly further comprises a plurality of encoded beads randomly mixed with the cells in the designated are, each bead displaying on its surface a ligand or an antibody, the code of each bead uniquely corresponding to the ligand or the antibody on its surface, the proximity of the cells with the ligand or the antibody allowing the binding interaction therebetween,

disassembling the mixed assembly of the cells and the encoded beads; and

detecting the binding interaction by analyzing the formation of clusters composed of the cells and the encoded beads, the binding indicating the presence or absence of a cellular receptor or an antigen specific for the ligand or the antibody.

75. The method of claim 1, further comprising providing a planar array of encoded beads immobilized on the light-sensitive electrode, said array comprising substantially one layer of the encoded beads in a substantially non-random spatial arrangement and each bead displaying on its surface a ligand or an antibody, wherein the cells are moved across the array of beads, allowing the binding interaction to occur between ligand or the antibody and the cells, and detecting the binding by identifying the selective adhesion of the cells to beads in the array, with the binding indicating the presence or absence of a cellular receptor or an antigen specific for the ligand or the antibody.

76. The method of claim 73 or 74, in which more than one ligand or antibody is used, with each being attached to an encoded bead that is distinguishable by chemical or physical characteristics.

77. A method of detecting a cellular response to an analyte, the method involving an assembly of cells and comprising:

providing a dynamically formed planar assembly of cells comprising substantially one layer of cells in a designated area on the substrate;

contacting the cells with an analyte; and

detecting a cellular response to the analyte.

78. A method of determining a cellular response to an analyte, the method involving an assembly of cells and comprising: providing a planar assembly of cells formed according to claim 55 or 56; contacting the cells with an analyte; and detecting a cellular response to the analyte.

79. The method of claim 77 or 78, wherein one or more analytes are being tested for its ability to inducc the cells to secrete one or more biologically active substances, and

the detection step comprises detecting the presence of the biologically active substances.

80. The method of claim of 79, wherein the biologically active substance comprises cytokine.

81. A method of determining a cellular response to an analyte, the method comprising:

providing the apparatus according to claim 25 or 31;

providing a planar array of encoded beads on the first electrode, each bead displaying a ligand on its surface and uniquely identifying the ligand, said ligand being attached to the bead in a releasable manner;

providing a planar assembly of the cells on the second electrode, with the gap area separating the encoded beads from the assembled cells; and

releasing the ligand and monitoring a cellular response, the proximity of the cells to the bead array permitting determination of the identity of the ligand inducing the cellular response.

82. The method of claim 77 or 78 , wherein the analyte comprises a drug molecule or a ligand.

83. The method of claim 77 or 78 , in which the cellular response being detected is an expression of a particular gene, said expression being determined by detecting for the presence of an intracellular reporter gene product.

84. The method of claim 83, in which the expression of the intracellular reporter gene yields intracellular fluorescence.

85. The method of claim 77 or 78, in which the cellular response being detected is selected from the group consisting of: morphological change of the cells, change in the cell membrane permeability, and a change in the chemiotaxis response by monitoring the movement of cells.

86. A method of determining the interaction between cells, comprising:

providing an apparatus of claim 25 or 31, in which the electrolyte solution containing at least one type of cell is located in said gap;

introducing at least one cell of second type into the electrolyte solution and allowing the cell of the first type and the cell of the second type to interact;

generating an electric field at the interface to move the cells into a designated area on the second electrode; and detecting the interaction between the two cell types by ascertaining the formation of clusters.

87. A method of determining the interaction between cells, comprising:

providing an apparatus of claim 25 or 31, in which the electrolyte solution contains at least one cell of a first cell type;

generating an electric field at the interface to move the cell of the first type into a designated area on the second electrode;

introducing at least one cell of a second type into the electrolyte solution, and moving said cell to the designated area on the second electrode in which the first cell type is located, thereby allowing the interaction between the two cell types to occur; and

detecting the interaction.

88. A method of determining the interaction between cells, comprising:

providing an apparatus of claim 25 or 31, in which the electrolyte solution contains at least one cell of a first cell type and one cell of a second cell type, each cell type being located in a different designated area on the second electrode;

generating an electric field at the interface to move the cell of the first type into a designated area occupied by the cell of the second type, thereby allowing the interaction between the different cell types to occur; and

detecting the interaction.

89. The method according to any one of claims 86-88, in which the interaction between the cell types is detected by a method selected from the group consisting of: detecting a morphological change of cells, detecting a change in a cell surface marker by use of a ligand or an antibody; detecting a change in a protein

expression, detecting a change in the RNA level, detecting a change in cell permeability, and by detecting adhesion of cells.

90. The method of claim 86 or 88, in which the detection measure pertains to determining a change in the protein expression or RNA changes, wherein the cells are fixed after the changes to assist detection.

91. A method of conducting a bioassay involving an assembly of cells, comprising:

providing an apparatus of claim 25 or 31, in which the electrolyte solution containing a plurality of cells is located in the gap area;

introducing at least one analyte into the electrolyte solution, and allowing the binding interaction to occur between the analyte and the cells;

forming an assembly of the cells in a designated are on the light-sensitive electrode; and

detecting the binding between the analyte and the cells.

92. A method of detecting a cellular response to an analyte, the method involving an assembly of cells and comprising:

providing an apparatus of claim 25 or 31, in which the electrolyte solution containing a plurality of cells is located in the gap area;

introducing at least one analyte into the electrolyte solution and allowing the interaction to occur between the analyte and the cells;

forming an assembly of the cells in a designated are on the light-sensitive electrode; and

detecting a cellular response to the analyte.

93. The method of claim 91 or 92, wherein the interaction between the cells and the analyte occur before the formation of the assembly.

94. The method of claim 92, in which the analyte comprises an antibody or a ligand.

95. An apparatus for conducting a bioassay, comprising:

a first electrode positioned in a first plane, and a second electrode positioned in a second plane different from the first plane, the first and the second electrode each comprising a planar electrode and said electrodes being in a substantially parallel alignment,

a gap between the first and the second electrode, the gap being capable of containing an electrolyte solution in which a plurality of cells are suspended;

wherein the second electrode comprises a light-sensitive electrode,

and wherein the first and the second electrode are configured so that when a voltage is applied between the electrodes, with the electrolyte solution in the gap, an electric field is generated at an interface between the second electrode and the electrolyte solution,

a means for introducing one or more analytes or cells into the gap; and

a means for detecting the interaction between the analyte and the cells, or between the different types of cells,

the apparatus being capable of controlling movement of the cells at the interface between the second electrode and the electrolyte solution when the electric field is generated at the interface and the interface is illuminated with a predetermined light pattern.

96. An apparatus for manipulating the movement of cells suspended at an interface between an electrode and an electrolyte solution, comprising:

a first electrode positioned in a first plane, and a second electrode positioned in a second plane different from the first plane, the first and the second electrode each comprising a planar electrode and being in substantially parallel alignment;

a gap between the first and the second electrode, the gap being capable of containing an electrolyte solution in which a plurality of cells are suspended, wherein the second electrode is patterned to modify its properties and wherein the first and the second electrode are configured so that an electric field is generated when a voltage is applied between the first and the second electrode, with the electrolyte solution containing the cells located in the gap, and wherein the patterning of the second electrode modifies its properties affecting the local distribution of the interfacial electric field;

a means for introducing one or more analytes into the gap; and

a means for detecting the interaction between the analytess and the cells, or between different types of cells.



FIG. Ia

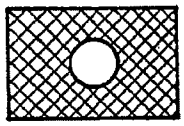


FIG. Ib

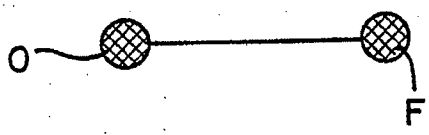


FIG. Ic

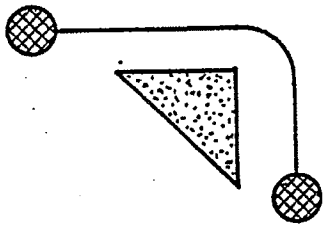


FIG. Id

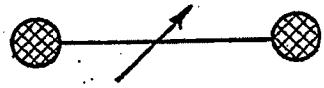


FIG. Ie

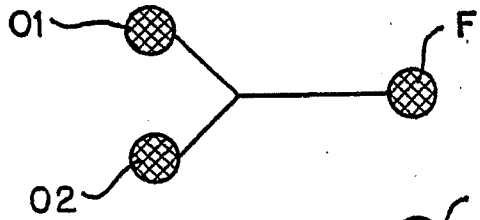


FIG. If

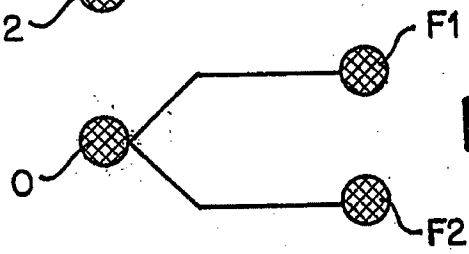


FIG. Ig

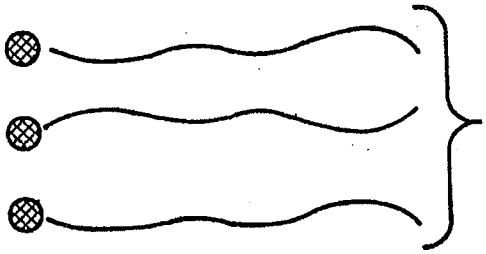


FIG. Ih

FIG. 2A

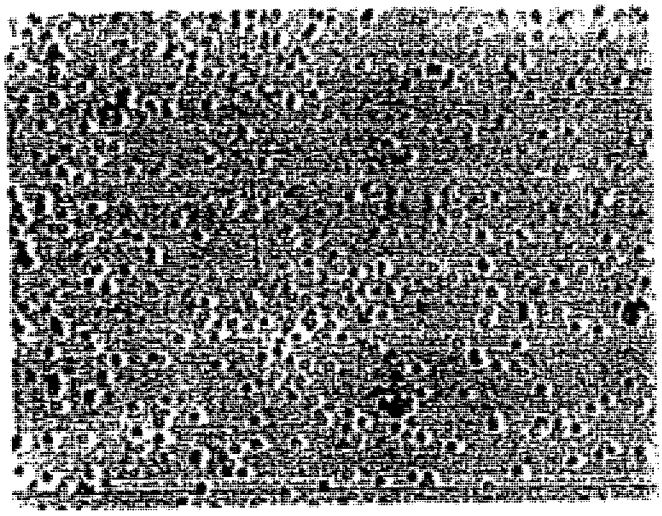


FIG. 2B

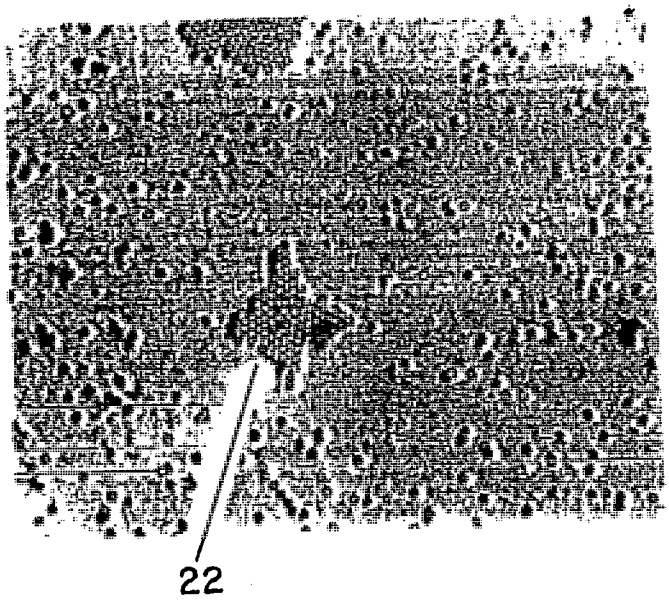


FIG. 2C

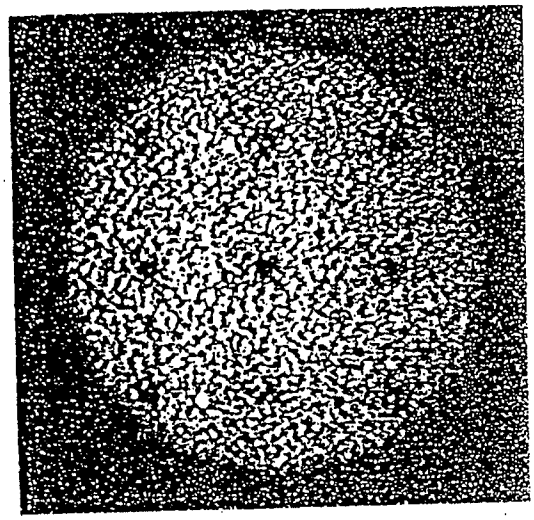


FIG. 2D

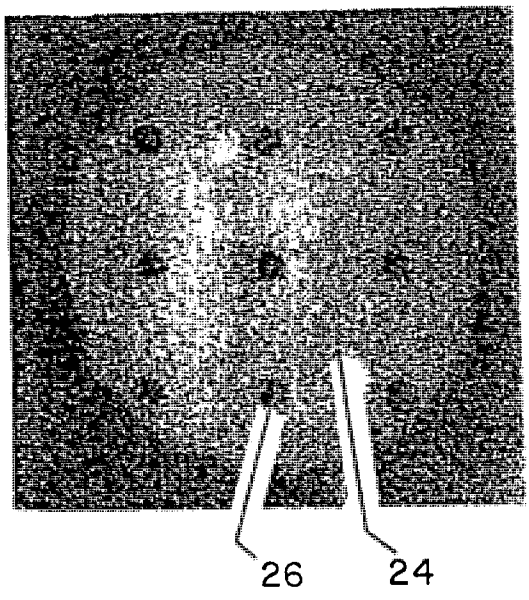


FIG. 3a

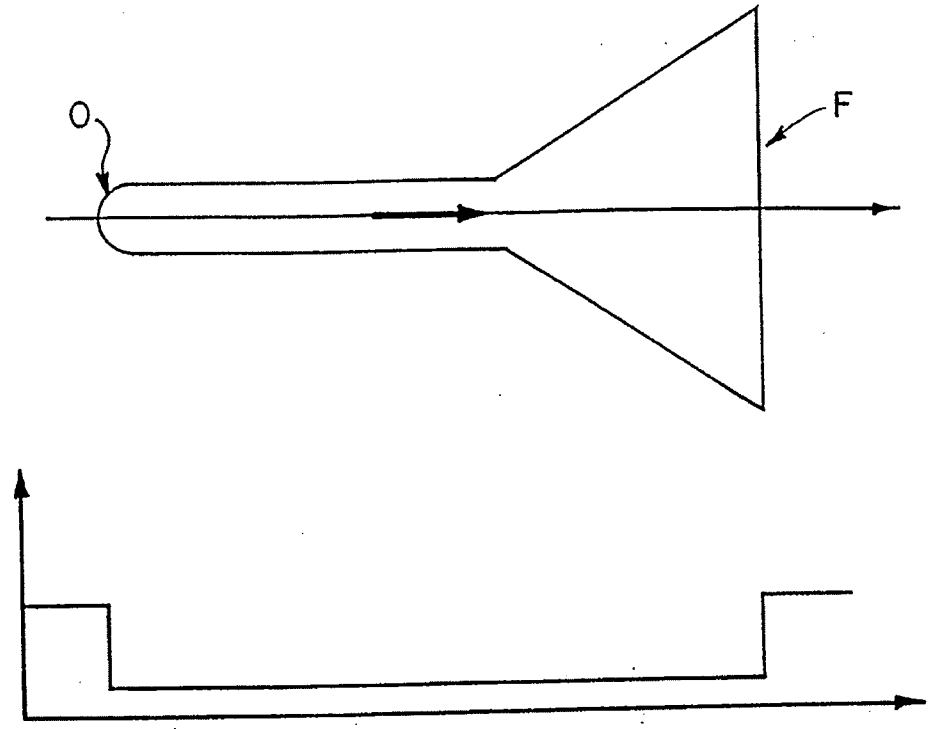


FIG. 3b

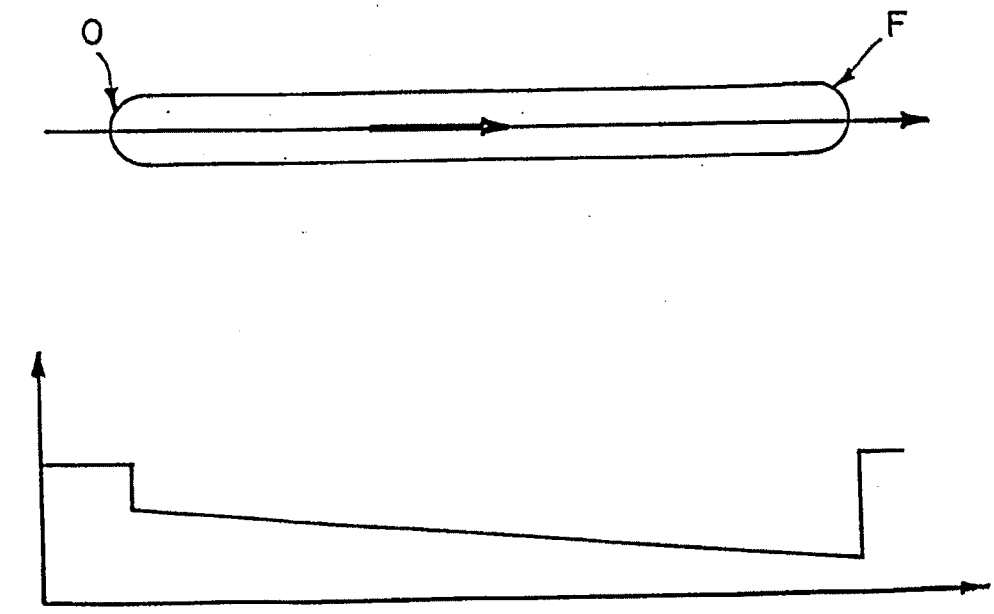


FIG. 3C

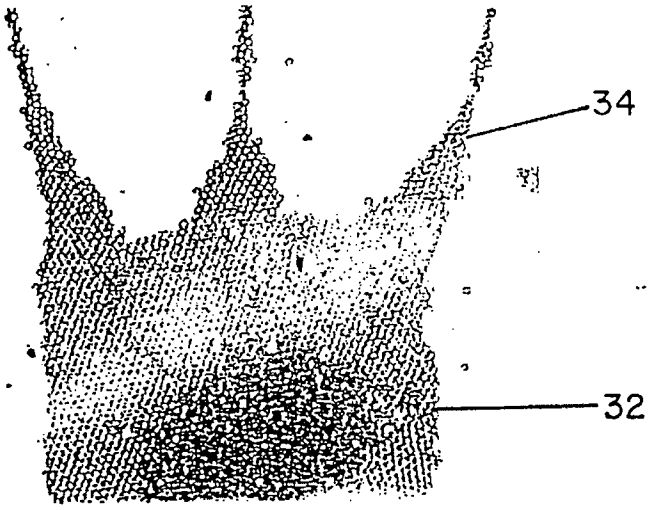


FIG. 3D

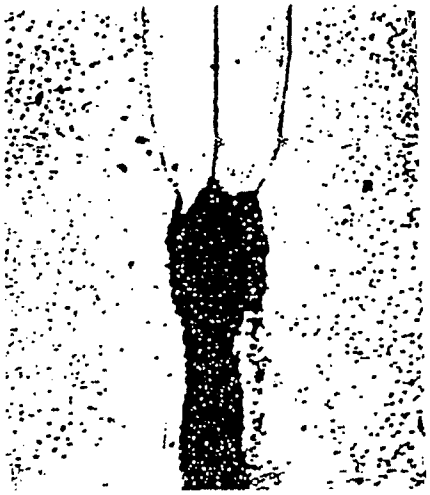


FIG. 4A

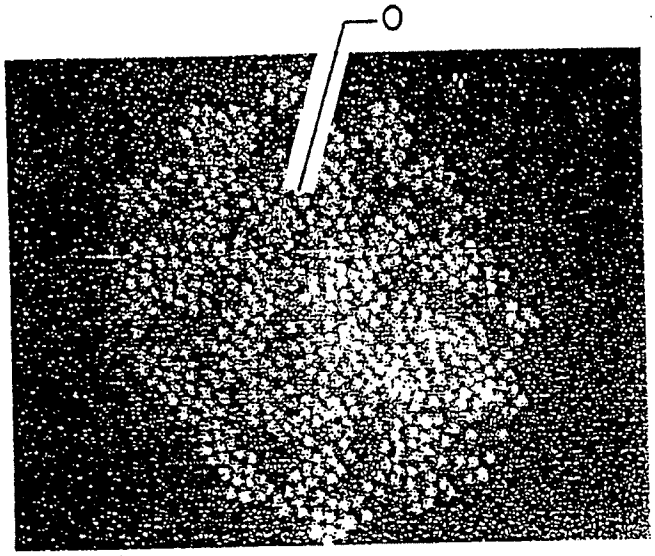


FIG. 4B

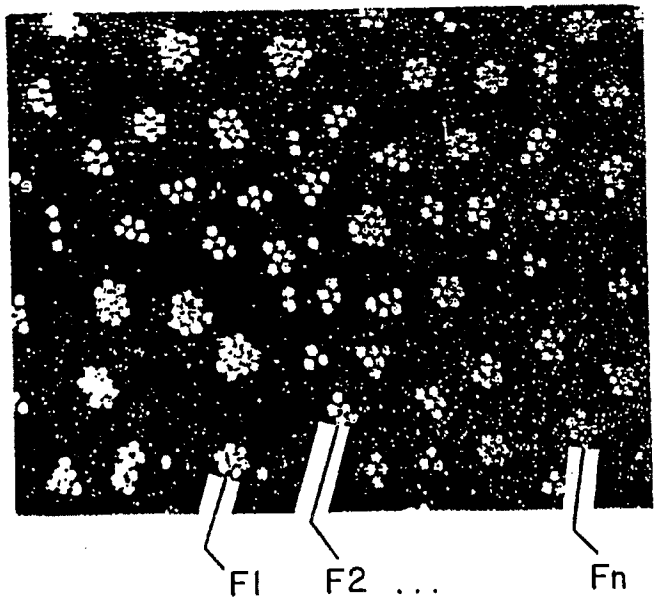
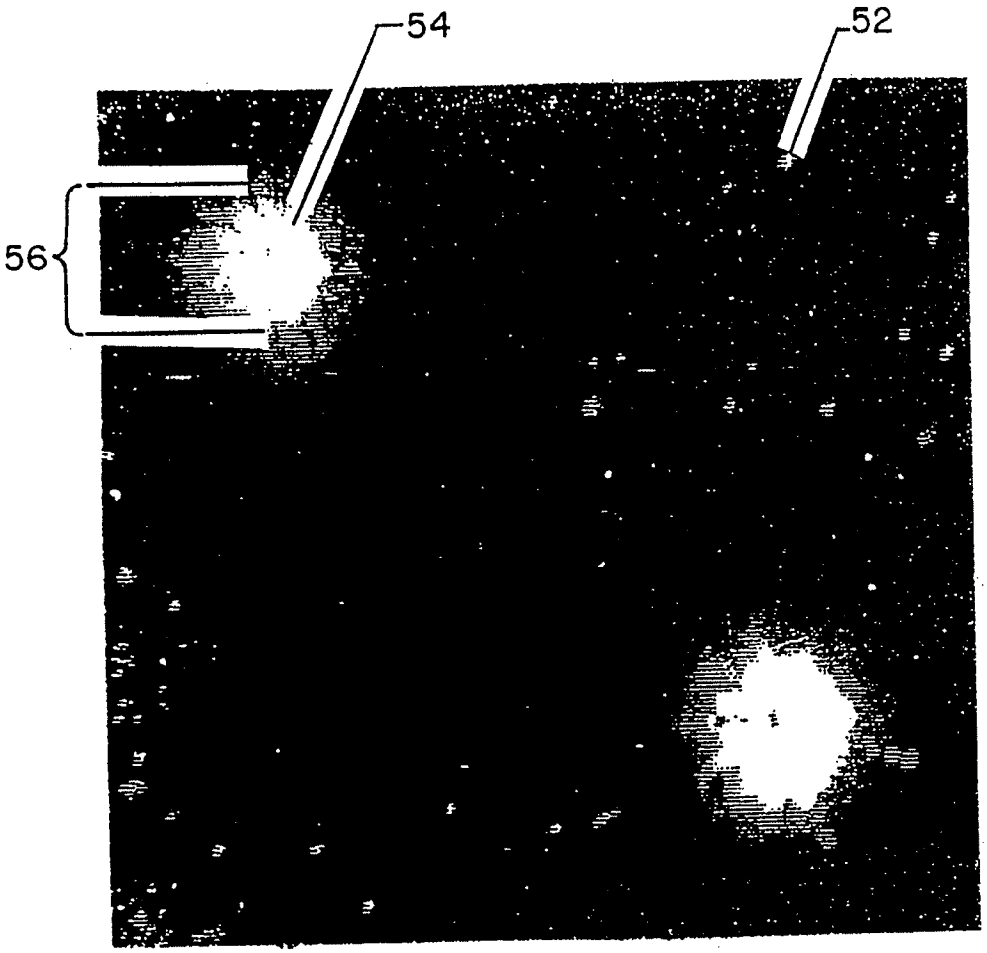


FIG. 5



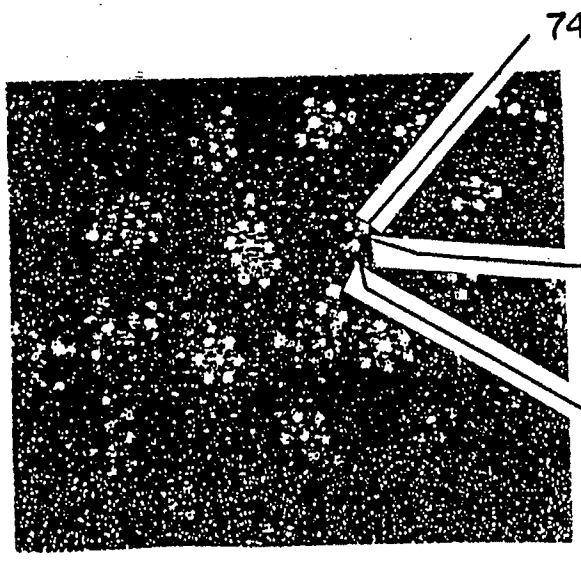


FIG. 7

FIG. 6a

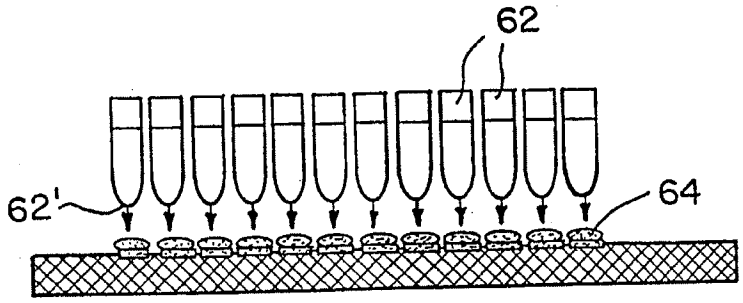


FIG. 6b

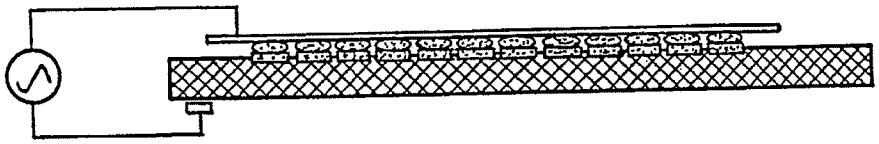


FIG. 6c

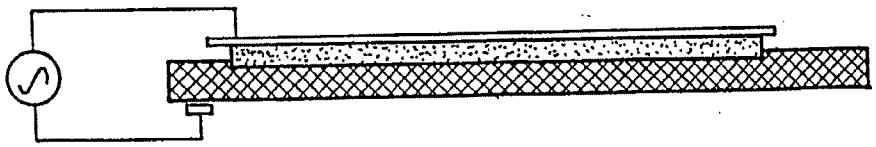


FIG. 8

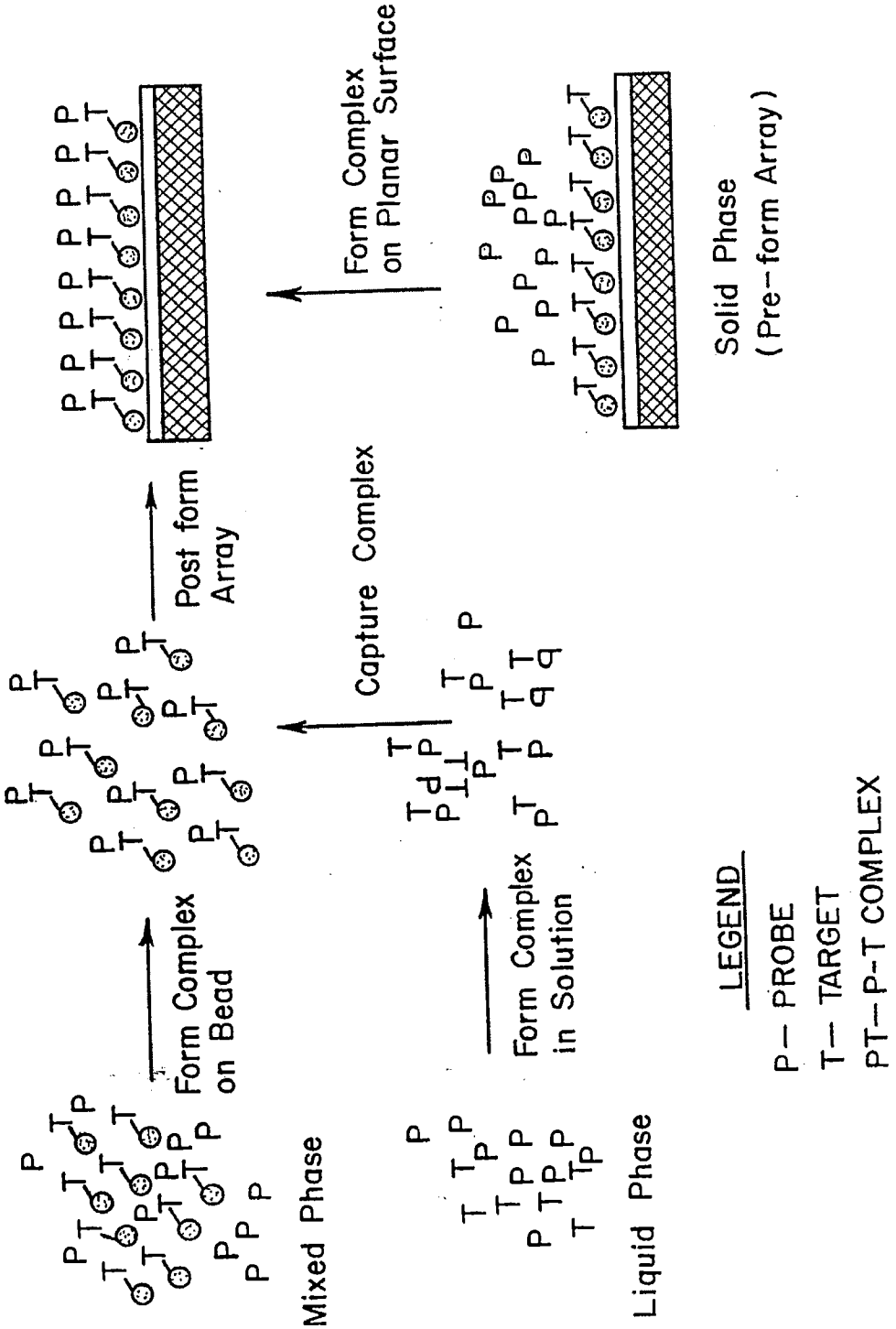

Mixed Phase
Form Complex on Bead
Liquid Phase
Form Complex in Solution
Capture Complex
Post form Array
Solid Phase
(Pre-form Array)
Form Complex on Planar Surface
LEGEND
P— PROBE
T— TARGET
PT— P-T COMPLEX

FIG. 9a

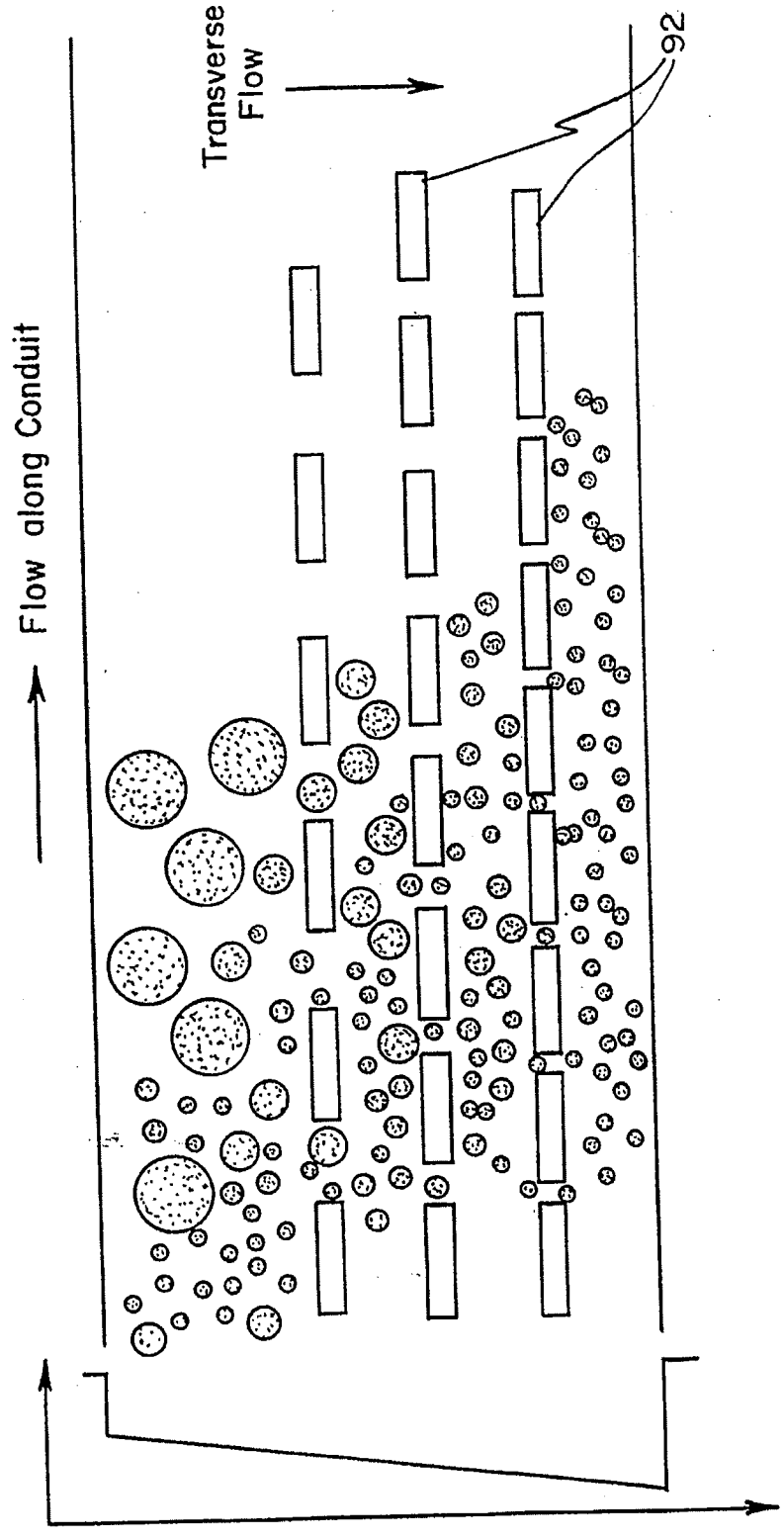


FIG. 9b

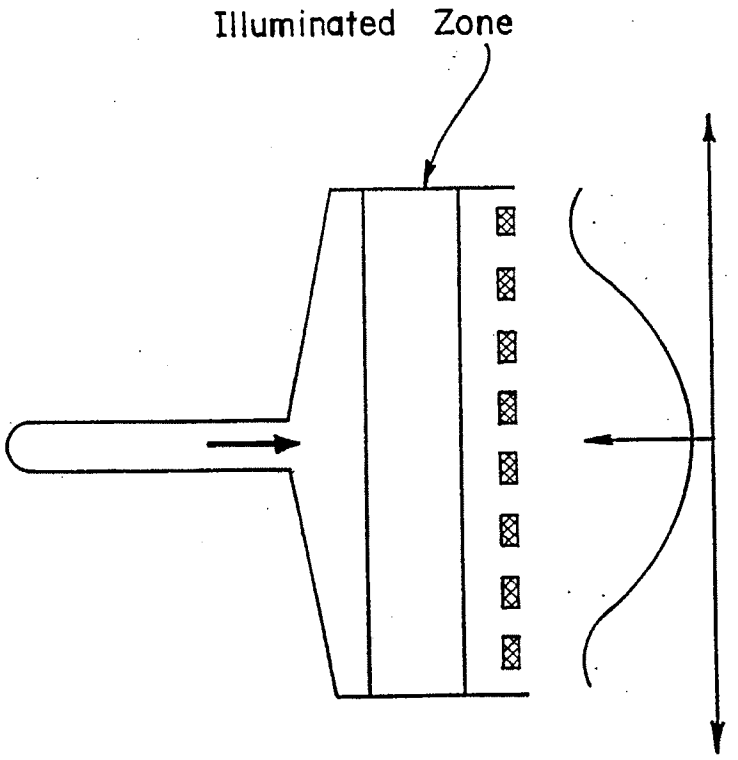


FIG. 9c

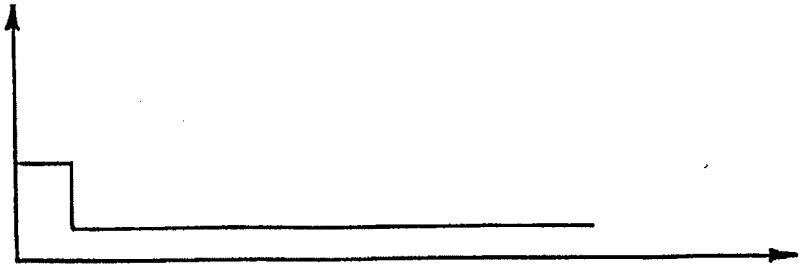


FIG. 10

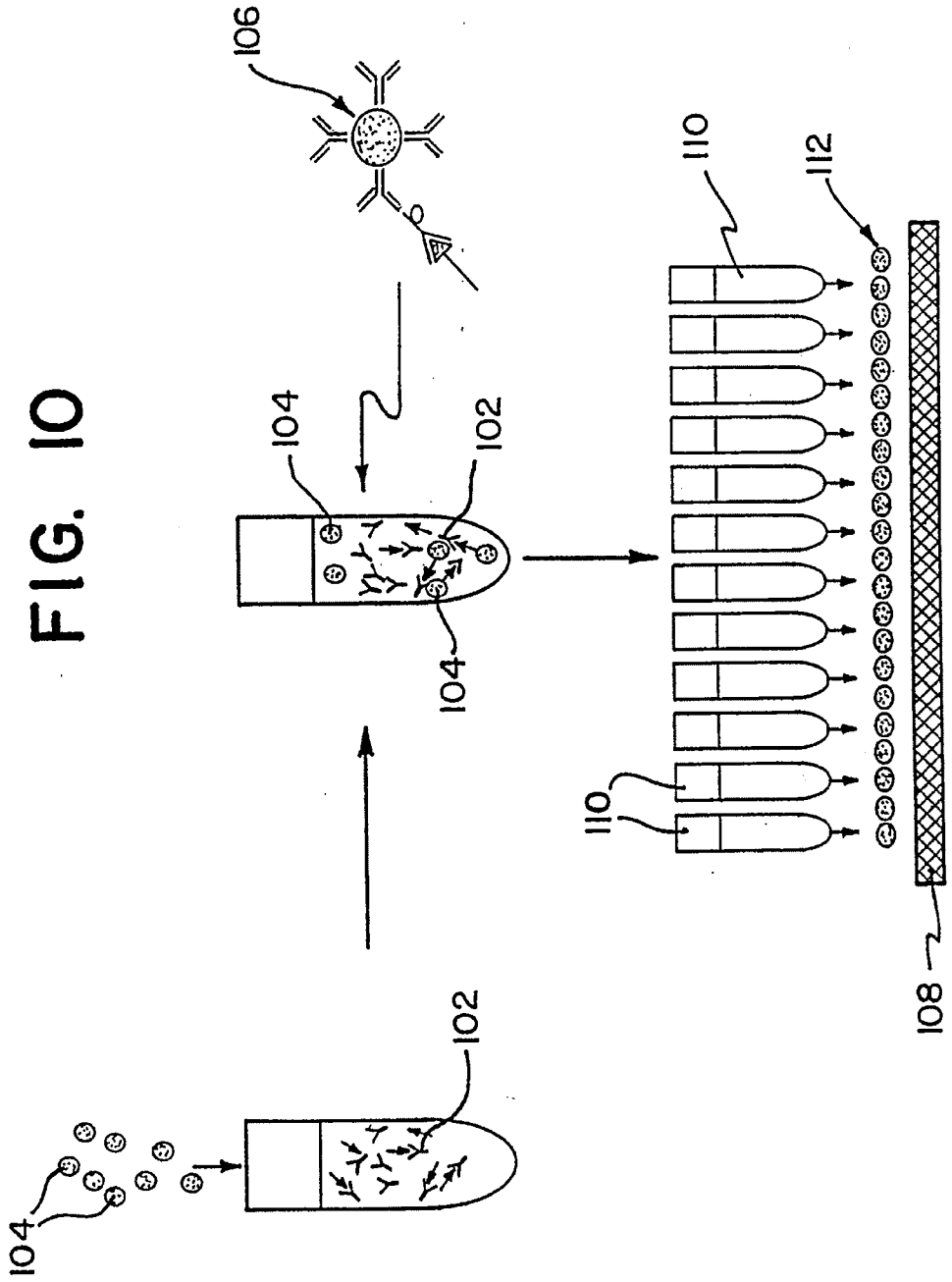


FIG. 11a

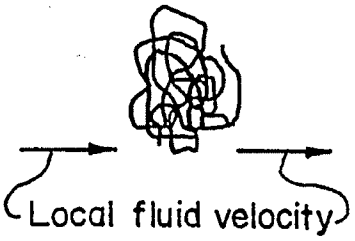


FIG. 11b

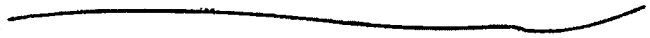


FIG. 11c

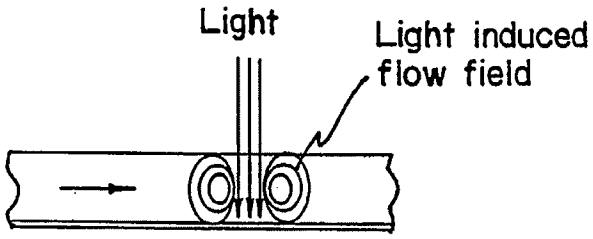


FIG. 11d

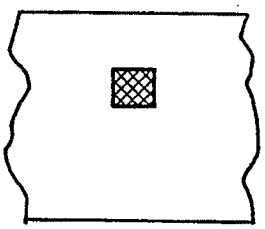


FIG. 11e

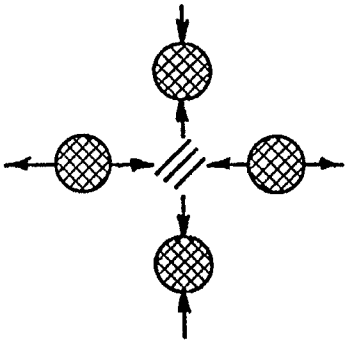


Figure 12

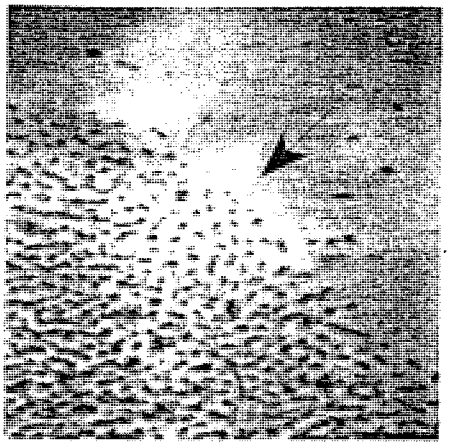


Figure 13

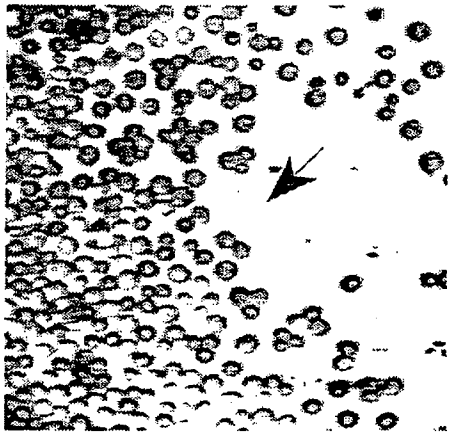


Figure 14

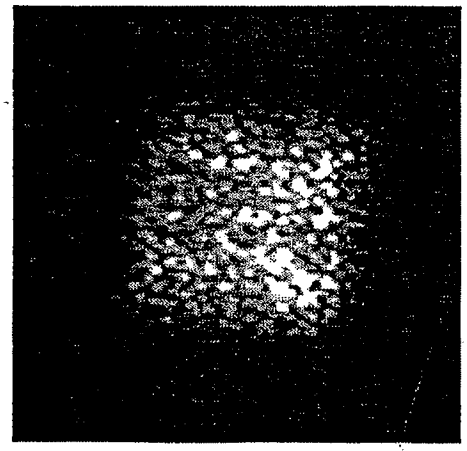


Figure 15

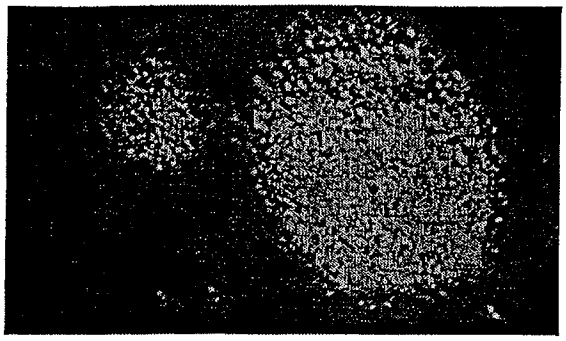


Figure 16

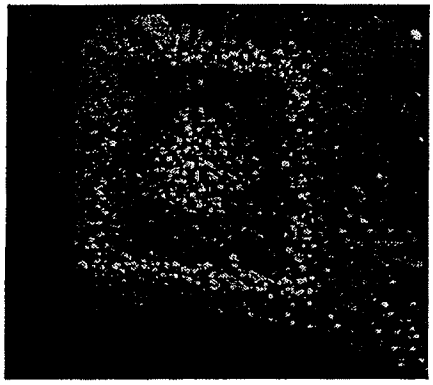


Figure 17

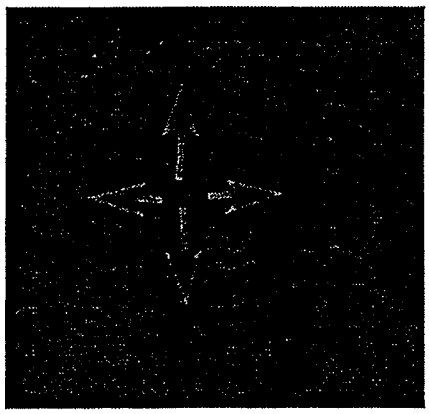


Figure 18

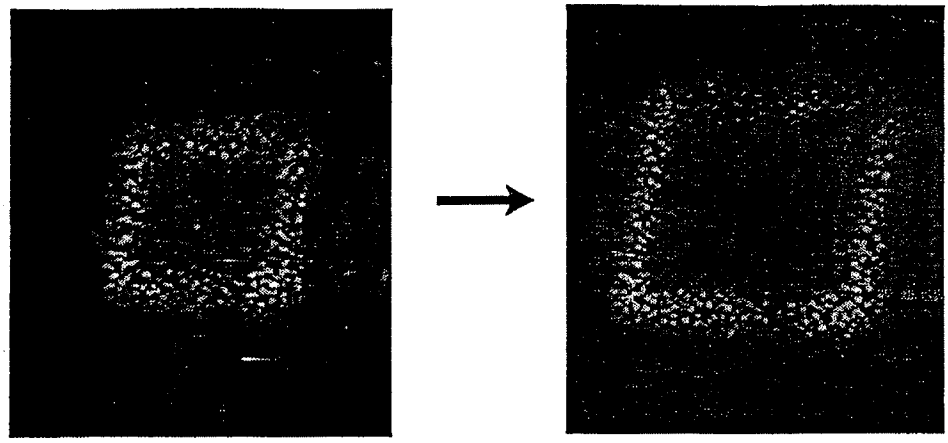


Figure 19

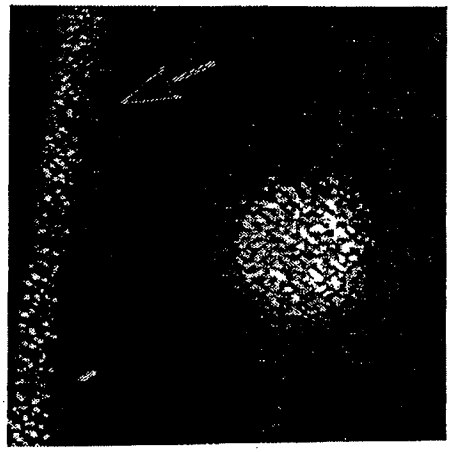


Figure 20

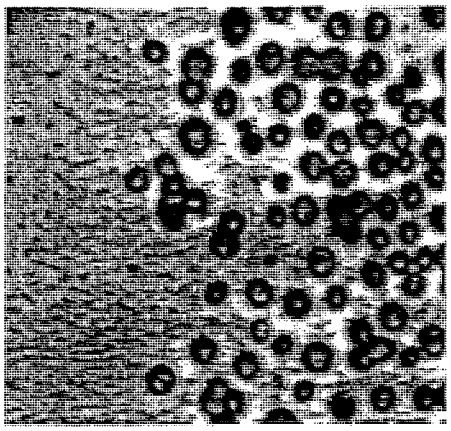


Figure 21

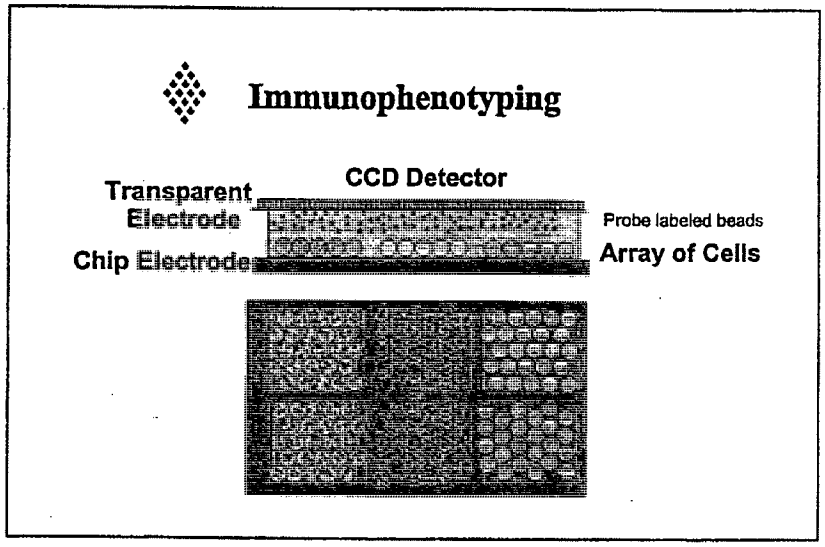


Figure 22

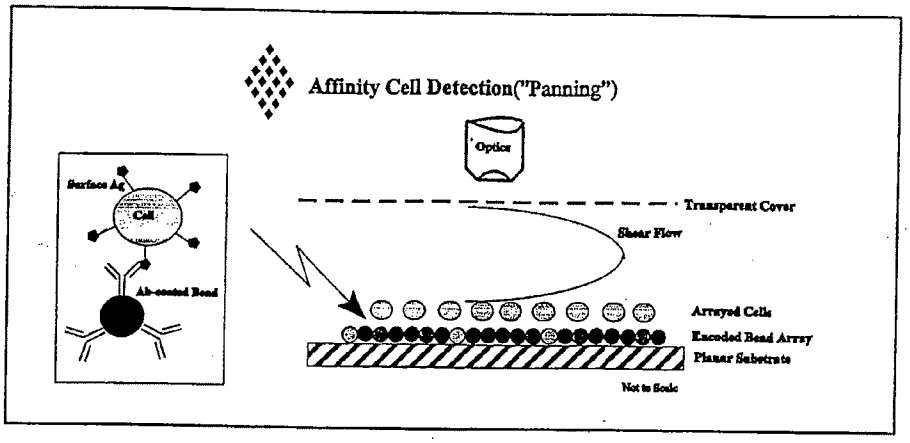


Figure 23

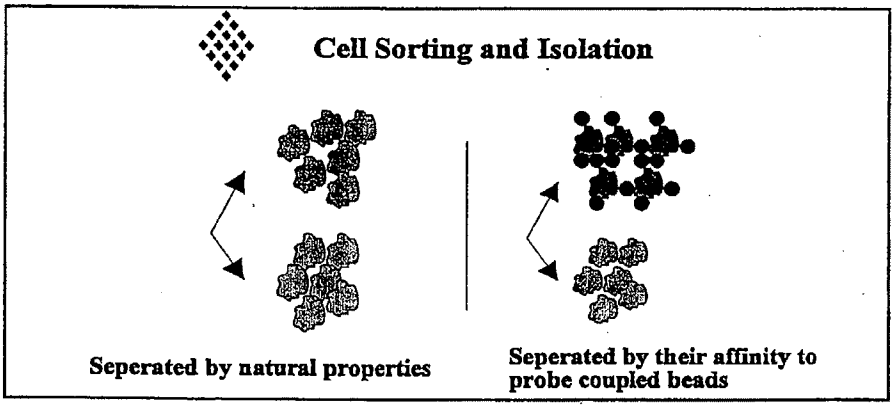


Figure 24

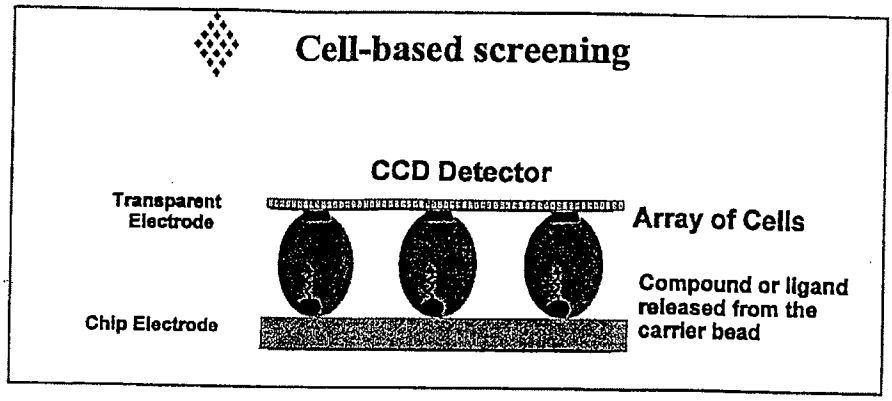


Figure 25

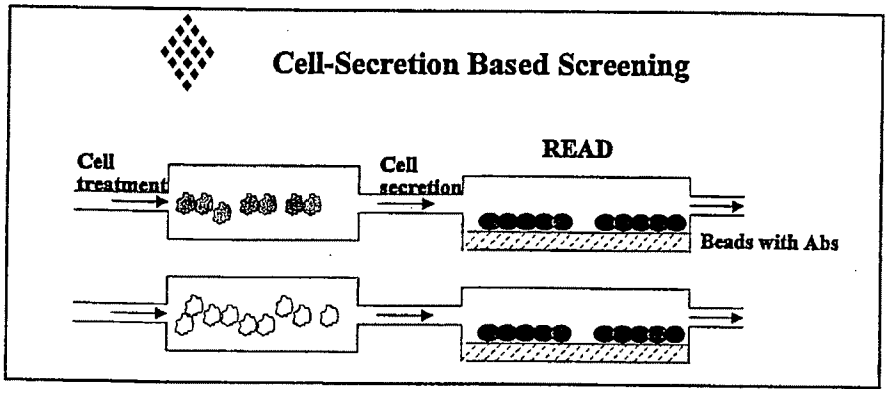


Figure 26

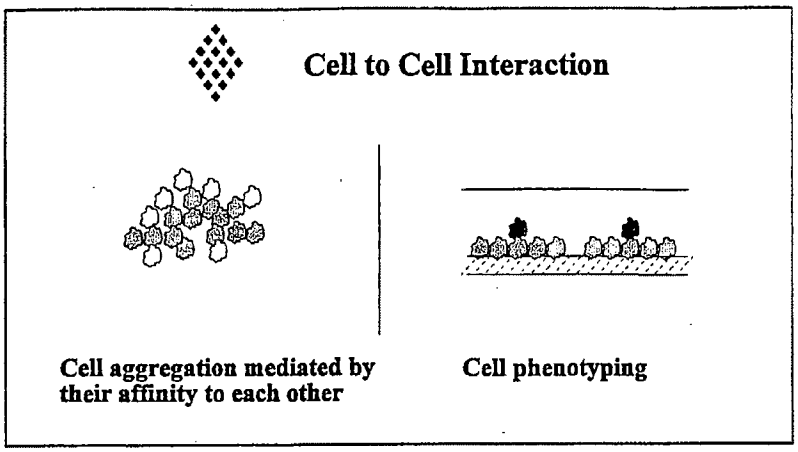
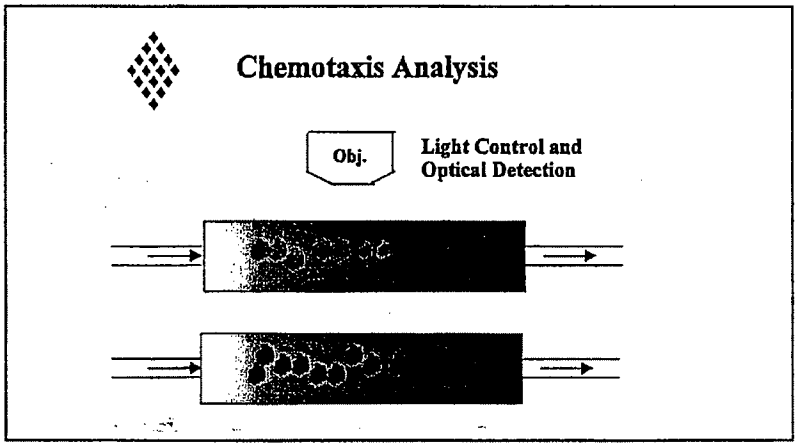


Figure 27



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/14957

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, STN.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	US 6,017,696 A (HELLER) 25 January 2000, col. 10, line 56-col. 20, line 41.	1-94
A	US 5,447,440 A (DAVIS et al.) 05 September 1995, col. 4, line 28-col. 8, line 34.	1-94
A	US 4,832,814 A (ROOT) 23 May 1989, col. 2, line 58-col. 3, line 51.	1-94
A	US 5,858,804 A (ZANZUCCHI et al.) 12 January 1999, col. 2, line 21-col. 8, line 62.	1-94

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

10 AUGUST 2000

Date of mailing of the international search report

03 OCT 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer

PENSEE T. DO

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/14957

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/14957

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

G01N 27/00, 27/02, 21/29, 15/06, 27/26, 33/543, 21/03; C12M 1/26; B29D 22/00;
H01M 4/36, 6/30, 6/36, 2/38, 6/04, 4/02; B01L 3/00; B01J 19/08.

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

204/193, 194, 400, 406, 407, 409, 410, 298.16; 422/50, 68.1, 82.01, 82.02, 82.03, 82.05, 99, 186, 186.13; 427/10; 428/36, 913; 429/104, 111, 113, 67, 126, 209, 232; 435/7.2, 7.21, 395, 401, 402, 403, 243, 292.1, 808; 436/518, 523, 524, 528, 531, 149, 150, 151, 165, 805, 806, 807.

B. FIELDS SEARCHED
Minimum documentation searched
Classification System: U.S.

204/193, 194, 400, 406, 407, 409, 410, 298.16; 422/50, 68.1, 82.01, 82.02, 82.03, 82.05, 99, 186, 186.13; 427/10; 428/36, 913; 429/104, 111, 113, 67, 126, 209, 232; 435/7.2, 7.21, 395, 401, 402, 403, 243, 292.1, 808; 436/518, 523, 524, 528, 531, 149, 150, 151, 165, 805, 806, 807.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

Group I, claim(s)1-34, and 75, drawn to a method and an apparatus of manipulating cells suspended at an interface between an electrode and an electrolyte solution.
Group II, claim(s) 24-34, 81, 86-91, drawn to an apparatus and a method for determining an interaction between cells.
Group III, claim(s) 35-74, 76-80, 82-85, drawn to an assembly, a method of making the assembly and a method of using the assembly.

The inventions listed as Groups I, II and III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the three groups are drawn to different methods and apparatus.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 January 2001 (04.01.2001)

PCT

(10) International Publication Number
WO 01/00875 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/US00/17421

(22) International Filing Date: 23 June 2000 (23.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/339,892 25 June 1999 (25.06.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/339,892 (CON)
Filed on 25 June 1999 (25.06.1999)

(71) Applicant *(for all designated States except US)*: **MOTOROLA, INC.** [US/US]; 1301 Algonquin Road, Schaumburg, IL 60196 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **KLIMECKI, Walter** [US/US]; 950 North Gilbert Road, Gilbert, AZ 85234 (US). **MARACAS, George** [US/US]; 2613 Bighorn Avenue, Phoenix, AZ 85048 (US).

(74) Agent: **MACHONKIN, Richard, A.**; McDonnell Boehnen Hulbert & Berghoff, Suite 3200, 300 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/00875 A2

(54) Title: NOVEL METHODS AND PRODUCTS FOR ARRAYED MICROSPHERE ANALYSIS

(57) Abstract: Among other things, the invention provides components, devices, and methods for arrayed microsphere analysis (AMA), and more particularly, for fluorescent arrayed microsphere analysis (FAMA).

NOVEL METHODS AND PRODUCTS FOR ARRAYED MICROSPHERE ANALYSIS

TECHNICAL FIELD OF THE INVENTION

The present invention provides, among other things, components, devices, and methods for arrayed microsphere analysis (AMA), and more particularly, for fluorescent arrayed microsphere analysis (FAMA).

BACKGROUND OF THE INVENTION

The identification and characterization of the constituent genes of the human genome involves two principal efforts. First, genes must be identified and sequenced. The gene sequence provides valuable clues as to the evolutionary derivation of newly identified genes. through sequence similarities with other known homologues. This type of sequence-homology information also provides clues as to possible functional relevance of novel genes. Second, and equally important sequence information, is that obtained by the analysis of gene expression. Gene expression analysis provides information regarding the functional relevance of a given gene product with respect to temporal, biochemical, and location-specific criteria. Thus, the knowledge of the levels of mRNA for a particular gene at particular times, under specific biochemical conditions, in a specific tissue, allows one to make predictions about the function of a gene for which functional significance is unknown; or to elucidate the fine points of biochemical pathway response and control for genes with well-characterized functions.

Presently, the techniques available to analyze mRNA levels are cumbersome. This has restricted the number of genes whose expression can be analyzed at a given time. One technique used commonly is RNA Northern Blot Analysis. In this technique RNA is isolated from a population of cells. The RNA is then size fractionated using gel electrophoresis. The fractionated RNA is then transferred to a membrane, where the RNA is immobilized. The membrane is then reacted with an excess of detection-labeled DNA ("probe") of known sequence, such that the probe specifically hybridizes with mRNA of the gene of interest. The greater the number of copies of mRNA for a particular gene, the greater the amount of hybridized, labeled probe on the membrane, yielding a larger signal.

While Northern Blot Analysis has proven an extremely effective method of RNA analysis, it often involves 72-96 hours of assay time to yield results, with analysis limited to 10-15 genes per experiment. In addition, since the signal from one probe molecule is well

below the minimum detectable signal of the assay, the measurement of low levels of RNA is variable, and ranges from difficult to impossible. The difference between a negative signal and a minimally detectable signal is essentially an analog gradation which is dependent upon the background signal ("noise") of that particular experiment on that particular day. The signal is also highly dependent upon the efficiency of the reactions which label the probe. Other RNA analysis techniques, such as the RNAse Protection Assay, have the potential of improved sensitivity, but are equally as cumbersome, potentially even more complex, and do not address the problem of clear identification of low level signals.

Thus, in an effort to bypass the cumbersome nature of RNA analysis techniques, recent efforts have utilized the concept of "DNA Microarrays" to allow the analysis of RNA levels from extremely large numbers of genes in a single experiment (Schena et al., Science, 270, 467-70 (1995)). The principle of these array-based RNA analysis techniques involves attaching short oligonucleotide probes, with specificity to particular genes, to a solid substrate such as a glass slide or an acrylamide pad. Because of their small size, thousands of distinct oligonucleotides can be attached to a small area, in defined locations. RNA is isolated from two distinct populations of cells and kept separate. One population is considered a control. The RNA from this population is reverse-transcribed in the presence of fluorescence-labeled nucleotides of a single color (e.g., green). The other population of cells is the "test" population which is treated similarly, except that the fluorescence-labeled nucleotide is a distinct color (e.g., red). The RNA samples are then mixed 1:1, and the mixture is hybridized to the immobilized oligonucleotides. Fluorescence at each oligonucleotide location is measured and the ratio of the fluorescence (i.e., in this example, green fluorescence as compared to red fluorescence) at a location is a measurement of the ratio of control mRNA to test mRNA. The overall fluorescence intensity reflects the quantity of mRNA at that site.

Indisputably, array analysis has greatly increased the number of genes which can be analyzed for expression in a single experiment. In spite of this, the array system still suffers from the problem of "analog" signal gradation, and consequently does not offer reliable, low level RNA analysis.

The present invention seeks to overcome these disadvantages attendant the prior art. The invention provides a novel method of RNA analysis which maintains the power of array analysis, while providing a discrete, "digital" signal of invariable intensity. This optimally is accomplished without significantly increasing the difficulty or labor input relative to currently used array-based systems of RNA analysis. These and other objects and advantages

of the present invention, as well as additional inventive features, will be apparent from the description of the invention provided herein.

BRIEF SUMMARY OF THE INVENTION

The present invention provides, among other things, components, devices, and methods for arrayed microsphere analysis (AMA), and more particularly, for fluorescent arrayed microsphere analysis (FAMA).

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts DNA (or cDNA) including one or more biotin-conjugated nucleotide(s) that is produced from a sample by isolating its DNA and labeling the DNA with at least one biotin-conjugated nucleotide, or that is produced from a sample by reverse transcription of sample RNA in the presence of at least one biotin-conjugated nucleotide. *Symbols*: B, biotin label.

Figure 2 depicts the synthesis of probe DNA and attachment to microspheres. *Symbols*: Circles labeled 1-6, microspheres that optionally differ in color.

Figure 3 depicts the placement of the probe/microspheres of **Figure 2** in separate microlocations of a multi-compartmentalized container (e.g., in separate wells of a microtiter plate). *Symbols*: Circles labeled 1-6, microspheres that optionally differ in color.

Figure 4 depicts the hybridization of the probe/microspheres of **Figure 3** in separate microlocations of a multi-compartmentalized container (e.g., in separate wells of a microtiter plate) with the biotinylated DNA (or cDNA) depicted in **Figure 1** to obtain probe/microsphere/DNA (or cDNA) hybrids. *Symbols*: B, biotin label; Circles labeled 1-6, microspheres that optionally differ in color.

Figure 5 depicts the addition to the probe/microsphere/DNA (or cDNA) hybrids depicted in **Figure 4** of streptavidin-conjugated magnetic particles to obtain streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids in the presence on non-bound probe/microsphere/DNA (or cDNA) hybrids. *Symbols*: B, biotin label; Circles labeled 1-6, microspheres that optionally differ in color; Stars, streptavidin-conjugated magnetic particles.

Figure 6 depicts the streptavidin-conjugated magnetic particles/probe/microsphere/DNA or (cDNA) hybrids that are selectively captured and removed from non-bound probe/microsphere/DNA (or cDNA) hybrids by exposing the mixture depicted in **Figure 5** to a magnetic field, followed by washing. *Symbols*: B, biotin

label; Circles labeled 1-6, microspheres that optionally differ in color; Stars, streptavidin-conjugated magnetic particles.

Figure 7 depicts recovery of the microspheres following removal (e.g., degradation) of bound DNA from the magnetic particles/probe/microsphere/DNA (or cDNA) hybrids of **Figure 6** and washing. *Symbols*: Circles labeled 2, 3, 4, or 6, microspheres that optionally differ in color.

Figure 8 depicts a fluidic switching assembly having a multi-compartmentalized container (10) such as a microtiter plate, and a micropipettes or other tubing system (20) having the same footprint and the container and providing means of inputting reagants and components to the container, as well as a means for removing items from the container for routing (30), e.g., to the detection system.

Figure 9 depicts an integrated fluidic switching assembly having a multi-compartmentalized container (10) such as a microtiter plate or other miniature reaction chambers, and one or more channels (34 and 38) as a means of providing reagants and components to the container, as well as a means for removing items from the container for routing (30), e.g., to the detection system, and a separately housed magnetic particle positioning system (40) located on top of the micropipettes or other tubing system (20) having the same footprint and the container and which provide means of inputting reagents and components to the container, as well as a means for removing items from the container.

Figure 10 depicts an integrated fluidic switching assembly having a multi-compartmentalized container (10) such as a microtiter plate or other miniature reaction chambers, and one or more channels (34 and 38) as a means of providing reagants and components to the container, as well as a means for removing items from the container for routing (30), e.g., to the detection system, and a separately housed magnetic particle positioning system (40) located on top of the micropipettes or other tubing system (20) having the same footprint and the container and which provide means of inputting reagents and components to the container, as well as a means for removing items from the container. This device differs from that depicted in Figure 9 in that the reaction chambers themselves are present inside of a housing, and in intimate association with the channels and micropipettes or other tubing system, thus allowing a "microchip" type assembly.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention provides, among other things, components, devices, and methods for arrayed microsphere analysis (AMA), and more particularly, for fluorescent arrayed

microsphere analysis (FAMA). This invention offers a novel method of nucleic acid (e.g., cellular RNA or DNA) analysis which maintains the power of array analysis, while providing a discrete, "digital" signal of invariable intensity. This is accomplished without significantly increasing the difficulty or labor input relative to currently used array-based systems of DNA/RNA analysis.

The AMA/FAMA method is summarized graphically in **Figures 1-7**, which are exemplary only, and in no way limiting of the invention. As can be seen from these figures, and as explained in greater detail below, **Figure 1** depicts DNA (or cDNA) including one or more biotin-conjugated nucleotide(s) that is produced from a sample by isolating its DNA and labeling the DNA with at least one biotin-conjugated nucleotide, or that is produced from a sample by reverse transcription of sample RNA in the presence of at least one biotin-conjugated nucleotide(s). **Figure 2** depicts the synthesis of probe DNA and attachment to microspheres. **Figure 3** depicts the placement of the probe/microspheres of **Figure 2** in separate microlocations of a multi-compartmentalized container (i.e., in separate wells of a microtiter plate), and **Figure 4** depicts their hybridization with the biotinylated DNA (or cDNA) generated from the sample to obtain probe/microsphere/DNA (or cDNA) hybrids. As depicted in **Figure 5,** streptavidin-conjugated magnetic particles are added to the probe/microsphere/DNA (or cDNA) hybrids to obtain streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids in the presence on non-bound probe/microsphere/DNA (or cDNA) hybrids. **Figure 6** depicts the streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids that are selectively captured and removed from non-bound probe/microsphere/DNA (or cDNA) hybrids by exposing the mixture depicted in **Figure 5** to a magnetic field, followed by washing, and **Figure 7** depicts recovery of the microspheres following optional removal (e.g., degradation) of bound DNA from the magnetic particles/probe/microsphere/DNA (or cDNA) hybrids.

Thus, in an especially preferred embodiment, the present invention provides, among other things, methods for analyzing the expression of a plurality of genes, e.g., present in a sample. Generally, the methods Arrayed Microsphere Analysis (AMA), and Fluorescent Arrayed Microsphere Analysis (FAMA), preferably comprise the following steps:

(a) isolating from the sample the RNA produced by the genes;

(b) reverse transcribing the RNA into cDNA in the presence of one or more biotin-conjugated nucleotide(s) to obtain biotin-conjugated cDNA;

(c) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to said cDNA of interest, wherein the probes are modified to allow end-attachment to labeled microspheres;

(d) obtaining probe/microspheres by attaching to the modified single-stranded DNA probes microspheres of a certain type, with microsphere type being determined by the combination, to provide a unique identification in the method of the microsphere, of:

(i) microsphere diameter, the microsphere being spherical; and

(ii) microsphere color, wherein microsphere color is determined by the presence within the microsphere of one or more dyes, with each dye being of a different color, different intensity, or different color and different intensity;

(e) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container and performing all subsequent steps in each the separate microlocation;

(f) hybridizing the biotin-conjugated cDNA to the probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/ cDNA hybrids;

(g) washing the probe/microsphere/cDNA hybrids;

(h) obtaining streptavidin-conjugated magnetic particles and incubating the particles with the probe/microsphere/cDNA hybrids under conditions such that they bind to the biotin-conjugated cDNA present in the probe/microsphere/cDNA hybrids, and streptavidin-conjugated magnetic particles/ probe/microsphere/cDNA hybrids are obtained;

(i) using a magnet to capture the streptavidin-conjugated magnetic particles/ probe/microsphere/cDNA hybrids; and

(j) assessing the streptavidin-conjugated magnetic particles/ probe/ microsphere/cDNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the expression of a plurality of genes present in the sample.

The key steps and considerations involved in AMA/FAMA are discussed below as follows: (1) sample (DNA or cDNA) preparation; (2) probe preparation; (3) microsphere (e.g., fluorescent microsphere, FM) preparation; (4) probe/microsphere attachment; (5) probe/microsphere placement to generate array; (6) probe/microsphere and DNA (or cDNA) hybridization; (7) selective capture of steptavidin-conjugated magnetic particle/probe/microsphere/DNA (or cDNA) hybrids; (8) potential for recovery of analyzed

DNA (or cDNA); (9) assessment of captured microspheres; and (10) miniaturization of AMA/FAMA assay.

Sample Preparation

According to the invention, the expression (or representation, as described below) is analyzed of a plurality of genes present in a sample. A "plurality of genes" means at least two genes, preferably from at least 2 to at least about 100,000 genes, and optimally, from at least 2 genes to as many genes as can feasibly be assessed in a particular array experiment. Feasibility assessment is based on cost, ease of manipulation, availability of instrumentation and software for analysis, sensitivity necessary or desired, etc. "Specific genes of interest" are those present in the sample whose expression (or representation) is being assessed in a particular experiment. Of course, a sample may contain many more genes than those whose expression (or representation) is being analyzed in an individual experiment, and separate experiments using the same sample can be employed to analyze expression (or representation) of different genes.

According to the invention, a "sample" desirably is any sample obtained from any natural source (i.e., isolated from nature). According to the invention, a sample comprises ribonucleic acid (i.e., RNA) when expression is being assessed, and comprises deoxyribonucleic acid (i.e., DNA) when representation is being assessed. In particular, preferably the sample comprises the RNA or DNA of a human or non-human animal. However, the sample nucleic acid can be taken from any organism. For instance, the sample can be taken from mammals (such as humans, non-human primates, horses, dogs, cows, cats, pigs, or sheep), from viruses, from plants, from microorganisms, from other living things in the environment, etc.

The biological sample (i.e., a sample containing DNA or RNA) can be isolated by any of a variety of means known to those skilled in the art. For instance, blood, tissue, or other bodily fluid samples can be taken from one or more individuals. Total RNA and/or messenger RNA (i.e., mRNA) desirably is isolated from a cell population to be analyzed. In general, mRNA should prove to be a superior sample for use in the method of the invention of analyzing expression as compared to total RNA. Any appropriate method such as is known in the art for isolation of total RNA or mRNA can be employed (see, e.g., Chirgwin et al., Biochemistry, 18, 5294 (1979), and others). Any commercially available kit for carrying out isolation of total RNA or mRNA similarly can be employed (e.g., Oligotex-dT resin, Qiagen, Valencia, California). Similarly, when representation of a particular DNA is being

assessed, desirably any method for its isolation and/or any commercially available kits or reagents can be employed.

The RNA from the sample is reverse transcribed, using reverse transcriptase enzyme, typical reaction buffers and deoxynucleotide triphosphates which consist of at least one biotin-conjugated nucleotide; for example, dATP, dCTP, dGTP and biotin conjugated dTTP (or, biotin-conjugated dUTP substituted for biotin-dTTP). Any appropriate method such as is known in the art for RNA reverse transcription can be employed (see, e.g., Friedman et al., Nucleic Acids Res. 4(10), 3455-71 (1977)). Similarly, any commercially available kit for carrying out RNA reverse transcription similarly can be employed (e.g., StrataScript RT-PCR kit, Stratagene, La Jolla, California).

Following cDNA synthesis, optionally the cDNA is separated from the RNA. This separation preferably is carried out by digesting the RNA. Alternately, desirably the biotin-conjugated cDNA is separated from the RNA by means of the biotin incorporated in the biotin-conjugated cDNA. For instance, by purifying the biotin-conjugated cDNA by binding to streptavidin beads and/or streptavidin packed in a column, such as is known in the art.

As indicated previously, in some applications, the nucleic acid of interest in the sample is DNA instead of RNA (i.e., cDNA transcribed from RNA, including from mRNA). For instance, RNA analysis is conducted where gene expression is being assayed. However, the presence and nature of the genomic DNA itself may be of interest in cases where gene copy number is being assessed, gene composition (i.e., presence of absence of mutations) is being assessed, etc. In these cases, biotinylated DNA suitable for analysis can be derived from the sample using methods commonly used by those familiar with the art. These methods include incorporation of biotin-derived nucleotide triphosphates onto the 3' end of the DNA using terminal deoxynucleotide transferase. Alternately, restriction endonucleases may be used to digest the double-stranded DNA in such a conformation that a 5' overhang is generated, which can then be a substrate for the incorporation of biotin-derived deoxynucleotide triphosphates by DNA polymerase, such as Klenow fragment, forming a blunt end of double-stranded DNA. Any other appropriate methods similarly can be employed.

Thus, in an especially preferred embodiment, the present invention provides, among other things, methods for analyzing the representation of a plurality of genes, e.g., present in a sample. This method also is a preferred method of the invention for carrying out Arrayed Microsphere Analysis (AMA), and Fluorescent Arrayed Microsphere Analysis (FAMA), and preferably comprises the following steps:

(a) isolating a sample to be tested for a DNA of interest;

(b) biotinylating the sample DNA to obtain biotin-conjugated DNA;

(c) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to the DNA of interest, wherein the probes are modified to allow end-attachment to labeled microspheres;

(d) obtaining a plurality of microspheres of differing types, with microsphere type being determined by the combination to provide a unique identification of the microsphere of:

(i) microsphere diameter, the microsphere being spherical; and

(ii) microsphere color;

(e) obtaining probe/microspheres by attaching the microspheres to the modified single-stranded DNA probes ;

(f) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each of the separate microlocation;

(g) hybridizing the biotin-conjugated DNA to the probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/DNA hybrids;

(h) washing the probe/microsphere/DNA hybrids;

(i) obtaining streptavidin-conjugated magnetic particles and incubating the particles with the probe/microsphere/DNA hybrids under conditions such that they bind to the biotin-conjugated DNA present in the probe/microsphere/DNA hybrids, and streptavidin-conjugated magnetic particles/probe/microsphere/DNA hybrids are obtained;

(j) using a magnet to capture the streptavidin-conjugated magnetic particles/probe/microsphere/DNA hybrids; and

(k) assessing the streptavidin-conjugated magnetic particles/probe/microsphere/DNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the representation of the DNA present in the sample.

Incorporation of a biotinylated nucleotide into the sample DNA or cDNA reverse transcribed from sample RNA allows subsequent conjugation with streptavidin (as well as providing a possible means of purifying nascent cDNA). While it is preferred according to the invention that biotinylation be employed as a means of providing later conjugation, any appropriate means that can be employed in the methods and devices according to the

invention can be used instead of biotinylation. If biotinylation is changed as the means of labeling, then, as later described, the binding partner for biotin (which is attached to a magnetized particle, i.e., preferably streptavidin) also must be changed. Binding partners analogous to (i.e., having the same function as) biotin/streptavidin are well known in the art.

In general, and as later described, it is desirable that only a single sample be analyzed in one microlocation of a multi-compartmentalized container. However, numerous genes (i.e., expression and/or representation) can be assessed in that microlocation, as will become apparent from the description of the invention provided herein.

Probe Preparation

Desirably a "probe" according to the invention is a single stranded oligonucleotide or polynucleotide, or any nucleic acid that can function as a probe as described herein. Single-stranded DNA probes, with sequences complimentary to the DNA (or cDNA) sequences from specific genes of interest (i.e., "target genes") are obtained. These probes thus are capable of specific hybridization to the cDNA of interest. Preferably the probe ranges from about 1 to about 5000 bases in length, desirably from about 1 to about 500 bases in length, especially from about 1 to about 300 bases in length, and particularly from about 1 to about 10 bases in length.

For relatively short probe sequences (i.e., <40-50 bases), synthesis desirably is accomplished on a DNA synthesizer (e.g., on a Cruachem PS 250 DNA Synthesizer, or other automated DNA synthesizer) using standard chemistries (e.g., phosphoramidite chemistry) as discussed, for instance, in Beucage et al., Tetrahedron, 48, 2223-2311 (1992), U.S. Patents 4,415,732, 4,458,066, 4,725,677, 4,973,679, and 4,980,960. Typically, synthetic oligonucleotides desirably are further purified, e.g., by HPLC followed by a 20% polyacrylamide gel/7 M urea. For longer probes, specific single stranded DNA preferably is synthesized using the Polymerase Chain Reaction (PCR), followed by an isolation technique to obtain purified populations of the desired single-stranded probe. Alternately, the oligonucleotides or polynucleotides of any size can be purchased as commercially prepared.

Regardless of the manner in which probe is prepared, preferably, during probe synthesis, provisions are made for one end of the probe to be anchored to the labeled microsphere (e.g., the fluorescent microsphere, FM). This desirably is accomplished by using an amine-conjugated base at one end of the probe, although other means of modification that provide for end-attachment also can be employed. 5' amine modification is a standard modification by most oligonucleotide synthesis facilities.

Preferably, the DNA probes are present in an array (i.e., as described further below). Since the probes are synthesized or otherwise obtained, they are of a known sequence, and are selected to be complementary to the DNA targets present in a sample being assessed. In an array according to the invention, desirably the sequence and/or position of every probe on the array is known, such that the probe array can be employed in sequencing and diagnostic applications, in forensics analyses, determinations of paternity, veterinary applications (e.g., thoroughbred testing), and the like.

In particular, preferably in the method of the invention, a different probe is employed for each microlocation of the array. Optionally, single probe is employed for hybridization at each array microlocation. As further described below, it also is desirable in some embodiments that more than one probe is hybridized at each array microlocation. Along these lines, desirably from 3 to 10 different probes are hybridized at each array microlocation. However, more generally, more than one probe optionally is hybridized at each microlocation, with the only limit on the number of probes that can be so hybridized being the ability to obtain and successfully detect such hybridization.

Should it be necessary or desirable according to a particular application, the probe itself also can include a detectable label. It is not necessary, however, according to the invention that each probe be labeled. In fact, it is important that any label that may be used for the probe does not interfere with the capture and/or detection of the microspheres according to the invention. Thus, it is desirable that the probe not be labeled with a fluorescent or luminescent label, and that the probe not include biotinylated nucleotides. With those provisos considered, any of the conventionally used methods of labeling (e.g., radioisotopes, enzymic, redox or other electrochemical labels, or other labels such as ligands, antigens, and the like) can be used according to the present invention to optionally label the probe and as set out, for instance, in: U.S. Patents 4,563,417, 4,581,333 and 4,582,789; EP Applications 119448 and 144914; and Prober et al., Science, 238, 336-340 (1987). Means of probe labeling, and other means of labeling to detect hybridization and other bioconjugation events are well known in the art.

Microsphere Preparation

Once the probes as described above have been obtained, the present invention calls for obtaining probe/microspheres by attaching to the modified single-stranded DNA probes microspheres of a certain type. According to the invention, "microsphere type" desirably is determined by the combination, to provide a unique identification in the method of the

invention, of: (i) microsphere diameter, the microsphere being spherical; and (ii) microsphere color.

The "diameter" of the microsphere preferably ranges from about 0.001 microns to about 100 microns, and most desirably ranges from about 0.02 microns to about 20 microns. The microsphere can be of any material that would function in the methods of the invention. Preferred materials for the microspheres include, but are not limited to, materials such as latex and polystyrene. The microspheres can be synthesized or purchased from commercial suppliers, e.g., from Molecular Probes, Eugene, Oregon, as well as other vendors.

The "color" of the microsphere color is determined by the presence within the microsphere of one or more dyes, with each dye being of a different color, different intensity, or different color and different intensity. Namely, the dye preferably is either a fluorescent or luminescent dye, such as are known in the art, e.g., AMCA, fluorescein, Rhodamine 6G, The Alexa Dye Series from Molecular Probes (Eugene, Oregon), Cy3, tetramethylrhodamine, lissamine rhodamine B, Texas Red, organic luminescent dyes, as well as others. Each dye desirably is of a color that can be distinguished from other color dyes using conventional instrumentation, such as flow cytometers and fluorescent particle counters made, for instance, by Beckman-Coulter, Fullerton, California. Alternately, each dye desirably can be of the same color, but of an intensity of color that can be distinguished from other intensity of color dyes using conventional instrumentation. In some embodiments, it may be desirable to use a single dye per microsphere to identify the microsphere. Preferably, more than one dye (e.g., fluorescent and/or chemiluminescent; same or different colors; same or different intensities of color) can be used per microsphere to provide a unique identification of the microsphere. The most important consideration for the microspheres is that the dyes in each microsphere are present alone or in combinations or at intensity levels which can be detected with minimal overlap by the assay detection system (i.e. each dye or dye combination can be easily, discretely identified by the analysis instrumentation).

Thus, a plurality of microspheres can be used in the invention wherein the different types of microspheres each have: (1) the same size but differ in color; (2) the same color but differ in size; (3) differing size and differing color; (4) the same size and color and differ in intensity of color; (5) the same size but differ in color and intensity of color; (6) the same color but differ in size and intensity of color; (7) differing size and differing color and differing intensity of color; or (8) any other combination of size/color/intensity that uniquely identifies the microsphere. Since luminescent and fluorescent dyes are detected by different means, the combination in a microsphere of luminescent and fluorescent dyes further can be

employed to uniquely identify the microsphere. A particularly preferred microsphere according to the invention, however, is a fluorescent microsphere (i.e., FM). Furthermore, the combination within a microsphere of more than one color and/or intensity of color dye adds additional options for unique identification of the microsphere.

As further described below, it is important to the invention that each microsphere be capable of being uniquely identified. However, placement in the array, and attachment to a particular probe, further provide another level of complexity. This additional level of complexity means that it is possible that a microsphere of a defined type (i.e., size, color, intensity of color) can be used more than once in an array, for instance, by attachment to differing probes known to be present in different parts of the array.

Probe/Microsphere Attachment

Probe attachment to microspheres can be accomplished by any one of several possible techniques. For instance, one such technique employs microspheres with carboxyl groups on their surface. These microspheres desirably have their surface-carboxyl groups activated with a carbodiimide. They then optimally are reacted with the amine-conjugated probes, forming a covalent, amide bond (e.g., as described in Molecular Probes Product Information Sheet, "Working with FluoSpheres® Fluorescent Microspheres", Section 7, "Covalent Coupling of Proteins to Carboxylate-Modified Microspheres", ©1997, as well as by other means).

The attachment of the probes to the microspheres desirably is at a ratio of probe number per microsphere so as to allow the most accurate and feasible correlation between microsphere number enumerated at the conclusion of the assay and the number of nucleic acid molecules which are the target of that particular probe in the sample to be assayed. The number of probes to be attached per microsphere is experimentally determined, and preferably is varied to accommodate individual expression levels. For example, for a very low abundance mRNA or DNA (i.e., a low level expression or low copy number gene), a low probe/microsphere ratio would produce a system in which microsphere number would more closely reflect actual number of nucleic acid copies. At the extreme end, one probe attached per one microsphere would produce a 1:1 correspondence between microsphere number and nucleic acid copy number. Alternatively, in order to avoid "saturating" the number of available probe-microspheres by high-expression level genes, a higher probe/microsphere ratio would be desirable. Of course, the number of probes attached per microsphere is a factor that must be assessed in the end stage analysis of gene expression. In cases where it is

anticipated that the gene has a high level of expression, it may be desirable to dilute the sample, or run the assay on various dilutions of sample, and still employ a one probe per microsphere. Alternately, it may be desirable to run an assay in duplicate, or triplicate, and compare results, for instance, with the first assay containing all probe/microsphere combinations at a ratio of one probe per microsphere, the second assay containing all probe/microsphere combinations at a ratio of ten probes per microsphere, and the third assay containing all probe/microsphere combinations at a ratio of one hundred probes per microsphere. These are all obvious and routine manipulations that are well within the skill and understanding of those conducting the assays.

Probe/Microsphere Placement to Generate Array

The method of the invention desirably is carried out by placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each of the separate array microlocations. Thus, according to the invention, each probe/microsphere is placed in a specific microlocation of a multi-compartmentalized container. A "multi-compartmentalized container" is a multiwell plate (e.g., a microtiter plate having separate wells) or other suitable container as described below. Preferably each microlocation of the multi-compartmentalized container is separately addressable, e.g., to facilitate probe/microsphere placement.

A "microtiter plate" is a plate comprising many microlocations which is commercially available. Generally, a microtiter plate comes in a variety of sizes and configurations, e.g., 96-wells, 384-wells, etc. Any appropriately sized microtiter plate can be employed in the invention. Furthermore, multi-compartmentalized containers having microlocations even smaller than those present on commercially available microtiter plates also can be used. For instance, the reactions can be carried out in separation microlocations of a microchip, e.g., as where the microchip's surface is etched to create separate indentations, or where holes are created within,with each indentation or hole serving as a microlocation where an independent probe hybridization reaction can take place, or in microwells cast in plates, elastomers or metals, etc.

The material of the multi-compartmentalized container is made optionally is any solid substrate that can be employed in the invention, e.g., film, glass, Si, modified silicon, ceramic, plastic, resins, or any type of appropriate polymer such as (poly)tetrafluoroethylene, (poly)vinylidenedifluoride, or polydimethylsiloxane (PDMS). Preferred substrates according to

the invention are glass and plastic. The solid substrate can be any shape or size, and can exist as a separate entity or as an integral part of any apparatus (e.g., cuvette, plate, vessel, device, and the like). It further is assumed that the material of the substrate is such that it does not interfere with the hybridization reaction, and can withstand any necessary manipulations, e.g., heating, centrifugation, etc., that are needed in order to effect hybridization.

By "array" (or "matrix") is meant an arrangement of microlocations (i.e., "locations") on the substrate. The microlocations can be arranged in two dimensional arrays, three dimensional arrays, or other matrix formats. The number of microlocations can range from one microlocation to a plurality of microlocations (e.g., from two to hundreds of thousands).

Desirably an array is comprised of microlocations in rows and columns. Microlocations can be of any shape, and preferably, are either round, square, or rectangular. The thickness and dimensions of the multi-compartmentalized container, and/or the arrays produced according to the invention can vary dependent upon the particular needs of the user.

Thus, according to the invention, specific probes, attached to specifically colored and sized microspheres are then placed into an arraying container such as a multi-well plate (e.g., having 96-, 384-, or 1536- wells). As discussed above, while it is possible that more than one different probe/microsphere could have the same microsphere size and color, desirable each probe within a single microlocations (e.g., well) will be attached to a microsphere of a unique combination of color and size. Thus, for example, one well could contain probes for the BRCA1 gene and the BRCA2 gene and the c-Myc gene. There could be thousands of BRCA-1-specific probe/microspheres in the array, thousands of BRCA-2-specific probe/microspheres in the array, etc. Importantly, it may be desirable for a certain experiment that all the BRCA-1-specific probe/microspheres in the array will be of the same color-size-intensity combination, and that color-size-intensity combination is not repeated within that well. Consequently, the combination of color, size, intensity and well location (i.e., microlocation, or exact position on the multi-compartmentalized container) of microspheres uniquely identify those microspheres as the probes for a specific gene. It is this novel utilization of color, size, intensity and well location which confers the array capability to AMA/FAMA. The theoretical underpinnings of the invention further are described in Example 1.

The absolute number of probe/microspheres which can be placed in a single microlocation desirably is experimentally optimized in the development of the assay for particularized uses. This number will depend upon the volume of the microlocation (i.e., the

volume of the well), the number of probe/microsphere replicates per gene, probe length, as well as a variety of other factors (i.e., factors inherent to each particularized use).

By comparison, as a less preferred embodiment, but still another way to do the assays according to the invention, the position of one microsphere in the assay relative to any other could be randomized. This would be accomplished by having unique identifiers for microsphere type (e.g., size, color, and/or intensity of color), as previously described, and only using a particular type of microsphere to attach to a specific type of probe (e.g., using a particular microsphere only once in an assay). In this embodiment, hybridization could be carried out in a container that is not compartmentalized (e.g., in a fluid-filled sac). The number of recovered particular microspheres (e.g., size, color, and/or intensity) in this case would tell the number of hybridizations per particular probe. Of course, this embodiment allows a substantially lesser number of nucleic acid species to be assessed.

The present invention accordingly further provides an improvement of an array-based hybridization method for nucleic acids, the improvement comprising the use of fluorescent or luminescent microspheres for attachment to single stranded probes to obtain probe/microspheres, the probes directed to (i) genes of interest present in the sample, or (ii) cDNA reverse-transcribed from RNA produced by specific genes of interest present in the sample,

wherein the combination of color, and size of probe/microspheres uniquely identify those probe/microspheres as specifically hybridizing to:

(a) the specific gene itself, and the level of signal detected corresponds to the level of representation of the gene in the sample; or

(b) cDNA reverse-transcribed from RNA produced by a specific gene, and the level of signal detected corresponds to the level of expression of the gene.

Furthermore, the invention also desirably provides an improvement of an array-based hybridization method for nucleic acids, the improvement comprising the use of fluorescent or luminescent microspheres for attachment to single stranded probes to obtain probe/microspheres, the probes directed to (i) genes of interest present in the sample, or (ii) cDNA reverse-transcribed from RNA produced by specific genes of interest present in the sample,

the hybridization method is carried out by placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container and performing all subsequent steps in each separate microlocation,

wherein the combination of color, size, and position of microlocation of the probe/microspheres uniquely identify those probe/microspheres as specifically hybridizing to:

(a) the specific gene itself, and the level of signal detected corresponds to the level of representation of the gene in the sample; or

(b) cDNA reverse-transcribed from RNA produced by a specific gene, and the level of signal detected corresponds to the level of expression of the gene.

Probe/Microsphere and DNA (or cDNA) Hybridization

According the invention, preferably the biotin-conjugated DNA (or cDNA) produced from the sample is hybridized to the probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/DNA (or cDNA) hybrids. This preferably is done by dividing up sample DNA (or cDNA) into aliquots for distribution into each microlocation (e.g., each well of a microtiter plate). Probe/microspheres desirably are mixed with the biotin-conjugated DNA (or cDNA) and hybridized under standard conditions using standard buffers within each well (as further described below). Washing of the probe/microsphere/DNA (or cDNA) hybrids is then carried out to wash non-hybridized DNA (or cDNA) will be washed away from hybridized DNA (or cDNA) using standard stringency hybridization and washing conditions. Desirably, all the probe/microspheres are collected following the wash step(s) using centrifugation to "pellet" the probe/microspheres. At this point some of the pelleted probe/microspheres will be hybridized to DNA (or cDNA), while some of the recovered probe microspheres will not. Selective capture of those pelleted probe/microspheres which are hybridized to DNA (or cDNA) is done in the next step of the process.

Generally, according to the invention, hybridization is carried out by standard means. Reaction conditions for hybridization include provision of appropriate salts and buffers with each enzyme incubation. The reaction conditions for hybridization desirably are maintained such that the probe/microspheres stably and specifically hybridize to the target biotin-conjugated DNA (or cDNA). The hybridization reaction mixture preferably is maintained under hybridizing conditions for a time period sufficient for the probe/microspheres to hybridize to complementary nucleic acid sequences present in the biotinylated DNA (or cDNA) sample to form probe/microsphere/DNA (or cDNA) hybrids. Such "hybridizing conditions" (i.e., "hybridization conditions") includes subjecting a hybridization reaction mixture to time, temperature, and pH conditions needed to allow the probe/microspheres to hybridize with the complementary nucleic acid sequence of the sample DNA (or cDNA).

Such time, temperature, and pH conditions are well known in the art, and depend, for instance, on the length and composition (e.g., guanidine and cytosine content) of the probe to be hybridized, the degree of complementarity between the probe and the sample DNA (or cDNA), the stringency of the hybridization employed, and the presence of salts or additional reagents in the hybridization mixture such as may affect the kinetics of the hybridization.

As defined herein, "stably" (or stable) hybridizing means that the hybridization desirably has a T_m greater than the temperature under which the hybridization is to be carried out (i.e., typically from about 20°C to about 40°C). "Specific" hybridization means that the length and/or sequence complexity of the probes employed for hybridization is/are sufficient to prevent any non-desirable spurious hybridization such as might occur between sequences that are only partially complementary.

Typical hybridization conditions include the use of solutions buffered to pH values ranging from about 4 to about 9, hybridization temperatures ranging from about 18°C to about 70°C, and for time periods ranging from about 0.5 seconds to 24 hours. Hybridization preferably is carried out for from about 15 to about 30 minutes at room temperature in a solution that comprises about 1.5 M NaCl and 10 mM EDTA, although other hybridization conditions also desirably can be employed.

Typically, preferably hybridization is carried out in a hybridization mixture that contains up to about 50% deionized formamide and about 10% dextran sulfate in 2X SSC (with 20X SSC containing 175.32 g sodium chloride and 88.23 g sodium citrate, brought to 1 liter with distilled water), or an equivalent reaction mixture in which hybridization can be carried out. Optimally, prior to hybridization, the probe/microspheres are heated to approximately 95°C for from about 1 to about 10 minutes in order to eliminate any secondary structure formation in the strands and/or any binding between strands. The heated mixture preferably is allowed to cool to about room temperature prior to use in hybridization. Desirably the sample DNA (or cDNA) is prehybridized in the hybridization mixture at about 50°C for about 5 minutes. The probe/microspheres are then added, and incubation is allowed to proceed, preferably at a temperature that ranges from about room temperature to about 37°C (depending on probe characteristics), desirably for at least from about 10 seconds to overnight, preferably for about 2 hours, and even more preferably for at least about 10 seconds to about 10 minutes, to effect the hybridization of the probe/microsphere to the hybridizing strand of the DNA (or cDNA) sample. Optionally, the hybridization temperature can be increased or decreased from room temperature.

Preferably, washing of the probe/microsphere/DNA (or cDNA) hybrids is carried out. The cooled mixture preferably is washed to remove unhybridized DNA from the mixture, desirably which is accomplished by successive incubations in wash solution (e.g., as is well known in the art), wherein fresh wash solution is employed for each wash. Alternately the wash preferably is accomplished by continuous flow of fresh wash solution over the probe/microsphere/DNA (or cDNA) hybrids. Typical characteristics of washing include a temperature of approximately 65 °C, a low salt concentration, and the presence of a detergent such as sodium dodecyl sulfate. For instance, post-hybridization washes of about 5 minutes each can be done as follows: preferably 4X SSC for up to two washes, optimally followed by 2X SSC for up to two washes, desirably followed by a 0.1 X SSC wash, and optionally followed by a 0.05X SSC wash. Wash conditions used in the invention can vary depending on many factors that have been well described in the art, e.g., probe length, probe composition, etc. Furthermore, it may be desirable to carry out the wash at room temperature, and using a higher salt concentration (or even the same buffer as is employed for hybridization).

Selective Capture of Probe/Microsphere/DNA (or cDNA) Hybrids

According to the invention streptavidin-conjugated magnetic particles are employed to selectively capture the probe/microsphere/DNA (or cDNA) hybrids. Such streptavidin-conjugated magnetic particles can be purchased from any appropriate commercial vendor (or synthesized), and desirably, are obtained from Dynal, Lake Success, New York. Preferably, the streptavidin-conjugated magnetic particles are incubated with the probe/microsphere/DNA (or cDNA) hybrids under conditions such that they bind to the biotin-conjugated DNA (or cDNA) present in the probe/microsphere/DNA (or cDNA) hybrids, and streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids are obtained. Preferably this incubation is carried out with an excess of streptavidin-conjugated magnetic particles using standard conditions. Binding can be carried out by any appropriate means, and particularly, can be carried out by the means described by the commercial vendor of the streptavidin-conjugated magnetic particles (e.g., in the Dynal product literature, the binding buffer is approximately 1M NaCl).

Following standard wash steps (e.g., with buffer), preferably a magnet is employed to capture the streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids. The magnet optionally can be a permanent-type magnet, or can be an electromagnetic field. The magnetic field applied to the mixture consolidates a "pellet"

comprised of streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids. Optionally, the non-bound magnetic particles are washed away from this pellet (e.g., with buffer). Subsequently, the pellet consolidated by the magnetic field is isolated.

Assessment of Captured Microspheres

In the next step of the method, the captured microspheres are assessed. Namely, using methods analogous to flow cytometric fluorescent particle detection, preferably each well is aspirated of its streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrid, and these hybrids desirably are passed through a fluorescence analysis system capable of simultaneously measuring the dye color and intensity via fluorescence emission (e.g., at one or more excitation wavelengths) or luminescence, and the microsphere diameter via a measurement such as electrical conductance or low angle forward light scatter. Preferably, the number and type of microsphere captured per microlocation is determined, and this information is assessed along with information regarding the type of probe employed per microlocation, to analyze the expression of a plurality of genes present in the sample.

It may be desirable to purify the microspheres from other attached components prior to microsphere analysis. This can be done by a variety of means. For instance, optionally the streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrid pellet is exposed to denaturing conditions under a magnetic field in order to remove hybridized DNA (or cDNA) conjugated to the magnetic particles from the streptavidin-conjugated magnetic particles/probe/microsphere/ DNA (or cDNA) hybrids. The probe/microspheres (i.e., the fraction not attracted to the electric field) are then collected by centrifugation. Optionally, the probe can be degraded, e.g., with DNAse. The resultant microsphere optionally is washed, and subjected to further analysis.

Thus, the invention desirably provides a means of assessing the streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the expression of a plurality of genes present in the sample.

Potential for Recovery of Analyzed DNA (or cDNA)

In some instances, post-analysis recovery of the sample DNA (or cDNA) is desirable. For example, in homology-based hybridization experiments, in which the exact sequence of the DNA (or cDNA) is unknown at every base, recovered sample could be further analyzed

for exact sequence data, either by direct sequencing or by using the recovered sample as a template for further amplification followed by sequencing. In another example, the assay could be used to screen for rare events (e.g., mutation events), such as DNA or mRNA from rare tumor cells in a predominately normal tissue sample. In these types of applications, it could be valuable to evaluate the DNA (or cDNA) bound to the microsphere as a way to verify that the positive signal, in fact, came from the mRNA of concern. In cases in which the outcome of the assay would direct future therapy of a patient, such verification would minimize the chance of false positive signals.

Two considerations are key to this sample recovery. First, the hybridized DNA (or cDNA) must be retained on the microsphere throughout analysis. Second, a technique must be used to allow sample recovery. Fluorescence Activated Cell Sorting (FACS) is a common technology that is available in most hospitals and universities, with equipment for carrying out FACS being sold by Becton-Coulter and other commercial vendors. FACS uses the principle of flow cytometry, and adds electrically induced deflection of certain droplets to collect them. FACS using the principle of flow cytometry and the electrostatic deflection of selected particles into a container desirably is one technique which could be employed according to the invention to accomplish selective recovery of sample DNA.

The biotinylated DNA (or cDNA) can be recovered from the process as described above. Namely, as the streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrid pellet is exposed to denaturing conditions under a magnetic field, biotinylated DNA (or cDNA) will be recovered bound to the streptavidin-conjugated magnetic particles and subjected to further analysis, such as PCR or sequencing. Optionally, the DNA (or cDNA) can then be recovered from the streptavidin-conjugated magnetic particles, for instance, by treating with hot formamide (e.g., at a temperature of from about 70-95°C, optimally about 90°C).

Thus, the present invention optionally provides a method wherein the DNA (or cDNA) is recovered from the streptavidin-conjugated magnetic particles/probe/microsphere/DNA (or cDNA) hybrids. This allows further manipulation and assessment of the DNA (or cDNA).

Miniaturization of AMA/FAMA assay

An AMA/FAMA assay that detects and analyzes many targets in parallel requires handling of a large number of fluids and microspheres. While these requirements can be satisfied using conventional microtiter plate robotic fluid handling systems, scaling the

process up to detect over ten times more samples can become unwieldy. Two approaches described below optimally can be employed to increase throughput of this assay. Like numbering is employed in **Figures 8** and 9 to depict like components in the devices.

(a) Parallel fluidic switching assembly

A preferred parallel fluidic switching assembly is shown in **Figure 8**. As can be seen from this figure, a multi-compartmentalized container (10), preferably a conventional microtiter plate is used to perform the assay. In the simplest case, desirably conventional dispensing robotics are used to supply the microlocations (e.g., microtiter wells) with hybridization and washing reagents and probe microspheres. The fluidic switching assembly desirably consists of a series of micropipettes or other tubing arrangement (20) to supply the microlocations and which have the same footprint (i.e., configuration, so they components mesh) as that of the appropriate microtiter plate. A magnetic assembly (e.g., electromagnets located in the supplying micropipettes or tubing arrangement, or an assembly equivalent to electromagnets) optimally is programmed to magnetically attract labeled microspheres from specific wells and draw them up into the appropriate pipette or tube. Movement of the microspheres in a transport liquid (e.g., buffer) would draw the microspheres into the switching assembly and then provide routing (30) into the appropriate detection system. Using this preferred embodiment, optimally each microlocation (i.e., each well of the microtiter plate) is interrogated and analyzed under computer control without the need for robotic handling of the finished assay products.

Figure 8 is depicted as having a single port (30) for providing input and output to the multi-compartmentalized container. However, multiple ports can be configured in the device, for instance, multiple ports for input and/or output, and with the ports for inputting and aspirating off reagents and components (e.g., buffers for hybridization, probe/microspheres, magnetized particles, etc.) being separated. It further is possible that the micropipettes or other tubing arrangement itself provides a means for inputting and aspirating off reagents and components. Further not depicted in **Figure 8** is a configuration wherein a magnetic particle positioning system is housed separately from the series of micropipettes or other tubing arrangement, and is employed to correctly position magnetic particles (i.e., either before or after hybridization to probe/microspheres). In this scenario, essentially, the electromagnets are positioned not within the micropipettes or tubing arrangement itself, but are sequestered in a separate area, e.g., positioned above the micropipettes or other tubing arrangement (20), as depicted by (40) in **Figure 9**, and as further described below. Alternately, the electromagnets can be position below the multi-compartmentalized container, i.e., either

alone, or, in addition to electromagnets positioned above the micropipettes or other tubing arrangement.

(b) Integratcd fluidic switching assembly

As depicted in **Figure 9**, in this preferred embodiment, desirably the multi-compartmcntalized container (10) is a microtiter plate, or a plate containing an array of microlocations (i.e., "reaction chambers") having dimensions less that 5 mm x 5mm (i.e., smaller than the typical microtiter plater). The reaction chambers preferably are etched or molded into glass or plastic or other appropriate material and, when attached to the fluidic switching assembly, form an environment that is sealed from the ambient room conditions and thus reduces the incidence of contamination by unknown and unwanted environmental agents.

Preferably the fluidic switching assembly consists of one or more etched or molded channels (34 and 38) in glass or plastic that serve to move the fluids, microspheres and magnetic particles and other reagents (optimally using hydrostatic pressure), and serve as an inlet into the reaction chambers. Similarly, desirably the fluidic switching assembly consists of one or more etched or molded channels (30) in glass or plastic that serve to move the fluids, microspheres and magnetic particles and other reagent (optimally using hydrostatic pressure), and serve as an outlet from the reaction chambers, e.g., to a detection device. It is possible for a channel to serve as inlet and outlet, i.e., with appropriate partitioning.

These channels desirably connect to a series of micropipettes or other tubing system (20) having the same footprint and the multi-compartmentalized container, and which provide means of inputting reagents and components, etc. to the container, as well as a means for removing these items from the container. Within the channels and/or series of micropipettes or other tubing system (i.e., 20 both in the present device depicted in **Figure 9**, as well as that depicted in **Figure 8**), a series of micromachined valves desirably are employed to control the quantity and path of fluids/components through the system (i.e., from the reagent inputs into the appropriate reaction chamber, out of the appropriate reaction chamber and to the detection system). These opening and closing of these valves preferably is controlled by hydrostatic pressure, puffs of air, and/or are electrically activated.

A magnetic particle positioning system (40) located above the reaction vessel plate (or optimally is integrated in series of micropipettes and tubing (20), as depicted in **Figure 8**) preferably is used to stir, attract and repel the magnetized particles. While the magnetic particle positioning system is positioned above the reaction vessel plate (10) and above the

series of micropipettes or other tubing system (20) in **Figure 9**, this magnetic particle positioning system optionally can be positioned below the series of micropipettes or other tubing system (20). and below and directly attached to the reaction vessel plate (10). Furthermore, optionally there can be more than one magnetic particle positioning system, for instance, one positioned above and directly attached to the series of micropipettes or other tubing system (20). and one positioned below and directly attached to the reaction vessel plate (10).

This magnetic positioning system (i.e., as well as that depicted in **Figure 8**) desirably consists of a multitude of computer controlled electromagnets, or permanent magnets controlled by electromagnets, which control the trajectory of the magnetic particles through the system. Thus, desirably, a combination of magnetic positioning and/or fluid pressure are used to move the magnetic particles into the detection system for analysis. The electromagnets themselves desirably are any item comprised of metal wire wound around a ferrous core (or other magnetizable core) with enough repetitions of winding such that the item is capable of functioning as an electromagnet. These electromagnets and/or permanent magnets can be prefabricated and deposited into a device, or constructed as part of the device, e.g., using photolithography techniques. Preferably according to the invention, the number of electromagnets is equivalent to the number of microlocations (e.g., reaction chambers) of the multi-compartmentalized container, and directly controls input to and aspiration from each of the microlocations (e.g., reaction chambers) of the multi-compartmentalized container. Accordingly, the positioning of the magnets desirably mirrors that of the array. Alternately, a magnet can control more than one microlocation. Under this scenario, the positioning of the magnets desirably can mirror that of the array, or, can be substantially different from that of the array. When permanent magnets are employed, these magnets can be controlled, for instance, by energizing or de-energizing the magnet, i.e., as with an electric charge. Electromagnets can be controlled, for instance, by switching the field of the applied current. Preferably the electromagnets are controlled by an appropriate power source, e.g., a DC power source capable of providing from about 0 to about 50 milliAmps per electromagnet coil. Desirably, each of the electromagnets is independently controlled, e.g., making use of a multiplexer to control signal inputs/outputs, and computer programming, such as are well known in the art.

In addition to the device configurations shown in **Figures 8** and **9**, the device can have the configuration depicted in **Figure 10**. Namely, in this device, the reaction chambers themselves are present inside of a housing, and in intimate association with the channels and

micropipettes or other tubing system, thus allowing a "microchip" type assembly. For instance, the reaction chambers actually can be the wells of a microtiter plate, or etched onto the surface of a glass slide, or other appropriate material. Of course, the type, number, and location of the input/output channels, and the number and the location of the magnetic particle positioning system with respect to the chambers can vary, all as previously described.

Accordingly, as described above, the present invention desirably provides a fluidic switching assembly (i.e., parallel or integrated) that combines pressure and magnetic elements to move liquids and particles from a microtiter plate into a detection system. In particular, desirably the method of the invention is carried out using a multi-compartmentalized container that is integrated onto a fluidic switching system that contains computer controlled magnetic particle position control. Sequential energizing of adjacent electromagnets can be used to move magnetic particles among different well and through the fluidic system.

While the foregoing discusses two approaches that optimally can be employed to increase throughput of this assay, other approaches and instrumentation such as would be convenient for a particular application of the method of the invention optionally can be employed instead of these (e.g., modification of the apparatus and method described in PCT International Application WO 98/38490, herein incorporated in its entirety by reference).

Alternative means of Effecting Microsphere/Magnetic Particle Binding and Capture

An important part of the present invention is the manner in which binding between the sample nucleic acid and the magnetic particle is effected. The foregoing description is directed toward use of streptavidin and biotin as means of binding. However, as previously indicated, there are alternative means of ensuring such binding. All that is required is that there be provided a "first binding partner" and a "second binding partner" which together constitute a binding pair. Of course, further according to the invention, the first and second binding partner are attached to their respective molecules (i.e., either a nucleic acid, or a magnetic particle) in some fashion, e.g., for instance, by a covalent interaction (e.g., chemical linkage and/or fusion), or by a noncovalent interaction. In particular, preferably a first binding partner is attached to a nucleic acid of interest, and a second binding partner is attached to the magnetic particle.

Namely a "binding pair" is comprised of (1) a first binding partner that is capable of specifically binding to a second binding partner, and (2) a second binding partner that is capable of specifically binding to a first binding partner. The binding pair thus optimally

binds so as to form a complex. A "complex" of the first and second binding partner is any interaction, e.g., covalent or noncovalent, between the first and second binding partner, and, preferably, is a noncovalent interaction. Preferably according to the invention, this complex is one that can be dissociated – not spontaneously, but only with application of the appropriate conditions.

The complex is brought about by any means of contacting the first binding partner with the second binding partner. Such "contacting" can be done by any means known to those skilled in the art, and described herein, by which the apparent touching or mutual tangency of the first and second binding partner can be effected.

In a preferred embodiment of the invention one of the binding partners preferably is an antibody (e.g., a polyclonal, monoclonal, bispecific, and/or single-chain antibody). Such an antibody includes, but is not limited to, immunoglobulin molecules and immunologically active portions of immunoglobulin molecules such as portions containing a paratope (i.e., an antigen binding site), such that the antibody comprises, for example, either intact immunoglobulin molecules or portions thereof, such as those known in the art as Fab, Fab', $F(ab')_2$ and F(v). The antibody can be, for example, a monoclonal antibody, a polyclonal antibody, a single-chain antibody (e.g., directly fused to an oligonucleotide, for instance), and a bispecific antibody (e.g., that in and of itself can be a bispecific molecule having one paratope directed to an epitope of a first binding partner, and another paratope directed to an epitope of a second binding partner).

Preferred antibodies according to the invention are, of course, those that are directed against the other binding partner that constitutes the binding pair. The antibody can be produced by any suitable technique, e.g., conventional techniques for preparing monoclonal, polyclonal, single-chain, and bispecific antibodies, as well as more current recombinant DNA techniques that are familiar to those skilled in the art. In particular, bispecific antibodies can be made by a variety of means, e.g., chemical techniques (see, e.g., Kranz et al., Proc. Natl. Acad. Sci., 78, 5807 (1981)), for instance, disulfide cleavage reformation of whole IgG or, preferably, $F(ab')_2$ fragments; fusions of more than one clone to form polyomas that produce immunoglobulins having more than one specificity (see, e.g., Segal et al., In Current Protocols in Immunology, Coligan et al. (eds.), vol. 1, 2.13.1-2.13.16 (John Wiley & Sons, Inc. (1995))); or by genetic engineering.

In particular an especially preferred antibody according to the invention is that directed against the agent digoxigenin. Preferably, digoxigenin is attached to the sample nucleic acid, and the anti-digoxigenin antibody is attached to the magnetic particles.

Of course, it further is possible according to the invention to carry out capture of the sample nucleic acid without use of magnetic particles. This can be done, for instance, by binding the second binding partner not to a magnetic particle, but instead, to a coated surface (e.g., slab, bead, or other). Use of such a surface coated with the second binding partner optimally provides for capture on the coated surface of the first binding entity associated with the nucleic acid. Such capture on a surface still allows for the nucleic acid to be assessed either while still attached, or, following its dissociation from the coated surface. As an example, in the case where the coated surface is a well, e.g., a well of a microtiter plate, the further analysis of the nucleic acid (e.g., PCR) can be done directly in the well. Other variations and manipulations would be apparent to one skilled in the art. Alternately, as with release from the magnetic particles as previously described, release can be effected with use of the appropriate denaturant, such as hot formamide.

Thus, the present invention further provides a method for analyzing the expression of a plurality of genes present in a sample, wherein the method comprises the steps:

(a) isolating from the sample the RNA produced by the genes;

(b) reverse transcribing the RNA into cDNA;

(c) attaching a first binding partner to the cDNA to obtain a first binding partner-conjugated cDNA;

(d) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to the cDNA of interest, wherein the probes are modified to allow end-attachment to labeled microspheres:

(e) obtaining a plurality of microspheres of differing types, with microsphere type being determined by the combination to provide a unique identification of the microsphere of:

(i) microsphere diameter, the microsphere being spherical: and

(ii) microsphere color:

(f) obtaining probe/microspheres by attaching the microspheres to the modified single-stranded DNA probes;

(g) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each the separate microlocation;

(g) hybridizing the first binding partner-conjugated cDNA to the probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/cDNA hybrids:

(h) washing the probe/microsphere/cDNA hybrids;

(i) obtaining magnetic particles that are coated with a second binding partner and incubating the coated particles with the probe/microsphere/cDNA hybrids under conditions such that they bind to the first binding partner-conjugated cDNA present in the probe/microsphere/cDNA hybrids, and second binding partner-coated magnetic particles/probe/microsphere/cDNA hybrids are obtained;

(j) using a magnet to capture the second binding partner-coated magnetic particles/probe/microsphere/cDNA hybrids; and

(k) assessing the second binding partner-coated magnetic particles/probe/microsphere/cDNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the expression of a plurality of genes present in the sample.

Also, this method desirably can be employed for nucleic acid hybridization, in this instance, preferably comprising the steps of:

(a) isolating a sample to be tested for a DNA of interest;

(b) attaching a first binding partner to the DNA to obtain a first binding partner-conjugated DNA;

(c) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to the DNA of interest, wherein the probes are modified to allow end-attachment to labeled microspheres;

(d) obtaining a plurality of microspheres of differing types, with microsphere type being determined by the combination to provide a unique identification of the microsphere of:

(i) microsphere diameter, the microsphere being spherical; and

(ii) microsphere color;

(e) obtaining probe/microspheres by attaching the microspheres to the modified single-stranded DNA probes ;

(f) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each of the separate microlocation;

(g) hybridizing the first binding partner-conjugated DNA to the probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/DNA hybrids;

(h) washing the probe/microsphere/DNA hybrids;

(i) magnetic particles that are coated with a second binding partner and incubating the particles with the probe/microsphere/DNA hybrids under conditions such that they bind to the first binding partner-conjugated DNA present in the probe/microsphere/DNA hybrids. and second binding partner-coated magnetic particles/probe/microsphere/DNA hybrids are obtained;

(j) using a magnet to capture the second binding partner-coated magnetic particles/probe/microsphere/DNA hybrids; and

(k) assessing the second binding partner-coated magnetic particles/probe/microsphere/DNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the representation of the DNA present in the sample.

Preferably these methods are carried out where the first binding partner is digoxigenin, and the second binding partner is an anti-digoxigenin antibody. However, other preferred first and second binding partners similarly can be employed.

The following example further illustrates the present invention but, of course, should not be construed as in any way limiting its scope.

Example 1: Theoretical Underpinnings of the Invention

This Example describes some of the theoretical underpinnings of the invention.

Using distinct fluorescent dyes (and/or luminescent dyes) either alone or in combinations of two dyes per microsphere, the number of available array elements (number of genes which theoretically could be probed in one experiment) can be described by the following equation:

$$T_a = (L)(S)[0.5(X) + 0.5(X^2)]$$

where:

T_a = Total Array Elements

L = Number of Microlocations (e.g., number of wells)

S = Number of microsphere sizes

X = Number of Individual Fluorescent Colors used alone or in dual combination

For X, the color can be due to use of distinctly different dyes (i.e., different color dyes), or, use of the same dye at distinctly different intensity levels (which analytical instrumentation can easily distinguish). Using the above equation, as an example, with a 384-well plate, 15 available colors, and 3 microsphere sizes, the Total Array Elements =138,240. Of course other equations also could be used to describe the arrays according to the invention.

A gene with a relatively high level of expression leads to more "copies" of biotinylated DNA (or cDNA) and thus a greater number of specifically-identifiable captured microspheres.

Software can be employed to automatically associate microlocation position, fluorescence/luminescence, and size with particular probe and output number of microspheres (measurement of level of expression) for that particular probe (expression for that particular gene). For instance:

PROBE	GENE	NUMBER MICROSPHERES DETECTED
AG418	BRCA1	8 9 7 6
BCV54	MDR1	4 3 1
PVT39	UVrex1	2

Example 2: Advantages of the Invention and Illustrative Uses

This Example describes some of the advantages of the invention, and some illustrative uses.

In the present invention, the concurrent use of microlocation position in a multi-well plate, several distinct fluorescent "colors", and several distinct microsphere sizes, are combined to produce arrays with extremely large numbers of array elements. The use of magnetic capture to selectively capture the hybridized microspheres is advantageously employed. The number of spheres of a given microlocation position, fluorescent color, and size are analyzed to reflect the abundance of the nucleic acid species for which that element was designed to probe. These elements as well as potentially others, act in combination to produce the present unexpected and novel invention.

The improvements over the prior art that are accomplished by this invention include, but are not limited to, the following. First and foremost, hybridization is carried out in solution, i.e., as opposed to in the solid phase, which generally is less useful for analyzing rare sequences. AMA/FAMA provides a novel system with which to analyze the expression and/or representation of extremely large numbers of genes in a single experiment. Thus, the

level of mRNA for a specific gene in the sample is reflected in quantal, digital events, with genes of low abundance mRNA yielding a signal of equal strength and intensity as that of high abundance genes. Low abundance mRNAs simply generate fewer event numbers compared to high abundance mRNAs. Furthermore, the measured sample can be recovered following analysis for further experimentation. Still further improvements would be apparent to those skilled in the arts.

AMA/FAMA provides a methodology enabling analysis of genetic information in an array-type system, in which even genes with low level of expression produce strong, discrete signals. Typical analyses could use the flow cytometric equipment which is already commonplace in most hospitals and research laboratories. AMA (e.g., FAMA) also provides numerous other advantages.

Potential applications for the invention include, but are not limited to, basic research involving gene expression and clinical analysis of the expression of clinically relevant genes, for example, the expression of genes conferring resistance to anti-cancer drugs in a human tumor. Other applications are such as would be apparent or obvious to one skilled in the art. There are no restrictions or limitations as to the use of the methodology described herein. Basically, the invention can be employed in any experiment that requires hybridization-based analysis of gene expression (i.e., where RNA is the nucleic acid species of interest in a sample), or hybridization-based analysis of gene representation (i.e., where DNA, such as genomic DNA, is the nucleic acid species of interest in a sample).

Accordingly, the present invention provides not only a novel means of hybridization, but also, an improvement over other hybridization methods. In such an improvement of an array-based hybridization method for analyzing gene expression, the improvement preferably comprises the use of fluorescent or luminescent microspheres for attachment to single stranded probes directed to genes of interest present in the sample to obtain probe/microspheres, the hybridization method carried out by placing each probe/microsphere in a separate, specific microlocation position of a multi-compartmentalized container and performing all subsequent steps in each separate microlocation, wherein the combination of color, size, and microlocation position of probe/microspheres uniquely identify those probe/microspheres as specifically hybridizing to RNA for a specific gene, and the level of signal detected corresponds to the level of expression of the gene.

All of the references cited herein are hereby incorporated in their entireties by reference.

While this invention has been described with an emphasis upon a preferred embodiment, it will be obvious to those of ordinary skill in the art that variations in the preferred composition and method may be used and that it is intended that the invention may be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the following claims.

WHAT IS CLAIMED IS:

1. A method for analyzing the expression of a plurality of genes present in a sample comprising the steps of:

(a) isolating from said sample the RNA produced by said genes;

(b) reverse transcribing said RNA into cDNA in the presence of one or more biotin-conjugated nucleotide(s) to obtain biotin-conjugated cDNA;

(c) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to said cDNA of interest, wherein said probes are modified to allow end-attachment to labeled microspheres;

(d) obtaining a plurality of microspheres of differing types, with microsphere type being determined by the combination to provide a unique identification of said microsphere of:

(i) microsphere diameter, the microsphere being spherical; and

(ii) microsphere color;

(e) obtaining probe/microspheres by attaching said microspheres to said modified single-stranded DNA probes;

(f) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each said separate microlocation;

(g) hybridizing said biotin-conjugated cDNA to said probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/cDNA hybrids;

(h) washing said probe/microsphere/cDNA hybrids;

(i) obtaining streptavidin-conjugated magnetic particles and incubating said particles with said probe/microsphere/cDNA hybrids under conditions such that they bind to said biotin-conjugated cDNA present in said probe/microsphere/cDNA hybrids, and streptavidin-conjugated magnetic particles/probe/microsphere/cDNA hybrids are obtained;

(j) using a magnet to capture said streptavidin-conjugated magnetic particles/probe/microsphere/cDNA hybrids; and

(k) assessing said streptavidin-conjugated magnetic particles/probe/microsphere/cDNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe

employed per microlocation to analyze the expression of a plurality of genes present in the sample.

2. The method of claim 1, wherein microsphere color is determined by the presence within the microsphere of one or more dyes, with each dye being of a different color, different intensity, or different color and different intensity.

3. The method of claim 1, wherein said dye is a fluorescent or luminescent dye.

4. The method of claim 1, which further comprises a step following step (b) wherein said cDNA is separated from said RNA.

5. The method claim 4, wherein said separation is carried out by:

(a) digesting said RNA; or

(b) separating said biotin-conjugated cDNA from said RNA by means of biotin incorporated in said biotin-conjugated cDNA.

6. The method of claim 1, which further comprises a step following step (j) wherein:

(i) said cDNA present in said hybrid is degraded, and the resulting microsphere is washed; or

(ii) said cDNA is recovered from said streptavidin-conjugated magnetic particles/probe/microsphere/cDNA hybrids.

7. The method of claim 1, wherein one probe is hybridized at each array microlocation.

8. The method of claim 1, wherein more than one probe is hybridized at each array microlocation.

9. The method of claim 8, wherein a different probe is employed for each array microlocation.

10. The method of claim 8, wherein from 3 to 10 different probes are hybridized at each array microlocation.

11. A method for nucleic acid hybridization comprising the steps of:

(a) isolating a sample to be tested for a DNA of interest;

(b) biotinylating said sample DNA to obtain biotin-conjugated DNA;

(c) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to said DNA of interest, wherein said probes are modified to allow end-attachment to labeled microspheres;

(d) obtaining a plurality of microspheres of differing types, with microsphere type being determined by the combination to provide a unique identification of said microsphere of:

(i) microsphere diameter, the microsphere being spherical; and

(ii) microsphere color;

(e) obtaining probe/microspheres by attaching said microspheres to said modified single-stranded DNA probes;

(f) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each of said separate microlocation;

(g) hybridizing said biotin-conjugated DNA to said probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/DNA hybrids;

(h) washing said probe/microsphere/DNA hybrids;

(i) obtaining streptavidin-conjugated magnetic particles and incubating said particles with said probe/microsphere/DNA hybrids under conditions such that they bind to said biotin-conjugated DNA present in said probe/microsphere/DNA hybrids, and streptavidin-conjugated magnetic particles/probe/microsphere/DNA hybrids are obtained;

(j) using a magnet to capture said streptavidin-conjugated magnetic particles/probe/microsphere/DNA hybrids; and

(k) assessing said streptavidin-conjugated magnetic particles/probe/microsphere/DNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the representation of said DNA present in the sample.

12. The method of claim 11, wherein microsphere color is determined by the presence within the microsphere of one or more dyes, with each dye being of a different color, different intensity, or different color and different intensity.

13. The method of claim 11, wherein said dye is a fluorescent or luminescent dye.

14. The method of claim 11, which further comprises a step following step (j) wherein:

(i) said DNA present in said hybrid is degraded, and the resulting microsphere is washed; or

(ii) said DNA is recovered from said streptavidin-conjugated magnetic particles/probe/microsphere/DNA hybrids.

15. The method of claim 11, wherein one probe is hybridized at each array microlocation.

16. The method of claim 11, wherein more than one probe is hybridized at each array microlocation.

17. The method of claim 16, wherein a different probe is employed for each array microlocation.

18. The method of claim 16, wherein from 3 to 10 different probes are hybridized at each array microlocation.

19. In an improvement of an array-based hybridization method for nucleic acids, said improvement comprising the use of fluorescent or luminescent microspheres for attachment to single stranded probes to obtain probe/microspheres, said probes directed to (i) genes of interest present in said sample, or (ii) cDNA reverse-transcribed from RNA produced by specific genes of interest present in said sample,

wherein the combination of color, and size of probe/microspheres uniquely identify those probe/microspheres as specifically hybridizing to:

(a) the specific gene itself, and the level of signal detected corresponds to the level of representation of said gene in said sample; or

(b) cDNA reverse-transcribed from RNA produced by a specific gene, and the level of signal detected corresponds to the level of expression of said gene.

20. In an improvement of an array-based hybridization method for nucleic acids, said improvement comprising the use of fluorescent or luminescent microspheres for attachment to single stranded probes to obtain probe/microspheres, said probes directed to (i) genes of interest present in said sample, or (ii) cDNA reverse-transcribed from RNA produced by specific genes of interest present in said sample,

said hybridization method is carried out by placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container and performing all subsequent steps in each separate microlocation,

wherein the combination of color, size, and position of microlocation of said probe/microspheres uniquely identify those probe/microspheres as specifically hybridizing to:

(a) the specific gene itself, and the level of signal detected corresponds to the level of representation of said gene in said sample; or

(b) cDNA reverse-transcribed from RNA produced by a specific gene, and the level of signal detected corresponds to the level of expression of said gene.

21. A fluidic switching assembly that uses electromagnets or permanent magnets to control movement of materials into and out of a multi-compartmentalized container.

22. The assembly of claim 21, wherein said electromagnets or permanent magnets are computer controlled.

23. The assembly of claim 21, which is connected to a detection device.

24. The assembly of claim 21, having the structure depicted in Figure 8.

25. The assembly of claim 24, further comprising one or more ports for input and/or output.

26. The assembly of claim 24, further comprising a magnetic particle positioning system positioned below the multi-compartmentalized container.

27. The assembly of claim 21, having the structure depicted in Figure 9.

28. The assembly of claim 27, further comprising one or more ports for input and/or output.

29. The assembly of claim 27, further comprising a magnetic particle positioning system positioned below the multi-compartmentalized container.

30. The assembly of claim 21, having the structure depicted in Figure 9.

31. The assembly of claim 30, further comprising one or more ports for input and/or output.

32. The assembly of claim 30, further comprising a magnetic particle positioning system positioned below the multi-compartmentalized container.

33. The method of claim 1, which further comprises use of a fluidic switching assembly that uses electromagnets or permanent magnets to control movement of materials into and out of said multi-compartmentalized container.

34. The method of claim 11, which further comprises use of a fluidic switching assembly that uses electromagnets or permanent magnets to control movement of materials into and out of said multi-compartmentalized container.

35. A method for analyzing the expression of a plurality of genes present in a sample comprising the steps of:

(a) isolating from said sample the RNA produced by said genes;

(b) reverse transcribing said RNA into cDNA;

(c) attaching a first binding partner to said cDNA to obtain a first binding partner-conjugated cDNA;

(d) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to said cDNA of interest, wherein said probes are modified to allow end-attachment to labeled microspheres;

(e) obtaining a plurality of microspheres of differing types, with microsphere type being determined by the combination to provide a unique identification of said microsphere of:

(i) microsphere diameter, the microsphere being spherical; and

(ii) microsphere color;

(f) obtaining probe/microspheres by attaching said microspheres to said modified single-stranded DNA probes;

(g) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each said separate microlocation;

(g) hybridizing said first binding partner-conjugated cDNA to said probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/cDNA hybrids;

(h) washing said probe/microsphere/cDNA hybrids;

(i) obtaining magnetic particles that are coated with a second binding partner and incubating said coated particles with said probe/microsphere/cDNA hybrids under conditions such that they bind to said first binding partner-conjugated cDNA present in said probe/microsphere/cDNA hybrids, and second binding partner-coated magnetic particles/probe/microsphere/cDNA hybrids are obtained;

(j) using a magnet to capture said second binding partner-coated magnetic particles/probe/microsphere/cDNA hybrids; and

(k) assessing said second binding partner-coated magnetic particles/probe/microsphere/cDNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the expression of a plurality of genes present in the sample.

36. The method of claim 35, wherein said first binding partner is digoxigenin, and said second binding partner is an anti-digoxigenin antibody.

37. A method for nucleic acid hybridization comprising the steps of:

(a) isolating a sample to be tested for a DNA of interest;

(b) attaching a first binding partner to said DNA to obtain a first binding partner-conjugated DNA;

(c) obtaining a plurality of single-stranded DNA probes that are capable of specific hybridization to said DNA of interest, wherein said probes are modified to allow end-attachment to labeled microspheres;

(d) obtaining a plurality of microspheres of differing types, with microsphere type being determined by the combination to provide a unique identification of said microsphere of:

(i) microsphere diameter, the microsphere being spherical; and

(ii) microsphere color;

(e) obtaining probe/microspheres by attaching said microspheres to said modified single-stranded DNA probes ;

(f) placing each probe/microsphere in a separate, specific microlocation of a multi-compartmentalized container so as to generate an array, and performing all subsequent steps in each of said separate microlocation;

(g) hybridizing said first binding partner-conjugated DNA to said probe/microspheres under conditions sufficient to effect specific hybridization and obtain probe/microsphere/DNA hybrids;

(h) washing said probe/microsphere/DNA hybrids;

(i) magnetic particles that are coated with a second binding partner and incubating said particles with said probe/microsphere/DNA hybrids under conditions such that they bind to said first binding partner-conjugated DNA present in said probe/microsphere/DNA hybrids, and second binding partner-coated magnetic particles/probe/microsphere/DNA hybrids are obtained;

(j) using a magnet to capture said second binding partner-coated magnetic particles/probe/microsphere/DNA hybrids; and

(k) assessing said second binding partner-coated magnetic particles/probe/microsphere/DNA hybrids to determine the number and type of microsphere captured per microlocation, and considering this information along with the type of probe employed per microlocation to analyze the representation of said DNA present in the sample.

38. The method of claim 37, wherein said first binding partner is digoxigenin, and said second binding partner is an anti-digoxigenin antibody.

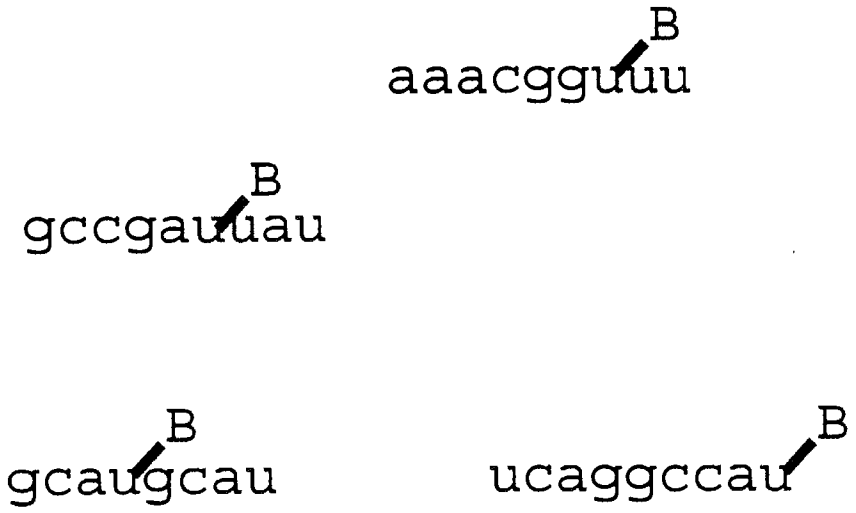


Figure 1

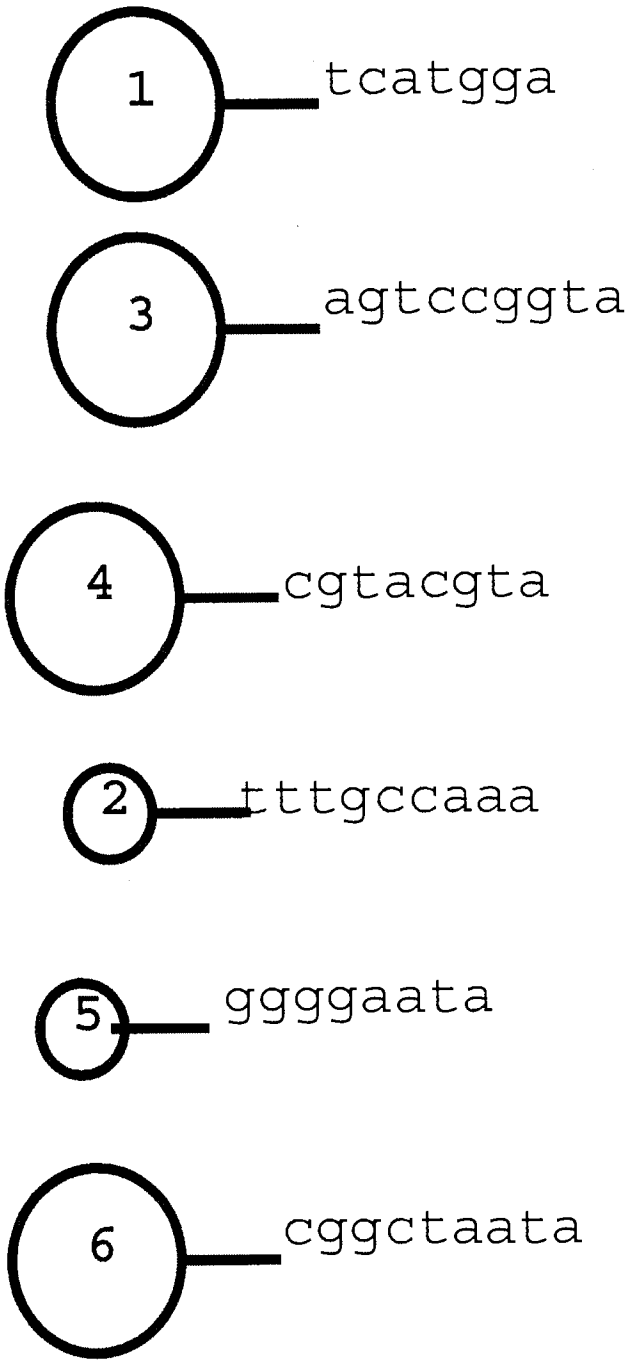

1
tcatgga
3
agtccggta
4
cgtacgta
2
tttgccaaa
5
ggggaata
6
cggctaata

Figure 2

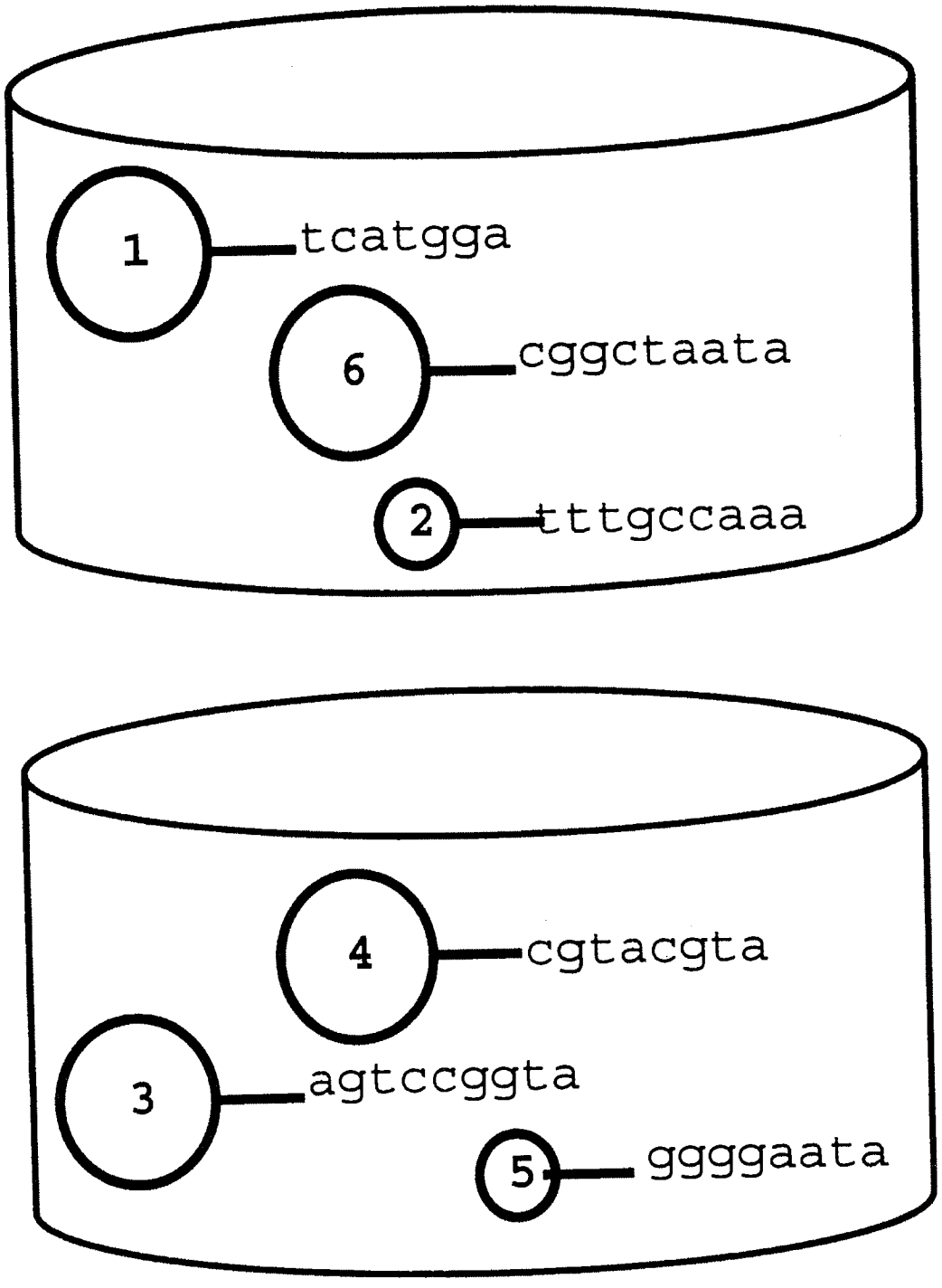

1
tcatgga
6
cggctaata
2
tttgccaaa
4
cgtacgta
3
agtccggta
5
ggggaata

Figure 3

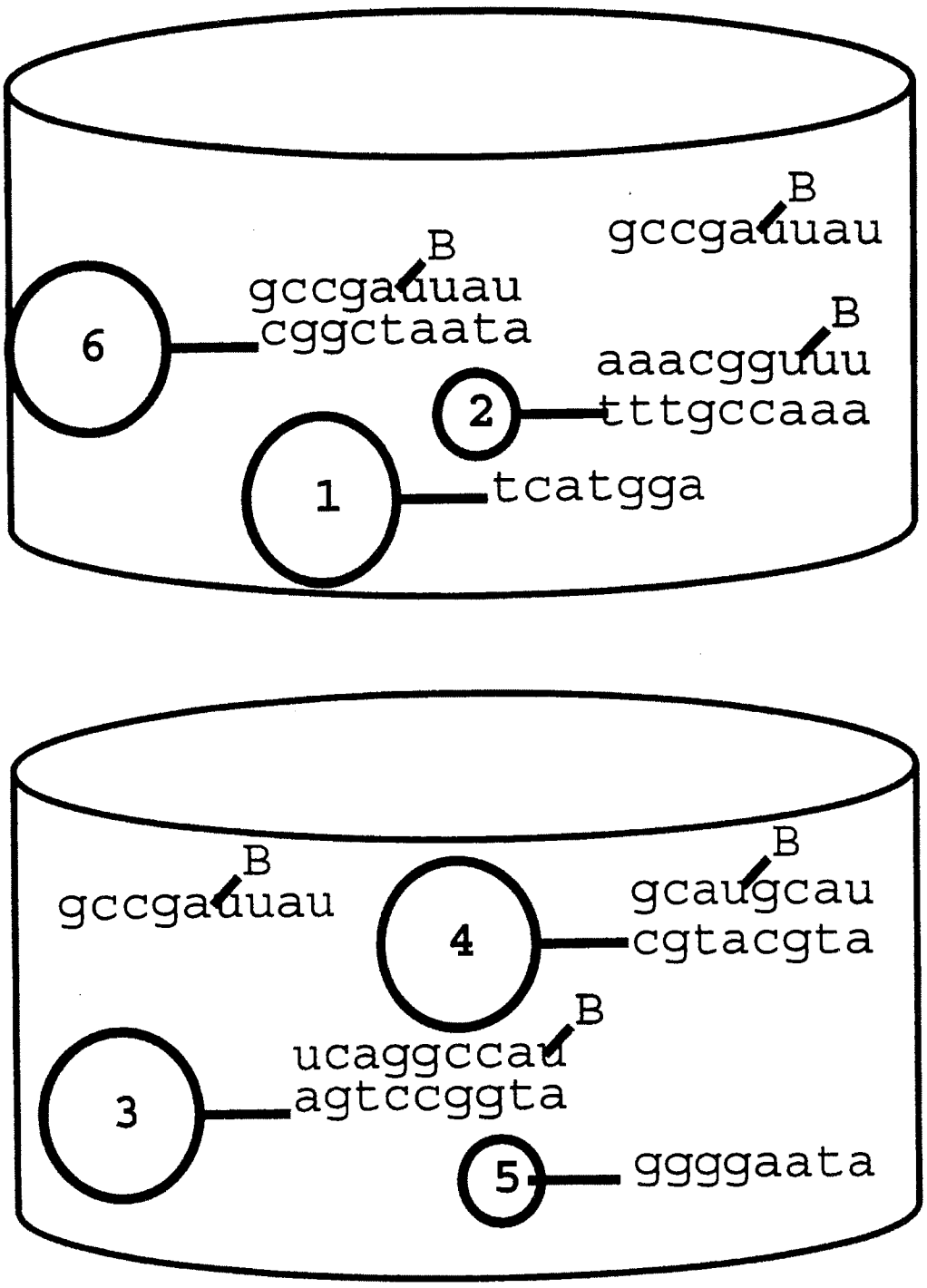

B
gccgauuau
B
gccgauuau
cggctaata
6
B
aaacgguuu
2
tttgccaaa
1
tcatgga
B
gccgauuau
B
gcaugcau
4
cgtacgta
B
ucaggccau
3
agtccggta
5
ggggaata

Figure 4

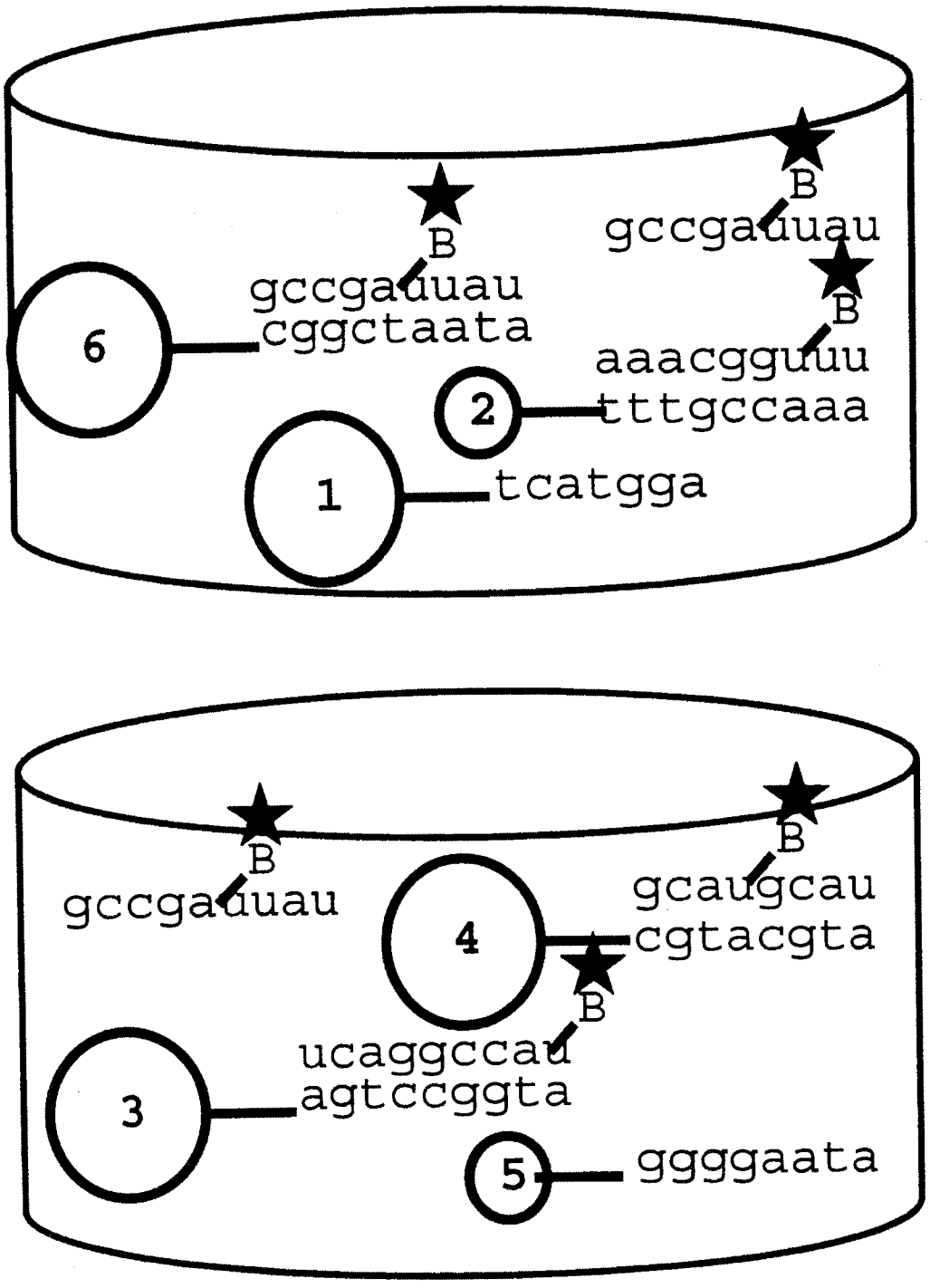

B
gccgaduuau
B
gccgaduuau
cggctaata
6
B
aaacgguuu
2
tttgccaaa
1
tcatgga
B
gccgaduuau
B
gcaugcau
4
cgtacgta
B
ucaggccau
3
agtccggta
5
ggggaata

Figure 5

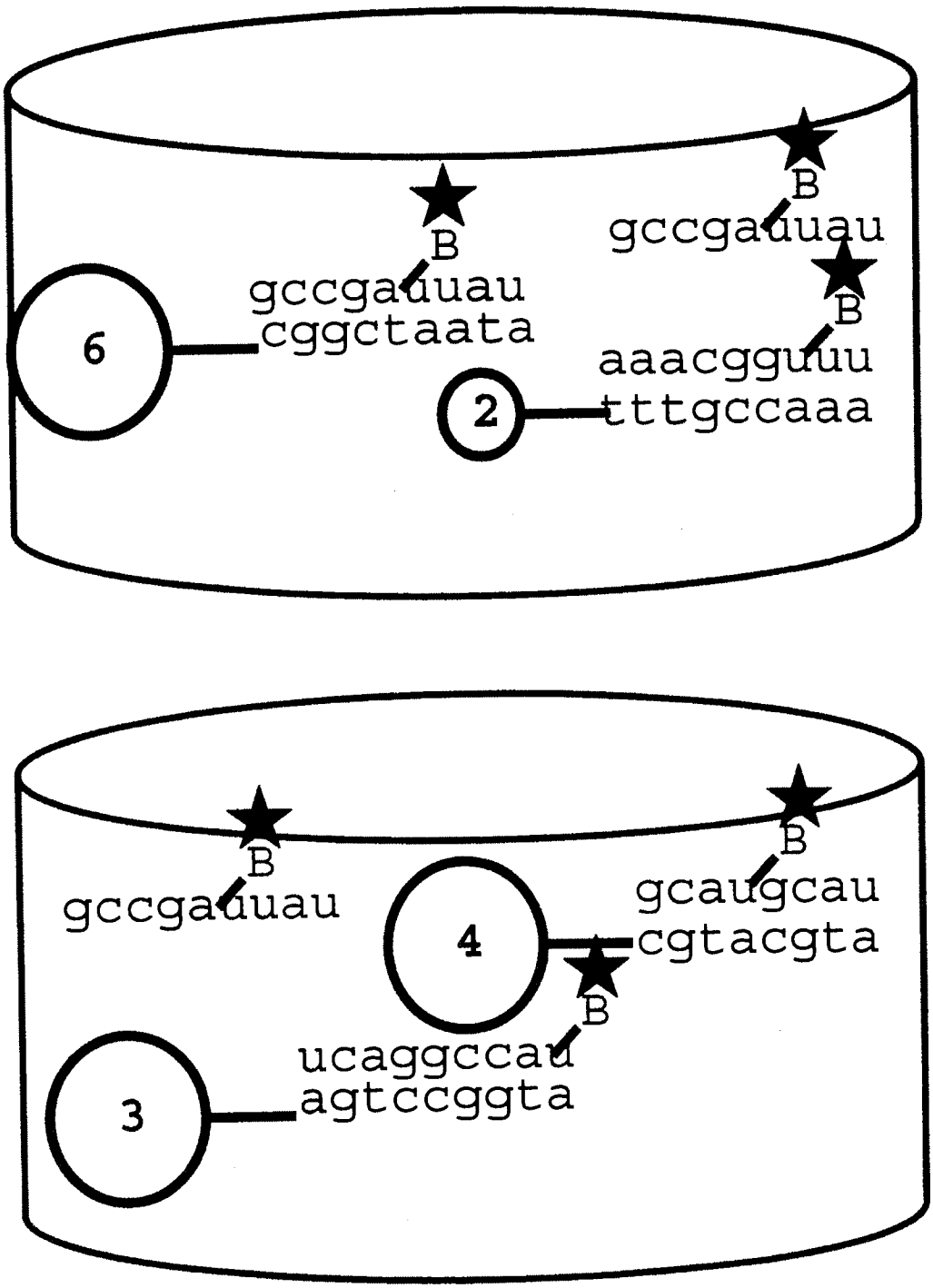

B
gccgauuau
B
gccgauuau
cggctaata
6
B
aaacgguuu
2
tttgccaaa
B
gccgauuau
B
gcaugcau
4
cgtacgta
B
ucaggccau
3
agtccggta

Figure 6

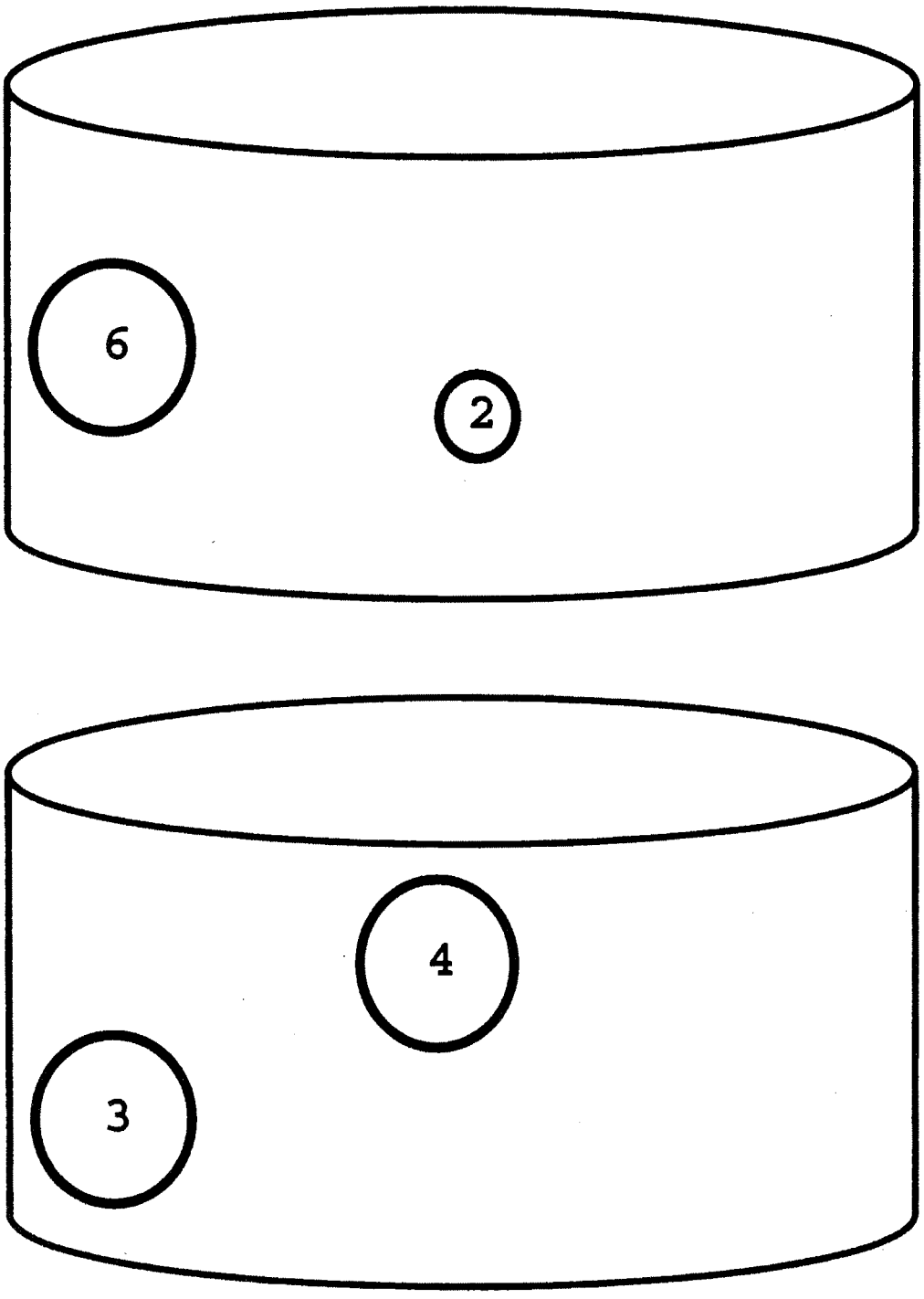


Figure 7

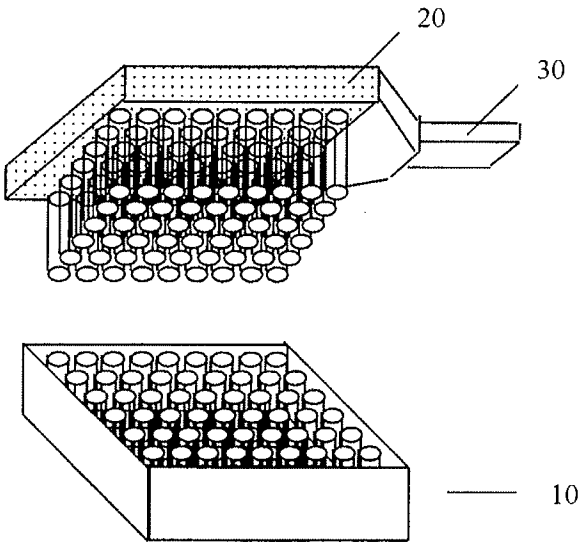


Figure 8

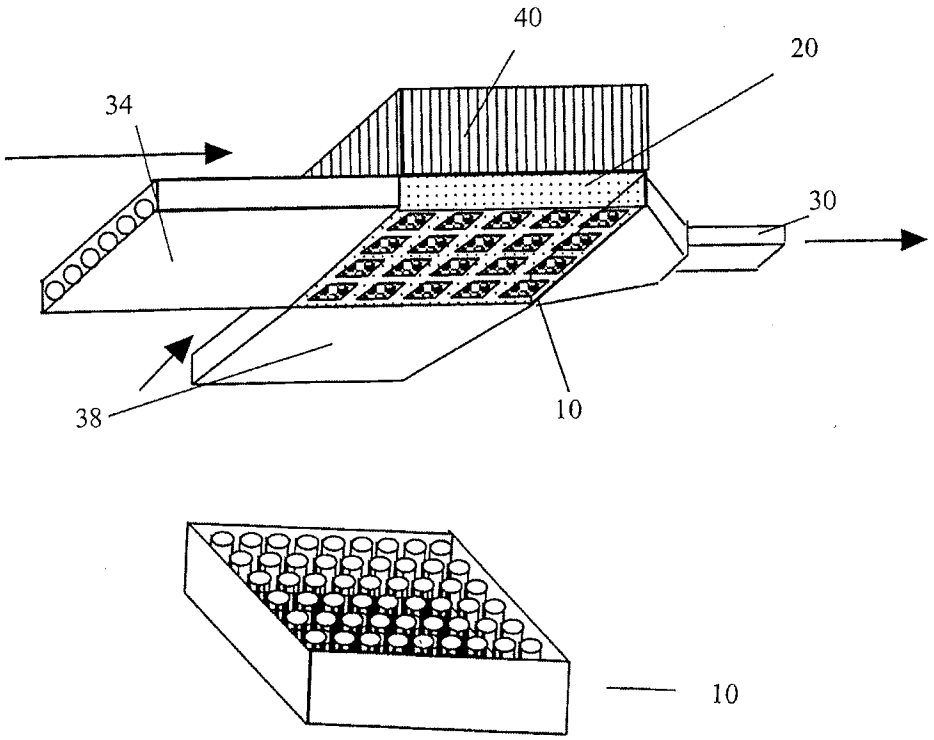


Figure 9

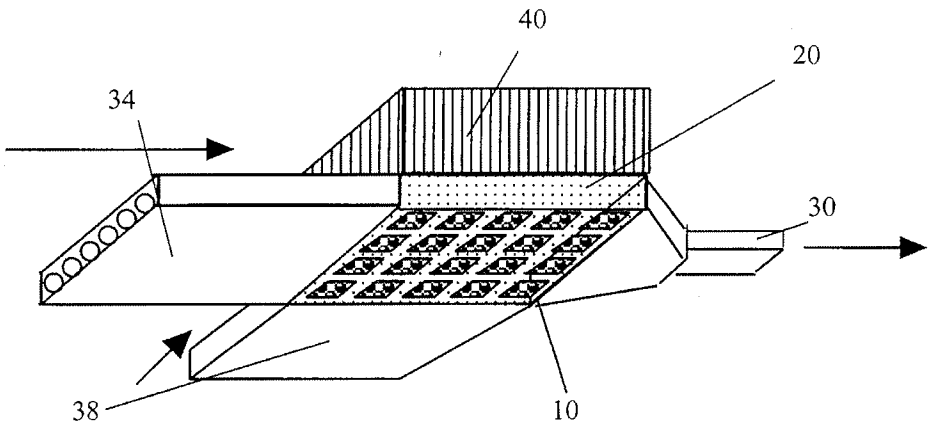


Figure 10

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 August 2012 (09.08.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/106385 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2012/023411

(22) International Filing Date:
31 January 2012 (31.01.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/437,854 31 January 2011 (31.01.2011) US
61/444,067 17 February 2011 (17.02.2011) US

(71) Applicant *(for all designated States except US)*: **APPRISE BIO, INC.** [US/US]; 248 Upper Terrace, San Francisco, CA 94117 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **NOLAN, Gary, P.** [US/US]; 248 Upper Terrace, San Francisco, CA 94117 (US).

(74) Agents: **WILLIS, Michael A.** et al.; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, CA 94304-1050 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: METHODS OF IDENTIFYING MULTIPLE EPITOPES IN CELLS

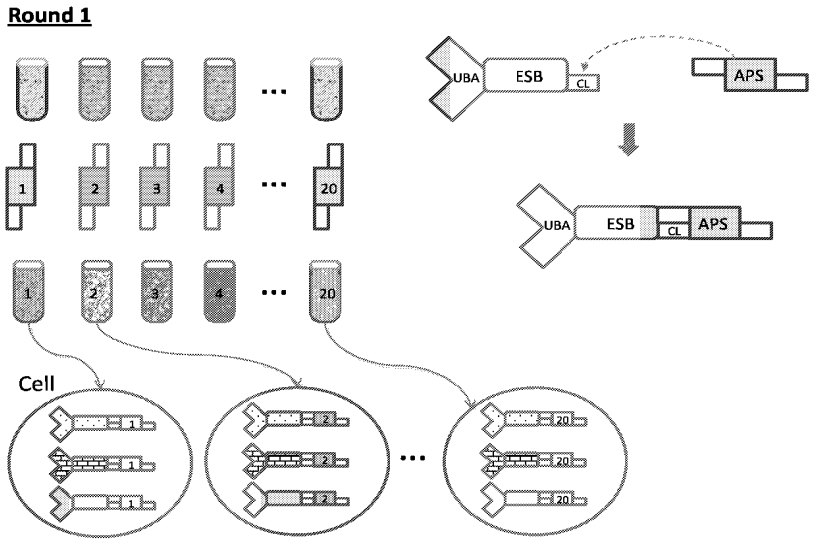


Figure 4

(57) **Abstract**: The invention provides methods, compositions, kits and devices for the detection of target molecules. In some embodiments, the invention allows for multiplexed target molecule detection.

WO 2012/106385 A2

METHODS OF IDENTIFYING MULTIPLE EPITOPES IN CELLS

CROSS-REFERENCE

[0001] This application claims priority to the following co-pending patent application: application Serial No. 61/437,854 filed January 31, 2011; which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0002] Although all cells in the human body contain the same genetic material, the same genes are not active in all of those cells. Alterations in gene expression patterns can have profound effects on biological functions. Furthermore, understanding the dynamics and the regulation of gene products (proteins), their variants, and interacting partners is essential in understanding, for example, the mechanisms behind genetic/and environmentally induced disorders or the influences of drug mediated therapies. This understanding can potentially become the underlying foundation for further clinical and diagnostic analyses. Therefore, identifying and quantifying the expression and regulation of genes and/or their products in cells can aid the discovery of new therapeutic and diagnostic targets.

[0003] Critical to these studies is the ability to qualitatively determine gene expression and specific variants of whole proteins (e.g., splice variants, point mutations, post-translationally modified versions, and environmentally/therapeutically-induced modifications) and the ability to view their quantitative modulation. Moreover, it is becoming increasing important to perform these analyses from not just one, but multiple target molecules in a cell. The methods available to date still require significant amounts of biological samples or will not provide cell specific information. Additionally, there are limited methods of multiplexed protein measurement technologies due to the additional challenges inherent in protein samples.

[0004] Thus, there exists a need for accurate and sensitive detection, identification and quantification of target molecules in every cell of a complex cell population and to retain cell specific information regarding that target molecule.

SUMMARY OF THE INVENTION

[0005] The invention relates generally to the field of detection, identification, and quantification of target molecules in a sample. The present invention relates in part to the detection, identification, and quantification of individual target molecules in single cells of a complex cell population while retaining cell specific information regarding that target molecule.

[0006] In some embodiments, the invention relates to methods for identifying whether a plurality of targets are in a plurality of cells comprising: binding to the targets a plurality of tags, wherein a tag comprises a code that represents a) the target identity and b) the identity of the cell in which tag is binding. In some embodiments, individual cell separation or isolation is unnecessary for the binding step. In some embodiments, tags comprise building blocks that are directly or indirectly associated with each other, for example through covalent binding or by association through affinity. In some embodiments, tags are formed through polymerization of building blocks in place. In some embodiments, multiple building blocks are added in a step. In some embodiment, a single building block is added at each step. In some embodiments, the cell is alive. In some embodiments, the cell is lysed or fixed.

[0007] In some embodiments, the invention relates to methods for identifying a single cell associated with a target comprising: binding to the target a tag, wherein the tag comprises a code that represents a) the target, and b) the single cell; wherein the during the binding the single cell is not isolated from a population of cells, and wherein the code that represents the single cell is unknown before the binding.

[0008] In some embodiments, the invention relates to methods for identifying a single cell associated with a target comprising: binding to the target a tag, wherein the tag comprises a code that represents a) the target, and b) the single cell; wherein the during the binding the single cell is isolated from a population of cells, and wherein the code that represents the single cell is unknown before the binding.

[0009] In some embodiments, the invention further comprises detecting the code, wherein individual cell separation or isolation is unnecessary for the detecting step. In some embodiments, each target is a protein or a nucleic acid. In some embodiments, the tag is a nucleic acid or a polypeptide. In some embodiments, the tag is comprises a series monomeric subunits that comprise a decipherable code. In some embodiments, the tag is a coded molecular constituent that can be decoded. In some embodiments, the tag comprises a combination of parts that can be decoded to determine the nature of the tag.

[0010] In some embodiments, the tag comprises a UBA. In some embodiments, the UBA is specific for one of the targets. In some embodiments, the tag comprises a UBA. In some embodiments, the UBA comprises an antibody. In some embodiments, the tag comprises a ESB. In some embodiments, the ESB comprises a common linker (CL). In some embodiments, the ESB codes the target identity. In some embodiments, the ESB comprises a nucleic acid. In some embodiments, the tag comprises an APS. In some embodiments, the APS is detectable a

detectably distinct coding unit. In some embodiments, during the binding step multiple APSs are added to the tag in an ordered manner during successive rounds of split pool synthesis. In some embodiments, the tag comprises at least 10 APSs. In some embodiments, the APS comprises a nucleic acid. In some embodiments, the tag comprises multiple APSs, an ESB, and a UBA linked by ligation. In some embodiments, multiple APSs, the ESB, and/or the UBA is capable of being linked Click chemistry. In some embodiments, the APS or the ESB comprises an amplification primer binding region. In some embodiments, the UBA, ESB, or APS is templatable. In some embodiment, the UBA, ESB, or APS is of a different discernable constituent (GPN: meaning one part of the code can be a nucleic acid, another can be a polypeptide, another can be a small molecule, etc.).

[0011] In some embodiments, the invention relates to compositions comprising: a) a first target molecule, b) a first unique binding agent (UBA) specific for the first target molecule, c) a first linkable UBA-dependent epitope specific barcode (ESB), and d) a plurality of ordered assayable polymer subunit (APS), wherein the order of APSs is detectable. In some embodiments, the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid.

[0012] In some embodiments, the invention relates to compositions comprising a population of particles each comprising at least a first target molecule, wherein the first target molecule is associated with: a) a first unique binding agent (UBA) specific for the first target molecule, b) a first linkable UBA-dependent epitope specific barcode (ESB), and c) a first plurality of ordered assayable polymer subunits (APS), wherein the plurality of ordered APSs associated with the first target molecule of a first particle in the population is detectably different than the plurality of ordered APSs associated with the first target molecule of a second particle in the population.

[0013] In some embodiments, the plurality of ordered APSs comprises 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 APSs. In some embodiments, the plurality of ordered APSs comprises more than 20 APSs. In some embodiments, the APSs are templatable. In some embodiments, at least one discrete particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead. In some embodiments, the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid,

and a phospholipid. In some embodiments, the first ESB comprises a first common linker (CL). In some embodiments, said first target molecule is directly bound to said first UBA and said first ESB is directly bound to said first UBA. In some embodiments, the plurality of ordered APS is formed by stepwise addition of APSs in separate rounds. In some embodiments, the APS added on each round is linked to the first complex. In some embodiments, the linking is in order of rounds. In some embodiments, the linking is performed through binding affinity. In some embodiments, the linking of an APS, an ESB, or a UBA is performed using chemical methods. In some embodiments, the chemical method comprises Click chemistry. In some embodiments, the linking is performed in the presence of Cu^{I}. In some embodiments, a UBA, an APS, or an ESB comprises nucleic acids. Some embodiments further comprise a first linking oligonucleotide comprising a first and a second complementary region to two components selected from a UBA, an APS, and an ESB. In some embodiments, a UBA, an APS, or an ESB is linked using a linking oligonucleotide comprising the first and the second complementary region to two components selected from a UBA, an APS, and an ESB. Some embodiments further comprise a second linking oligonucleotide comprising a third and a fourth complementary region to two components selected from a UBA, an APS, and an ESB. In some embodiments, a UBA, an APS, or an ESB is linked using a linking oligonucleotide comprising the third and the fourth complementary region to two components selected from a UBA, an APS, and an ESB. In some embodiments, the second and fourth complementary regions are identical. In some embodiments, the second and fourth complementary regions are identical. In some embodiments, the first or second complementary region is shared between two APSs within the plurality of APSs. In some embodiments, the linking is performed by ligation. In some embodiments, the linking oligonucleotide comprises a subcode encoding the origin of the APS or the ESB. In some embodiments, the APS has a subcode encoding the origin of the APS. In some embodiments, the ESB has a subcode encoding the origin of the ESB. In some embodiments, an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag. In some embodiments, unique counter tag is detectable. In some embodiments, the ESB is covalently linked to the linking oligonucleotide. In some embodiments, an APS or an ESB comprises an amplification primer binding region. In some embodiments, the APSs and ESB, when linked, are capable of encoding a secondary product. In some embodiments, the secondary product is an RNA or a peptide. In some embodiments, the APSs and ESB, when linked, comprises a polymerase start site. In some embodiments, the peptide comprises an affinity tag. In some embodiments, the affinity tag is a His-tag. In some embodiments, the UBA, ESB, or APS is templatable. In some embodiments, the composition further comprises a probe. In some embodiments, the probe is

attached to a surface. In some embodiments, the surface comprises an array. In some embodiments, the surface comprises a bead. In some embodiments, the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid. In some embodiments, the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits. In some embodiments, said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight.

[0014] In some embodiments, the invention relates to kits for labeling a target molecule of a cell in a population of cells with a cell origination barcode, comprising a) n sets of m assayable polymer subunits (APSs) each comprising a distinct package of information; wherein the packages of information are capable of being linked in an ordered fashion; b) a target molecule-specific unique binding agent (UBA).

[0015] In some embodiments, the invention relates to kits for labeling a target molecule of a cell in a population of cells with a cell origination barcode, comprising a) n sets of m assayable polymer subunits (APSs) each comprising a distinct package of information; wherein the packages of information are capable of being linked in an ordered fashion; b) a plurality of target molecule-specific unique binding agents (UBA) each linked with a UBA-specific epitope specific barcode (ESB).

[0016] In some embodiments, the invention relates to kits for labeling a target molecule of a cell in a population of cells with a cell origination barcode, comprising a) n sets of m assayable polymer subunits (APSs) each comprising a distinct package of information; wherein the packages of information are capable of being linked in an ordered fashion; b) a plurality of target molecule-specific unique binding agents (UBA); c) a plurality of UBA-specific epitope specific barcode (ESB), wherein each ESB is capable of linking with a designated UBA.

[0017] In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10. In some embodiments, m is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some embodiments, n is greater than 10. In some embodiments, m is greater than 20. In some embodiments, the first ESB comprises a first common linker (CL). In some embodiments, said ESB is capable of directly binding to said UBA. In some embodiments, said UBA is capable of directly binding to the target molecule. In some embodiments, at least two of the assayable polymer subunit (APS) sets are identical. In some embodiments, the APSs in a first set is linkable to the APSs in a second set. In some embodiments, the APSs in a first set is further linkable to the APSs in a second set in an ordered fashion. In some embodiments, an APS, an ESB, or a UBA is capable of being

linked using chemical methods. In some embodiments, the chemical method comprises Click chemistry. In some embodiments, the presence of Cu^{I} is required for linkage. In some embodiments, the kit components can assemble through affinity binding. In some embodiments, a UBA, an APS, or an ESB comprises nucleic acids. In some embodiments, the kits further comprise a first linking oligonucleotide comprising a first and a second complementary region to two components selected from a UBA, an APS, and an ESB. In some embodiments, the kits further comprise a second linking oligonucleotide comprising a third and a fourth complementary region to two components selected from a UBA, an APS, and an ESB. In some embodiments, the first and third complementary regions are identical. In some embodiments, the second and fourth complementary regions are identical. In some embodiments, an APS, an ESB, or a UBA is capable of being linked by ligation. In some embodiments, the linking oligonucleotide comprises a subcode encoding the original set of the APS or the ESB. In some embodiments, the APS has a subcode encoding the origin population of the APS. In some embodiments, the ESB has a subcode encoding the origin population of the ESB. In some embodiments, an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag. In some embodiments, the unique counter tag is detectable. In some embodiments, the ESB is covalently linked to the linking oligonucleotide. In some embodiments, an APS or an ESB comprises an amplification primer binding region. In some embodiments, the APSs and ESB, when linked, are capable of encoding a secondary product. In some embodiments, the secondary product is an RNA or a peptide. In some embodiments, the APSs and ESB, when linked, comprises a polymerase start site. In some embodiments, the peptide comprises an affinity tag. In some embodiments, the affinity tag is a His-tag. In some embodiments, the UBA, ESB, or APS is templatable. In some embodiments, the kit further comprises a probe. In some embodiments, the probe is attached to a surface. In some embodiments, the surface comprises an array. In some embodiments, the surface comprises a bead. In some embodiments, the plurality of UBAs comprises 2, 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600, 700, 800, 900, 1000 or more than 1000 UBAs. In some embodiments, the plurality of UBAs comprises up to 2000 UBAs. In some embodiments, the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid. In some embodiments, the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits. In some embodiments, said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight.

[0018] In some embodiments, the invention relates to methods for identifying target molecules sharing a common particle origin, comprising labeling a first plurality of targets of a first particle

in a population of x particles with a first origination barcode; and labeling a second plurality of targets of a second particle in a population of x particles with a second origination barcode; wherein each origination barcode comprises a set of n assayable polymer subunits (APS); wherein each of the n APSs in the first and second set of APSs is selected from a group comprising m different APSs; and wherein the first and second origination barcodes are detectably different from each other with a certainty of $c = 1- [(1 – 1/x)^(m^n)]$. In some embodiments, x is greater than 1,000,000. In some embodiments, c is greater than 99.9%. In some embodiments, c is greater than 99.99%. In some embodiments, c is greater than 99.999%. In some embodiments, c is greater than 99.9999%. In some embodiments, c is greater than 99.99999%. In some embodiments, n is 2, 3, 4, 5, 6, 7, 8, 9, or 10. In some embodiments, m is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some embodiments, n is greater than 10. In some embodiments, m is greater than 20.

[0019] In some embodiments, at least one discrete particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead. In some embodiments, the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid. In some embodiments, at least two groups comprising m different APSs are identical.

[0020] In some embodiments, n APSs are added in separate rounds. In some embodiments, the APSs of separate rounds are linked. In some embodiments, the linking is in order of rounds. In some embodiments, an appropriate n and/or m is selected based in a desired certainty level given a number of cells, x.

[0021] In some embodiments, the invention relates to methods of imparting a particle specific code to a component of a particle of a population of particles, the method comprising: linking a first ordered set of assayable polymer subunits (APS) to a first component of a first particle of a population of particles, wherein the order of the APSs is detectable. In some embodiments, the method further comprises detecting the first ordered set of APSs linked with the first component, thereby determining a particle origin of the first component. In some embodiments, the method further comprises linking a second ordered set of assayable polymer subunits (APS) to a second component of the first particle of a population of particles, wherein the order of the APSs is detectable. In some embodiments, the method further comprises detecting the second ordered set of APSs linked with the second component, thereby determining the particle origin of the second

component. In some embodiments, the first and the second ordered sets of APSs linked with the first and second components of the first particle are the same. In some embodiments, the method further comprises linking a third ordered set of assayable polymer subunits (APS) to a first component of second particle of a population of particles, wherein the order of the APSs is detectable. In some embodiments, the first ordered set of APSs linked with the first component of the first particle is different than the third ordered set of assayable polymer subunits linked with the first component of the second particle. In some embodiments, the method further comprises linking a component specific epitope specific barcode (ESB) to the first component. In some embodiments, the method further comprises linking a component specific ESB to the second component. In some embodiments, at least the particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead. In some embodiments, said at least one target molecule is directly bound to said first UBA and said ESB is directly bound to said UBA. In some embodiments, at least two of the assayable polymer subunit (APS) sets are identical. In some embodiments, each APS from the ordered set of APSs is linked to the first complex. In some embodiments, the linking is in order of rounds. In some embodiments, the UBA, ESB, or APS encodes a secondary product. In some embodiments, the secondary product is an RNA or a peptide. In some embodiments, the UBA, ESB, or APS is templatable. In some embodiments, the ESB further comprises a unique counter tag. In some embodiments, the quantity of the target molecule of the molecule is estimated using the counter tag. In some embodiments, the APS further comprises a round-specific subcode. In some embodiments, the detection further comprises determining the presence of an APS from a designated round. In some embodiments, detection is digital. In some embodiments, detection is indirect. In some embodiments, detecting comprises mass spectrometry. In some embodiments, detecting comprises nucleic acid sequencing. In some embodiments, detecting comprises peptide sequencing. In some embodiments, detecting comprises mass gel electrophoresis. In some embodiments, detecting comprises HPLC or other chromatographic separation. In some embodiments, detecting comprises detecting one or more signals associated with one or more individual APSs. In some embodiments, the signals are ordered. In some embodiments, detecting comprises using one or more probes. In some embodiments, the probe is attached to a surface. In some embodiments, the surface comprises an array. In some embodiments, the surface comprises a bead. In some embodiments, detecting comprises a separation. In some embodiments, the separation is multi-dimensional. In some embodiments, the separation resolves the first linkable UBA-dependent epitope specific barcode (ESB) from a second linkable UBA-dependent epitope specific barcode (ESB). In some embodiments, 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600,

700, 800, 900, 1000 or more than 1000 different target molecules are detected. In some embodiments, up to 2000 different target molecules are detected. In some embodiments, the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid. In some embodiments, the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits. In some embodiments, said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight. In some embodiments, the APSs are linked to through ligation or extension via polymerization. In some embodiments, a cell origination barcode (COB) is generated with the APSs from the ordered set of APSs. In some embodiments, each COB in said plurality of complexes has a detectable signal or sequence that distinguishes it from other COBs in said population of cells. In some embodiments, an APS, an ESB, or a UBA is linked using chemical methods. In some embodiments, the chemical method comprises Click chemistry. In some embodiments, the linking is performed in the presence of Cu^{I}. In some embodiments, a UBA, an APS, or an ESB comprises nucleic acids. In some embodiments, the linking of a UBA, an APS, or an ESB is performed using a linking oligonucleotide that comprises a first and a second complementary region to two components to be linked. In some embodiments, the first or second complementary region is shared between APSs within a population of APSs. In some embodiments, the first or second complementary region is distinct for two different round-specific sets of APSs. In some embodiments, the method further comprises ligation. In some embodiments, the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid. In some embodiments, the first ESB comprises a first common linker (CL). In some embodiments, the first ESB comprises a first common linker (CL). In some embodiments, an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag. In some embodiments, detection comprises detecting the unique counter tag. In some embodiments, the number of unique counter tags associated with a specific ESB is determined. In some embodiments, the number of detected unique counter tags relate to the initial quantity of the specific ESB. In some embodiments, the ESB is covalently linked to the linking oligonucleotide. In some embodiments, an APS or an ESB comprises an amplification primer binding region. In some embodiments, a COB encodes a peptide sequence. In some embodiments, the COB comprises a polymerase start site. In some embodiments, the peptide comprises an affinity tag. In some embodiments, the affinity tag is a His-tag.

[0022] In some embodiments, the invention relates to methods for detecting plurality of properties originating from a plurality of discrete particles, the method comprising: a) providing: i) a population of particles comprising at least a first target molecule; ii) a first unique binding agent (UBA) specific for the first target molecule; iii) a first linkable UBA-dependent epitope specific barcode (ESB); iv) a plurality of round-specific assayable polymer subunit (APS) sets, each set containing a plurality of APSs that are detectably distinct from each other; b) forming at least a first complex comprising said at least first target molecule, said first UBA probe, and said first ESB; c) performing n rounds of split pool synthesis, each round comprising; i) splitting the population of particles into m reaction volumes; ii) contacting one or more reaction volumes with an APS from the APS set specific for the round; iii) pooling two or more reaction volumes; d) detecting a plurality of properties from at least one particle from the population of particles; wherein at least one of the properties relate to a quantity or an identity for a target molecule associated with the particle.

[0023] In some embodiments, the invention relates to methods for detecting a plurality of properties originating from a plurality of discrete particles, the method comprising: a) providing: i) a population of particles comprising at least a first target molecule; ii) a first unique binding agent (UBA) specific for the first target molecule; iii) a first linkable UBA-dependent epitope specific barcode (ESB); and iv) a plurality of round-specific assayable polymer subunit (APS) sets, each set containing a plurality of APSs that are detectably distinct from each other; b) forming at least a first complex comprising said at least first target molecule, said first UBA probe, and said first ESB; c) performing n rounds of split pool synthesis, each round comprising: i) splitting the population of particles into m reaction volumes; ii) contacting one or more reaction volumes with an APS from the APS set specific for the round; and iii) pooling two or more reaction volumes; d) performing another round of split pool synthesis comprising steps c) i) and c) ii); e) detecting a plurality of properties from at least one particle from the population of particles; wherein at least one of the properties relate to a quantity or an identity for a target molecule associated with the particle.

[0024] In some embodiments, the split pool method is replaced by separation of particles, for example in microwells, or in microfluidic devices. In some embodiments, separated cells are labeled with cell origination barcodes. In some embodiments, cell origination barcodes are built in place by stepwise addition of building blocks.

[0025] In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some embodiments, n is more than 20. In some embodiments, m is different

between at least two rounds. In some embodiments, m is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some embodiments, m is more than 20. In some embodiments, at least one discrete particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead. In some embodiments, the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid. In some embodiments, the first ESB comprises a first common linker (CL). In some embodiments, said at least one target molecule is directly bound to said first UBA and said ESB is directly bound to said UBA. In some embodiments, at least two of the assayable polymer subunit (APS) sets are identical. In some embodiments, the APS added on each round is linked to the first complex. In some embodiments, the linking is in order of rounds. In some embodiments, the UBA, ESB, or APS encodes a secondary product. In some embodiments, the secondary product is an RNA or a peptide. In some embodiments, the UBA, ESB, or APS is templatable. In some embodiments, the ESB further comprises a unique counter tag. In some embodiments, the quantity of the target molecule of the molecule is estimated using the counter tag. In some embodiments, the APS further comprises a round-specific subcode. In some embodiments, the detection further comprises determining the presence of an APS from a designated round. In some embodiments, detection is digital. In some embodiments, detection is indirect. In some embodiments, detecting comprises mass spectrometry. In some embodiments, detecting comprises nucleic acid sequencing. In some embodiments, detecting comprises peptide sequencing. In some embodiments, detecting comprises detecting one or more signals associated with one or more individual APSs. In some embodiments, the signals are ordered. In some embodiments, detecting comprises using one or more probes. In some embodiments, the probe is attached to a surface. In some embodiments, the surface comprises an array. In some embodiments, the surface comprises a bead. In some embodiments, detecting comprises a separation. In some embodiments, the separation is multi-dimensional. In some embodiments, the separation resolves the first linkable UBA-dependent epitope specific barcode (ESB) from a second linkable UBA-dependent epitope specific barcode (ESB). In some embodiments, 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600, 700, 800, 900, 1000 or more than 1000 different target molecules are detected. In some embodiments, up to 2000 different target molecules are detected. In some embodiments, the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid. In some embodiments, the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits. In some embodiments, said APS

comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight. In some embodiments, the APSs are linked to through ligation or extension via polymerization. In some embodiments, a cell origination barcode (COB) is generated from APSs of round specific APS sets. In some embodiments, each COB in said plurality of complexes has a detectable signal or sequence that distinguishes it from other COBs in said population of cells. In some embodiments, an APS, an ESB, or a UBA is linked using chemical methods. In some embodiments, the chemical method comprises Click chemistry. In some embodiments, the linking is performed in the presence of Cu^{I}. In some embodiments, a UBA, an APS, or an ESB comprises nucleic acids. In some embodiments, the linking of a UBA, an APS, or an ESB is performed using a linking oligonucleotide that comprises a first and a second complementary region to two components to be linked. In some embodiments, the first or second complementary region is shared between APSs within a population of APSs. In some embodiments, the first or second complementary region is distinct for two different round-specific sets of APSs. Some embodiments further comprise ligation. In some embodiments, the linking oligonucleotide comprises a subcode encoding the origin population of the APS or the ESB. In some embodiments, the APS has a subcode encoding the round-specific set of the APS. In some embodiments, the ESB has a subcode encoding the origin presence of the ESB. In some embodiments, an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag. In some embodiments, detection comprises detecting the unique counter tag. In some embodiments, the number of unique counter tags associated with a specific ESB is determined. In some embodiments, the number of detected unique counter tags relate to the initial quantity of the specific ESB. In some embodiments, the ESB is covalently linked to the linking oligonucleotide. In some embodiments, an APS or an ESB comprises an amplification primer binding region. In some embodiments, a COB encodes a peptide sequence. In some embodiments, the COB comprises a polymerase start site. In some embodiments, the peptide comprises an affinity tag. In some embodiments, the affinity tag is a His-tag. In some embodiments, each of the reaction volumes created by the most recent splitting receives a different APS from the APS set.

[0026] In some embodiments, the invention provides methods for detecting at least one target molecule in a sample comprising the steps:(a) providing: (i) a population of cells potentially comprising at least one target molecule, (ii) a first UBA specific for a first target molecule, (iii) a first epitope specific barcode ESB specific for a region of the first UBA, where the ESB comprises a first common linker moiety, and (iv) a population of COB, where the population of COB comprises a second common linker moiety, where the second linker moiety is

complementary to the first common linker moiety is the first ESB; (b) forming at least a first complex comprising the at least one target molecule, the first UBA probe, and the first ESB, where the at least one target molecule is bound to the first UBA and the ESB is bound to the UBA (c) adding the population of COBs, where a second complex is formed with the least one target molecule, the first UBA probe, the first ESB, and a first COB, and where the second common linker moiety from the first COB is bound to the first linker moiety from the first ESB, and where the COBs from the population of COBs is associated with a cell from the population of cells; and (d) detecting the second complex or at least part of the third complex.

[0027] In some embodiments the invention provides methods for detecting at least one target molecule in a sample comprising the steps: (a) providing: (i) a population of cells potentially comprising at least one target molecule, (ii) a first unique binding agent (UBA) specific for a first target molecule, (iii) a first epitope specific barcode (ESB) specific for a region of the first UBA, where the ESB comprises a first common linker moiety, and (iv) a population of assayable polymer subunits (APSs), where the APSs comprises a second common linker moiety and a third common linker moiety, where the second linker moiety is complementary to the first common linker moiety is the first ESB; (b) forming at least a first complex comprising the at least one target molecule, the first UBA probe, and the first ESB, where the at least one target molecule is bound to the first UBA and the ESB is bound to the UBA; (c) splitting the population into two or more samples; (d) adding one APS from the population of APSs per sample to the two or more samples from step (c), where a second complex is formed with the least one target molecule, the first UBA probe, the first ESB, and a first APS, and where the second common linker moiety from the first APS is bound to the first linker moiety from the first ESB; (e) pooling the two or more samples from step (c) into one sample; (f) splitting the sample from step (e) into two or more samples; (g) adding one APS from the population of APSs per sample to the two or more samples from step (e), where a third complex is formed with the least one target molecule, the first UBA probe, the first ESB, the first APS, and the second APS, where the second common linker moiety from the second APS is bound to the third linker moiety from the first APS, and where the first APS and the second APS form a cell origination barcode (COB); and (c) detecting the third complex or at least part of the third complex. In some embodiments, the methods further comprise repeating steps (e), through (g).

[0028] In some embodiments, the methods further comprise detecting of a plurality of target molecules by forming a plurality of complexes in step (b), each complex comprising (i) at least one target molecule (ii) a first UBA and (iii) a first epitope specific barcode (ESB) specific for a

region of the first UBA, where the ESB comprises a first common linker moiety, where the at least one target molecule is bound to the first UBA and the ESB is bound to the UBA.

[0029] In some embodiments, each COB in the plurality of complexes has a detectable signal that distinguishes it from other COB in the population of cells.

[0030] In some embodiments, the complex is detected by sequencing or mass spectrometry. In some embodiments, the third complex is detected by a method comprising individually counting the presence of one or more molecules of the third complex where the presence of the one or more molecules of the third complex is indicative of the concentration of the target molecule in a cell. In some embodiments, the individually detecting further comprises detecting a digital signal.

[0031] In some embodiments, 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600, 700, 800, 900, 1000 or more than 1000 different target molecules are detected. In some embodiments, up to 2000 different target molecules are detected.

[0032] In some embodiments, the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid. In some embodiments, the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits.

[0033] In some embodiments, the APS is a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight. In some embodiments, the APS comprises a single-stranded nucleic acid hybridized to a complementary polynucleotide sequence having attached thereto a detectable label.

[0034] In some embodiments, the first APS is attached to the first ESB through ligation or extension via polymerization. In some embodiments, the second APS is attached to the first APS through ligation or extension via polymerization.

[0035] In some embodiments, the common linker moiety is a nucleic acid. In some embodiments, the ESB is attached to the USB.

[0036] In some embodiments, said first COB comprises a plurality of APS.

[0037] Some embodiments, further comprise detecting of a plurality of target molecules by a method comprising: forming a plurality of complexes in step (b), each complex comprising (i) at least one target molecule (ii) a first UBA and (iii) a first epitope specific barcode (ESB) specific for a region of said first UBA, wherein said ESB comprises a first common linker moiety, wherein said at least one target molecule is associated with said first UBA and said ESB is associated with said UBA.

[0038] In some embodiments, each COB in said plurality of complexes has a detectable signal or sequence that distinguishes it from other COBs in said population of cells.

[0039] In some embodiments, the linking of an APS, an ESB, or a UBA is performed using chemical methods. In some embodiments, the chemical method comprises Click chemistry. In some embodiments, the linking is performed in the presence of Cu^{I}. In some embodiments, linking of an ABS, an ESB, or a UBA is performed using binding affinity. In some embodiments, a UBA, an APS, or an ESB comprises nucleic acids. In some embodiments, the linking of a UBA, an APS, or an ESB is performed using a linking oligonucleotide that comprises a first and a second complementary region to two components to be linked. In some embodiments, the first or second complementary region is shared between APSs within a population of APSs. In some embodiments, the first or second complementary region is distinct for different populations of APSs. Some embodiments further comprise ligation. In some embodiments, the linking oligonucleotide comprises a subcode encoding the origin population of the APS or the ESB. In some embodiments, the APS has a subcode encoding the origin population of the APS. In some embodiments, the ESB has a subcode encoding the origin population of the ESB. In some embodiments, an individual APS, ESB, or linking oligonucleotide molecule comprises a unique tag. In some embodiments, detection comprises detecting the unique tag. In some embodiments, the number of unique tags associated with a specific ESB is determined. In some embodiments, the number of detected unique tags relate to the initial quantity of the specific ESB. In some embodiments, the ESB is covalently linked to the linking oligonucleotide. In some embodiments, an APS or an ESB comprises an amplification primer binding region. In some embodiments, a COB encodes a peptide sequence. In some embodiments, the COB comprises a polymerase start site. In some embodiments, the peptide comprises an affinity tag. In some embodiments, the affinity tag is a His-tag. In some embodiments, the two or more samples comprise at least 5 samples. In some embodiments, the two or more samples comprise at least 10 samples. In some embodiments, the two or more samples comprise at least 20 samples. In some embodiments, each of the samples created by the most recent splitting receives a different APS.

[0040] In some embodiments, the invention relates to methods for labeling an ESB linked target molecule of a cell in a population of cells with a cell origination barcode (COB), comprising: separating each cell into an individual reaction volume; and adding the COB to the ESB via chemical or affinity means. In some embodiments, the reaction volume is selected from the group consisting of a microbubble, a microdroplet, a well, a microwell, and an enclosure in a microfluidics device.

[0041] In some embodiments, the invention relates to methods comprising disassociating a variety of types of components originating from a cell and placing the components on a particle,

wherein the components are labeled on said particle. In some embodiments, the labeling comprises labeling according to cell origin. In some embodiments, the labeling comprises labeling according to component type.

[0042] In some embodiments, the signals in detection steps are ordered. In some embodiments, detecting comprises using one or more probes. In some embodiments, the probe is attached to a surface. In some embodiments, the surface comprises an array. In some embodiments, the surface comprises a bead. In some embodiments, detecting comprises a separation. In some embodiments, the separation is multi-dimensional.

[0043] In some embodiments, the invention provides methods for preparing at least one UBA, ESB and/or APS as described herein.

[0044] In some embodiments, the invention provides a population of UBAs, ESBs and/or APSs as described herein. In some embodiments, the invention provides kits comprising a population of UBAs, ESBs and APSs as described herein and instructions for its use.

INCORPORATION BY REFERENCE

[0045] All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0046] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0047] **Figure 1** depicts quantum of information representing a distinct signature (barcode) of the cell origin for each epitope.

[0048] **Figure 2** shows a graphical representation of one embodiment of the components of the epitope specific barcode and cells origin barcodes of the invention and their assembly.

[0049] **Figure 3** shows UBA-ESB-CL reagents of one embodiment of the invention.

[0050] **Figure 4** depicts labeling of cells in one embodiment of the invention with UBA-ESB-CL reagents.

[0051] **Figure 5** depicts ESB-COB reagents of one embodiment of the invention.

[0052] **Figure 6** depicts ESB-COB assembly according to one embodiment of the invention.

[0053] **Figure 7** depicts ESB-COB assembly according to another embodiment of the invention.

[0054] **Figure 8** depicts ESB-COB assembly according to another embodiment of the invention.

[0055] **Figure 9** depicts ESB-COB assembly according to another embodiment of the invention.

[0056] **Figure 10** depicts ESB-COB assembly according to another embodiment of the invention.

[0057] **Figure 11** depicts a peptide based ESB-COB readout according to one embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0058] The term "nucleic acid" refers to a nucleotide polymer, and unless otherwise limited, includes known analogs of natural nucleotides that can function in a similar manner (e.g., hybridize) to naturally occurring nucleotides.

[0059] The terms "polynucleotide", "nucleotide", "nucleotide sequence", "nucleic acid" and "oligonucleotide" are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three dimensional structure, and may perform any function, known or unknown. The following are non limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, intergenic DNA, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), small nucleolar RNA, ribozymes, complementary DNA (cDNA), which is a DNA representation of mRNA, usually obtained by reverse transcription of messenger RNA (mRNA) or by amplification; DNA molecules produced synthetically or by amplification, genomic DNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Polynucleotide sequences, when provided, are listed in the 5' to 3' direction, unless stated otherwise.

[0060] The term nucleic acid encompasses double- or triple-stranded nucleic acids, as well as single-stranded molecules. In double- or triple-stranded nucleic acids, the nucleic acid strands need not be coextensive (i.e., a double-stranded nucleic acid need not be double-stranded along the entire length of both strands).

[0061] The term nucleic acid also encompasses any chemical modification thereof, such as by methylation and/or by capping. Nucleic acid modifications can include addition of chemical groups that incorporate additional charge, polarizability, hydrogen bonding, electrostatic interaction, and functionality to the individual nucleic acid bases or to the nucleic acid as a whole. Such modifications may include base modifications such as 2′-position sugar modifications, 5-position pyrimidine modifications, 8-position purine modifications, modifications at cytosine exocyclic amines, substitutions of 5-bromo-uracil, backbone modifications, unusual base pairing combinations such as the isobases isocytidine and isoguanidine, and the like.

[0062] More particularly, in certain embodiments, nucleic acids, can include polydeoxyribonucleotides (containing 2-deoxy-D-ribose), polyribonucleotides (containing D-ribose), and any other type of nucleic acid that is an N- or C-glycoside of a purine or pyrimidine base, as well as other polymers containing normucleotidic backbones, for example, polyamide (e.g., peptide nucleic acids (PNAs)) and polymorpholino (commercially available from the Anti-Virals, Inc., Corvallis, Oreg., as Neugene) polymers, and other synthetic sequence-specific nucleic acid polymers providing that the polymers contain nucleobases in a configuration which allows for base pairing and base stacking, such as is found in DNA and RNA. The term nucleic acid also encompasses linked nucleic acids (LNAs), which are described in U.S. Pat. Nos. 6,794,499, 6,670,461, 6,262,490, and 6,770,748, which are incorporated herein by reference in their entirety for their disclosure of LNAs.

[0063] The nucleic acid(s) can be derived from a completely chemical synthesis process, such as a solid phase-mediated chemical synthesis, from a biological source, such as through isolation from any species that produces nucleic acid, or from processes that involve the manipulation of nucleic acids by molecular biology tools, such as DNA replication, PCR amplification, reverse transcription, or from a combination of those processes.

[0064] A nucleic acid "probe" is an oligonucleotide capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, generally through complementary base pairing, usually through hydrogen bond formation, thus forming a duplex structure. The probe binds or hybridizes to a "probe binding site." The probe can be labeled with a detectable label to permit facile detection of the probe, particularly once the probe has hybridized to its complementary target. Alternatively, however, the probe may be unlabeled, but may be detectable by specific binding with a ligand that is labeled, either directly or indirectly. Probes can vary significantly in size. Generally, probes are at least 7 to 15 nucleotides in length. Other probes are at least 20, 30, or 40 nucleotides long. Still other probes are somewhat longer,

being at least 50, 60, 70, 80, or 90 nucleotides long. Yet other probes are longer still, and are at least 100, 150, 200 or more nucleotides long. Probes can also be of any length that is within any range bounded by any of the above values (e.g., 15-20 nucleotides in length).

[0065] A primer or probe can be perfectly complementary to the target nucleic acid sequence or can be less than perfectly complementary. In certain embodiments, the primer has at least 65% identity to the complement of the target nucleic acid sequence over a sequence of at least 7 nucleotides, more typically over a sequence in the range of 10-30 nucleotides, and often over a sequence of at least 14-25 nucleotides, and more often has at least 75% identity, at least 85% identity, at least 90% identity, or at least 95%, 96%, 97%. 98%, or 99% identity. It will be understood that certain bases (e.g., the 3′ base of a primer) are generally desirably perfectly complementary to corresponding bases of the target nucleic acid sequence. Primer and probes typically anneal to the target sequence under stringent hybridization conditions.

[0066] Available binding interactions present in a mixture relying on affinity define specificity for two or more components' binding specificity. Generally, a high binding affinity for a first interaction in comparison to binding affinities of other available interactions that are available for one or more binding partners in the first interaction will lead to high specificity. Binding partners with high specificity form designated binding partners.

[0067] Reference will now be made in detail to particularly preferred embodiments of the invention. Examples of the preferred embodiments are illustrated in the following Examples section.

[0068] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. All patents and publications referred to herein are incorporated by reference in their entirety.

[0069] In some embodiments, the invention provides methods, compositions and kits for detection and quantification of individual target molecules in biomolecular samples. In some embodiments, the invention provides methods, compositions and kits for detection and quantification of individual target molecules in every cell of a complex cell population while retaining cell specific information regarding that target molecule. Thus in some embodiments, the invention provides methods, compositions and kits for detection and quantification of individual target molecules in a single cell basis in samples with complex cell populations. Thus in some embodiments, for each cell the amount of each target molecule associated with that cell is assayed. In particular the invention provides unique binding agents that are capable of binding individual target molecules. The invention also provides the use of epitope specific barcode to tag target molecules. The invention also provides the use of cell origination barcodes

to indicate the cell of origin. Through epitope specific barcodes and cell origination barcodes, the binding of unique binding agents to target molecules results in the identification of the target molecules. Methods of making and using such unique binding agents and/or epitope specific barcodes and/or cell origination barcodes are also provided. The methods and compositions described herein can be used in a wide variety of applications such as diagnostic, prognostic, quality control and screening applications. Some embodiments of the invention relate to methods, compositions and kits for individually tagging cells.

[0070] The term "epitope" and "target molecule" are used interchangeably herein to refer to the molecule of interest (parts of it or the whole molecule) being detected and/or quantified by the methods described herein.

[0071] Certain aspects of the invention relate to the detection of multiple target molecules. The methods described herein provide potential benefits in the areas of detection of multiple target molecules, quantification, and sensitivity. In some embodiments, the invention provides methods and compositions for the study of multiple protein measurements and/or multiple nucleic acid measurements that are sensitive and reliable.

[0072] Multiplexing within one sample at a single cell level is a key advantage of this approach. Multiplexing within one sample saves significant labor, reduces sample quantity requirements proportional to the number of measurements, and improves accuracy by elimination of errors compounded by separate sample handling and measurement steps. Furthermore, obtaining measurement of multiple target molecules in single cells in a complex cell population provides a better understanding of the physiological processes within each individual cell. In some embodiments, the methods described herein allow for the pooling of different samples together during processing to be analyzed at once. This offers throughput advantages and can accelerate the analysis of different samples.

[0073] In some embodiments, the invention provides unique binding agents (UBA) for the analysis of target molecules. In some embodiments, the invention provides an UBA population for use in a multiplexed assay. Each UBA in the population is specific for a target molecule. Thus, the UBA provides the specificity for the target molecule recognized in a cell. The binding of the target molecules to the UBAs is then detected using epitope specific barcodes (ESB) and cell origination barcodes (COB). Each ESB comprises a unique code that can be associated to a specific target molecule. Each COB comprises a unique code that can be associated to a specific cell of origin.

[0074] In some embodiments, the ESB are attached, directly or indirectly, to the UBA. In other embodiments, the ESBs bind to the UBAs in a cell or sample, e.g., as part of the assay

procedure. A unique COB is associated to the UBAs in a specific cell such that each COB can be associated to the target molecules bound to the UBAs in that cell. In some embodiments, the specific ESB/COB combination is referred as a quantum of information representing each target molecule or epitope (See Figure 1).

[0075] In some embodiments, the COB is composed of one or more assayable polymer subunit (APS). Certain aspects of the present invention relate to the selection of a library or population of designed (*e.g.*, synthetic sequences) APS. In some embodiments, the present invention provides a population of designed (e.g. synthetic) APS wherein said APS comprises an unique sequence and/or a detectable molecule, and wherein the combination of one or more different APS in each COB has a detectable signal or sequence that distinguishes it from other COBs in said population. In some embodiments, the invention provides APSs comprising unique sequences (e.g. synthetic) that hybridize to a unique complementary polynucleotide sequence having attached thereto a detectable label. In some embodiments, the APS are detected by sequencing. Accordingly, certain aspects of the present invention provide a population of unique COBs or ESB/COBs, each comprised of a unique APS-based combination, wherein each COBs or ESB/COBs in the population is distinct from the other COBs or ESB/COBs in the population. APSs generally are capable of forming a construct comprising the COBs. Any chemical structure allowing for such formation can be used for the APSs.

Unique Binding Agent (UBA)

[0076] UBAs are molecules or assemblies that are designed to bind with at least one target molecule, at least one target molecule surrogate, or both; and can, under appropriate conditions, form a molecular complex comprising the UBA and the target molecule. Examples of target molecules include, but are not limited to, proteins, nucleic acids, lipids, carbohydrates, ions, small molecules, organic monomers, and drugs. For convenience only, most of the embodiments described herein are explained in the context of UBAs that bind to a target protein or a target mRNA. However, these embodiments also can be applied to other target molecules. The terms "protein", "polypeptide", "peptide", and "amino acid sequence" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non amino acids or synthetic amino acids. The terms also encompass an amino acid polymer that has been modified, for example, by disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. As used herein the term "amino acid" refers to either natural and/or unnatural or synthetic amino acids, including but not

limited to glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics.

[0077] UBAs comprise at least one reaction portion that allow them to bind to or interact with at least one target molecule, at least one part of at least one target molecule, at least one target molecule surrogate, at least part of a target molecule surrogate, or combinations thereof; typically in a sequence-specific, a confirmation-specific manner, or both; for example but not limited to antigen-antibody binding, aptamer-target binding, and the like.

[0078] In certain embodiments, the UBAs comprise an identity portion or at least part of an identity portion, for example, an ESB, a COB, an ESB and/or a linker oligo. In certain embodiments, the UBAs comprise a capture region. In some embodiments, the capture region is used for the isolation of the UBA and/or immobilization of the UBA into a surface. The capture region can be an affinity tag, a bead, a slide, an array, a microdroplet, an enclosure in a microfluidic device or any other suitable capture region in the art. In some embodiments, the capture region is the ESB, for example the ESB can be a detectable bead such as a bead with a unique spectral signature (e.g. a bead that has been internally dyed with red and infrared fluorophores). Capture regions can define reaction volumes in which manipulation of compositions of the invention can take place.

[0079] In some embodiments, the UBA is an antibody. As used herein, the terms antibody and antibodies are used in a broad sense, to include not only intact antibody molecules, for example but not limited to immunoglobulin A, immunoglobulin G and immunoglobulin M, but also any immunoreactive component(s) of an antibody molecule that immunospecifically bind to at least one epitope. Such immunoreactive components include but are not limited to, FAb fragments, FAb' fragments, FAb'2 fragments, single chain antibody fragments (scFv), miniantibodies, diabodies, crosslinked antibody fragments, AffibodyTM, cyclotides, molecules, and the like. Immunoreactive products derived using antibody engineering or protein engineering techniques are also expressly within the meaning of the term antibodies. Detailed descriptions of antibody and/or protein engineering, including relevant protocols, can be found in, among other places, J. Maynard and G. Georgiou, Ann. Rev. Biomed. Eng. 2:339 76 (2000); Antibody Engineering, R. Kontermann and S. Dubel, eds., Springer Lab Manual, Springer Verlag (2001); U.S. Pat. No. 5,831,012; and S. Paul, Antibody Engineering Protocols, Humana Press (1995).

[0080] The skilled artisan will appreciate that antibody can be obtained from a variety of sources, including but not limited to polyclonal antibody, monoclonal antibody, monospecific antibody, recombinantly expressed antibody, humanized antibody, plantibodies, and the like; and can be obtained from a variety of animal species, including rabbit, mouse, goat, rat, human,

horse, bovine, guinea pig, chicken, sheep, donkey, human, and the like. A wide variety of antibodies are commercially available and custom-made antibodies can be obtained from a number of contract labs. Detailed descriptions of antibodies, including relevant protocols, can be found in, among other places, Current Protocols in Immunology, Coligan et al., eds., John Wiley & Sons (1999, including updates through August 2003); The Electronic Notebook; Basic Methods in Antibody Production and Characterization, G. Howard and D. Bethel, eds., CRC Press (2000); J. Goding, Monoclonal Antibodies: Principles and Practice, 3d Ed., Academic Press (1996); E. Harlow and D. Lane, Using Antibodies, Cold Spring Harbor Lab Press (1999); P. Shepherd and C. Dean, Monoclonal Antibodies: A Practical Approach, Oxford University Press (2000); A. Johnstone and M. Turner, Immunochemistry 1 and 2, Oxford University Press (1997); C. Borrebaeck, Antibody Engineering, 2d ed., Oxford university Press (1995); A. Johnstone and R. Thorpe, Immunochemistry in Practice, Blackwell Science, Ltd. (1996); H. Zola, Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000); and S. Hockfield et al., Selected Methods for Antibody and Nucleic Acid Probes, Cold Spring Harbor Lab Press (1993). Additionally, a vast number of commercially available antibodies, including labeled or unlabeled; polyclonal, monoclonal, and monospecific antibodies, as well as immunoreactive components thereof; custom antibody suppliers and the like can be found on the World Wide Web at, among other places, the Antibody Search page at biocompare.com, the Antibody Resource Page at antibodyresource.com, and the Antibody Explorer page at sigmaaldrich.com.

[0081] In some embodiments, the antibodies described herein are attached to a nucleic acid, e.g., linker oligo or a nucleic acid ESB. Methods to attach nucleic acids to antibodies are known in the art. Any suitable method to attach nucleic acids to antibodies is encompassed in the methods of the invention. The antibodies described herein can be attached to a nucleic acid by the methods described in Gullberg et al., PNAS 101 (22): pages 228420–8424 (2004); and Boozer et al, Analytical Chemistry, 76(23): pages 6967-6972 (2004), both incorporated herein by reference. The antibodies described herein can be attached to a nucleic acid by random amine attachment. In some embodiments, the antibodies described herein can be attached to a nucleic acid by random amine attachment using a 10 to 1 ratio of nucleic acid to antibody. The antibodies described herein can be attached to a nucleic acid by the methods described in Kozlov et al., Biopolymers 5: 73 (5): pages 621–630 (2004) incorporated herein by reference. The antibodies described herein can be attached to a nucleic acid by hydrazine chemistry. The antibodies described herein can be attached to a nucleic acid using tadpoles as described in

Nolan, Nature Methods 2, 11 - 12 (2005), incorporated herein by reference. The antibodies described herein can be attached to a nucleic acid by any suitable methods known in the art to generate engineered antibodies including the ones described herein.

[0082] In some embodiments, the UBA is an aptamer. Aptamers include nucleic acid aptamers (i.e., single-stranded DNA molecules or single-stranded RNA molecules) and peptide aptamers. Aptamers bind target molecules in a highly specific, conformation-dependent manner, typically with very high affinity, although aptamers with lower binding affinity can be selected if desired. Aptamers have been shown to distinguish between targets based on very small structural differences such as the presence or absence of a methyl or hydroxyl group and certain aptamers can distinguish between D- and L-enantiomers. Aptamers have been obtained that bind small molecular targets, including drugs, metal ions, and organic dyes, peptides, biotin, and proteins, including but not limited to streptavidin, VEGF, and viral proteins. Aptamers have been shown to retain functional activity after biotinylation, fluorescein labeling, and when attached to glass surfaces and microspheres.

[0083] Nucleic acid aptamers, including speigelmers, are identified by an in vitro selection process known as systematic evolution of ligands by exponential amplification (SELEX). In the SELEX process very large combinatorial libraries of oligonucleotides, for example 1014 to 1015 individual sequences, often as large as 60 -100 nucleotides long, are routinely screened by an iterative process of in vitro selection and amplification. Most targets are affinity enriched within 8 - 15 cycles and the process has been automated allowing for faster aptamer isolation. Peptide aptamers are typically identified by several different protein engineering techniques known in the art, including but not limited to, phage display, ribosome display, mRNA display, selectively infected phage technology (SIP), and the like. The skilled artisan will understand that nucleic acid aptamers and peptide aptamers can be obtained following conventional procedures and without undue experimentation. Detailed descriptions of aptamers, including relevant protocols, can be found in, among other places, L. Gold, J. Biol. Chem., 270(23):13581 84 (1995); S. Jayasena, Clin. Chem., 45:1628-50 (1999); V. Sieber et al., Nat Biotechnol. 16 (10):955-60 (1998); D. Wilson and J. Szostak, Ann. Rev. Biochem. 68:611-47 (1999); L. Jermutus et al., Eur. Biophys. J., 31:179-84 (2002); S S. Spada et al., Biol. Chem., 378:445-56 (1997); B. Wlotzka et al., Proc. Natl. Acad. Sci., 99:8898-8902 (2002).

[0084] In some embodiments the aptamer will be ligated or hybridized to nucleic acid such as a linker oligo or a nucleic acid ESB. The hybridization or ligation of aptamers can be done by any suitable method known in art. For example ligation can be performed enzymatically by at least one DNA ligase or at least one RNA ligase, for example but not limited to, T4 DNA ligase, T4

RNA ligase, Thermus thermophilus (Tth) ligase, Thermus aquaticus (Taq) DNA ligase, or Pyrococcus furiosus (Pfu) ligase. Ligation can also be performed by chemical ligation can, using activating and reducing agents such as carbodiimide, cyanogen bromide (BrCN), imidazole, 1-methylimidazole/carbodiimide/cystamine, N-cyanoimidazole, dithiothreitol (DTT) and ultraviolet light.

[0085] In some embodiments, the UBA is a peptoid. Peptoids are short sequences of N-substituted glycines synthetic peptides that bind proteins. In some embodiments, small size peptoids improve diffusion and kinetics of the methods described herein. Any suitable method known in the art to generate peptoids is encompassed in the methods described herein. See Simon et al., PNAS 15; 89(20): 9367–9371 (1992), incorporated herein by reference.

[0086] In some embodiments, the UBA is a nucleic acid sequence, e.g. an antisense DNA for a target mRNA. The nucleic acid sequence is preferably at least 15 nucleotides in length, and more preferably is at least 20 nucleotides in length. In specific embodiments, the target-specific sequence is about 10 to 500, 20 to 400, 30 to 300, 40 to 200, or 50 to 100 nucleotides in length. In other embodiments, the target-specific sequence is about 30 to 70, 40 to 80, 50 to 90, or 60 to 100, 30 to 120, 40 to 140, or 50 to 150 nucleotides in length.

Epitope Specific Barcode (ESB)

[0087] In some embodiments, the invention provides an epitope specific barcode (ESB). Each ESB comprises a unique code that can be associated to a specific target molecule. ESBs are molecules or assemblies that are designed to bind with at least one UBA or part of an UBA; and can, under appropriate conditions, form a molecular complex comprising the ESB, the UBA and the target molecule.

[0088] ESBs can comprise at least one identity identification portion that allow them to bind to or interact with at least one UBA; typically in a sequence-specific, a confirmation-specific manner, or both; for example but not limited to UBA-antibody binding, aptamer-target binding, and the like. In some embodiments, the ESB are attached, directly or indirectly, to the UBA. In other embodiments, the ESBs bind to the UBAs in a cell or sample, e.g., as part of the assay procedure.

[0089] In certain embodiments, the ESB is a solid surface or a capture region, for example, the ESB can be a detectable bead such as a bead with a unique spectral signature (e.g. a bead that has been internally dyed with red and infrared fluorophores). In some embodiments, the UBA is directly or indirectly attached to the capture region.

[0090] In certain embodiments, the ESBs comprise common linker moiety, for example, a linker oligo. In certain embodiments, the common linker oligo is complementary to a common linker oligo in the assayable polymer subunits (APSs) that form the cell origination barcode (COB).

[0091] In certain embodiments, the ESBs comprise a capture region. In some embodiments, the capture region is used for the isolation of the ESB and/or immobilization of the ESB into a surface. The capture region can be an affinity tag, a bead, a slide or an array. In some embodiments, the capture region is a detectable bead such as a bead with a unique spectral signature (e.g. a bead that has been internally dyed with red and infrared fluorophores

[0092] In some embodiments, the ESB is an antibody or fragment thereof, an aptamer, a nucleic acid, or peptoid, as described above.

[0093] In some embodiments, the ESB is a nucleic acid. In some embodiments, a part of the nucleic acid is amplified with branch chain or rolling circle approaches as known in the art.

[0094] In some embodiments, the ESB is a peptoid. Peptoids are short sequences of N-substituted glycines synthetic peptides that bind proteins. In some embodiments, small size peptoids improve diffusion and kinetics of the methods described herein. Any suitable method known in the art to generate peptoids is encompassed in the methods described herein. See Simon et al., PNAS 15; 89(20): 9367–9371 (1992), incorporated herein by reference.

[0095] In some embodiments, the ESB is a nucleic acid sequence, e.g. an antisense nucleic acid for a complementary target nucleic acid sequence. The nucleic acid sequence is preferably at least 15 nucleotides in length, and more preferably is at least 20 nucleotides in length. In specific embodiments, the target-specific sequence is about 10 to 500, 20 to 400, 30 to 300, 40 to 200, or 50 to 100 nucleotides in length. In other embodiments, the target-specific sequence is about 30 to 70, 40 to 80, 50 to 90, or 60 to 100, 30 to 120, 40 to 140, or 50 to 150 nucleotides in length.

Cell Origination Barcode (COB)

[0096] In some embodiments, the invention provides a cell origination barcode (COB). Each COB provides a unique code that can be associated to a specific cell of origin. In some embodiments, upon binding of the COB to a common linker moiety (e.g. common linker oligo) associated with an ESB, the COB code identifies the cells of origin of the target molecule to which the UBA/ESB complex is bound. Thus, in some embodiments the COBs of the invention comprise two main portions: (i) a sequence specific for a common linker moiety (e.g. common linker oligo) associated with a UBA/ESB probe; and (ii) an unique code that can be associated to a specific cell of origin.

[0097] In some embodiments, COBs are modular structures. In some embodiments, the COB comprises a plurality of different assayable polymer subunits (APS). In some embodiments, the COBs comprise a plurality of APSs attached in linear combination. In some embodiments, a COB is a molecular entity containing certain basic elements: (i) a plurality of APSs comprising label attachment regions attached in linear combination to form a backbone, and (ii) complementary polynucleotide sequences, comprising a label, which are complementary and are attached to the label attachment regions of the backbone. The term label attachment region includes a region of defined polynucleotide sequence within a given backbone that may serve as an individual attachment point for a detectable molecule. In some embodiments, the COBs comprise a plurality of different APSs attached in linear combination, wherein the APSs comprise small molecules of deterministic weight. In some embodiments, the COB comprises 2, 3, 4, 5, 6, 7, 8, 9, 10 or more unique APSs attached in a linear combination. In some embodiments, the COB comprises 4 or more APSs attached in linear combination.

[0098] In some embodiments, the plurality of APSs attached in linear combination can comprise uniquely designed nucleic acid sequences. In addition, the plurality of APSs attached in linear combination in the COBs can comprise at least one template, for example but not limited to, at least one nucleic acid sequence, such as at least part of a linear or linearizable viral genome, such as the genomes of adenovirus, hepatitis virus, herpes virus, rotavirus, and the like, or bacteriophages such as lambda, M13, ϕX-174, T-series bacteriophages, and the like, including derivatives thereof comprising cloning cassettes, polylinkers, and the like; plasmids, such as pBR322 and pUC series plasmids, etc., including derivatives thereof comprising cloning cassettes, polylinkers, and the like; synthetic templates; templates comprising artificial sequences; and the like. The skilled artisan will understand that virtually any piece of nucleic acid can serve as a template for fabricating a COB provided that it is large enough to include at least two APSs, or it can be combined with at least one other nucleic acid sequence so that the combined sequence is large enough to include at least two APSs. In some embodiments, ESBs, APSs, or COBs of the invention relate to templatable building blocks. In some embodiments, UBAs of the invention relate to templatable molecules.

[0099] In some embodiments, the COB also comprises one or more APSs containing a common linker moiety (e.g. common linker oligo). The common linker moiety can be directly or indirectly attached to the APSs. Thus, the common linker moiety can be covalently attached to a COB or the common linker moiety can be bound to the COB later in the assay. The term common linker moiety includes tandemly-repeated sequences of about 10 to about 25 nucleotides. The common linker moiety can be attached at either the 5' region or the 3' region of

a COB, and may be utilized for capture and immobilization of a COB for imaging or detection, such as by attaching to a solid substrate a sequence that is complementary to the common linker moiety.

[00100] The elements of a COB can be found in a single molecular entity (a singular COB), or two distinct molecular entities (a dual COB). Each molecular entity may be composed of one molecule or more than one molecule attached to one another by covalent or non-covalent means. In some embodiments, each component of a dual COB has a target molecule-specific UBA/ESB that binds to a different site on the same target molecule. When using a dual COB system one of the COBs may be unlabeled. In some embodiments, the unlabeled COB may comprise a capture region.

[00101] In various embodiments of the invention a COB is constructed from individual APS building blocks. In some embodiments, the APSs are combined linearly. In some embodiments, the COB constructed from the APSs maintains the order of the APS. In some embodiments, COBs comprise branched structures. In some embodiments, the branched COB structures comprise information about the order of APS addition. In various embodiments, the individual APS forming a COB can be decoded. In some embodiments, the order or order of addition of the APSs forming the COB can be decoded. Without being bound by theory, platforms allowing the APS addition order to be protected and decoded lend for a higher number of different COBs to be generated from equivalent number of APS building blocks. With a higher number of total COB molecule types, the likelihood of two cells/particles in a population being labeled with the same COB decreases. Thus, the methods of various embodiments of the invention allow for a higher statistical significance for the determination of cell/particle identity.

[00102] In some embodiments, the complementary polynucleotide sequences attached to an APS serve to attach detectable molecules, or label monomers, to the APS. The complementary polynucleotide sequences may be directly labeled, for example, by covalent incorporation of one or more detectable molecules into the complementary polynucleotide sequence. Alternatively, the complementary polynucleotide sequences may be indirectly labeled, such as by incorporation of biotin or other molecule capable of a specific ligand interaction into the complementary polynucleotide sequence. In such instances, the ligand (e.g., streptavidin in the case of biotin incorporation into the complementary polynucleotide sequence) may be covalently attached to the detectable molecule. Where the detectable molecules attached to an APS are not directly incorporated into the complementary polynucleotide sequence, this sequence serves as a bridge between the detectable molecule and the APS, and may be referred to as a bridging molecule, e.g., a bridging nucleic acid.

[00103] In some embodiments, the nucleic-acid based COBs, COB-ESB complexes, or COB/ESB/UBA complexes of the present invention comprise nucleic acids, which may be affinity-purified or immobilized using a nucleic acid, such as an oligonucleotide, that is complementary to a constant region of the COB (e.g. common linker moiety, capture region or affinity tag). As noted above, in some embodiments the COBs comprise at least one common linker moiety, which may serve as an affinity tag for purification and/or for immobilization (for example to a solid surface). The common linker moiety can comprises two or more tandemly-repeated regions of repeat nucleotides, such as a series of 15-base repeats. In such exemplary embodiments, the COB, whether complexed to an ESB, ESB/UBA, a target molecule/UBA/ESB or otherwise, can be purified or immobilized by an affinity reagent coated with a 15-base oligonucleotide which is the reverse complement of the repeat unit.

[00104] COBs, COB-ESB complexes, or COB/ESB/UBA complexes can be purified in two or more affinity selection steps. For example, in the embodiments in which the COB is attached to an ESB/UBA complex, the COB can comprise an affinity tag. In other embodiments when a dual COB is used, one COB can comprise a first affinity tag and the other COB can comprise a second (different) affinity tag. The COBs are mixed with the target molecules, and complexes comprising the two probes of the dual COBs are separated from unbound materials by affinity purification against one or both individual affinity tags.

[00105] In the first step, the mixture can be bound to an affinity reagent for the first affinity tag, so that only probes comprising the first affinity tag and the desired complexes are purified. The bound materials are released from the first affinity reagent and optionally bound to an affinity reagent for the second affinity tag, allowing the separation of complexes from COBs comprising the first affinity tag. At this point only full complexes would be bound. The complexes are finally released from the affinity reagent for the second affinity tag and then analyzed by the methods described herein. The affinity reagent can be any solid surface coated with a binding partner for the affinity tag, such as a column, bead (e.g., latex or magnetic bead) or slide coated with the binding partner. A variety of affinity tags known in the art may be used, e.g., to purify and/or immobilize COBs, COB-ESB complexes, or COB/ESB/UBA complexes. In some embodiments, a biotin anchor is attached to the COB, ESB and/or UBA, allowing immobilization of the COBs, COB-ESB complexes, or COB/ESB/UBA complexes on a streptavidin surface (e.g. coated slide or bead). In some embodiments, an affinity tag is attached to a UBA, e.g., to purify and/or immobilize the UBA. An affinity tag can be used for attachment to beads or other matrixes for a variety of useful applications including but not limited to purification. Examples of affinity tags and methods of making and/or attaching them to the

nucleic acids are described in US patent 7,473,767; US applications No. 10/542,458; 12/324,357; 11/645,270 and 12/541,131, incorporated herein by reference in their entirety.In some embodiments, at least two of ESB, UBAs, APSs, and COBs comprise different chemical compositions described herein or any other suitable composition known in the art. For example, an ESB can comprise a peptide and/or a nucleic acid, while an APS comprises a peptide or a peptoid. Any combination of described chemistries are envisioned within the scope of the invention.

[00106] In various embodiments, ESBs, APSs and/or COBs can be template, ordered, and or decoded. In some embodiments, the order of any of these subunits in a construct can be detected. ESBs, APSs and/or COBs can provide templates for the addition of any of another one of ESBs, APSs and/or COBs, for example via nucleic acid or any other suitable chemical complementarity known in the art. ESBs, APSs and/or COBs may encode secondary products, such as a nucleic acid encoding for another nucleic acid or a peptide. In some embodiments, the detection of the ESBs, APSs and/or COBs comprises decoding the same, for example generating an amplicon or expressing a peptide from the ESBs, APSs and/or COBs and sequencing or otherwise detecting the products.

[00107] The sequences, the weights or the signals provided by the label monomers associated with the various APS of the COB of a given cell allow for the unique identification of the COB. For example, when using nucleic acid sequences, a COB having a unique identity or unique sequence signature is associated with a UBA that recognizes a specific target molecule or a portion thereof. Detection of the COB sequence allows for the detection of the presence of the target molecule in the mixture (qualitative analysis). In another example, when using fluorescent labels, a COB having a unique identity or unique spectral signature is associated with a UBA that recognizes a specific target molecule or a portion thereof. Detection of the COB signal, such as the spectral code of a fluorescently labeled COB allows detection of the presence of the target molecule in the mixture (qualitative analysis). In yet another example, when using small molecules as per combinatorial synthesis procedure, a COB having a unique deterministic weight is associated with a UBA that recognizes a specific target molecule or a portion thereof. Detection of the COB deterministic weight (e.g., via mass spectrometry) allows detection of the presence of the target molecule in the mixture (qualitative analysis). Counting and or quantifying the codes (e.g. sequences, label monomers, or deterministic weights) associated with a given signature (e.g. spectral code, unique sequence, or unique deterministic weight) allow the counting or quantitation of all the molecules in the mixture associated the UBA coupled to the COB (quantitative analysis). UBA/ESB/COB complexes are thus useful for the diagnosis or

prognosis of different biological states (e.g., disease vs. healthy) by quantitative analysis of known biological markers.

[00108] Moreover, the exquisite sensitivity of single molecule detection and quantification provided by the COBs of the invention allows for the identification of new diagnostic and prognostic markers, including those whose fluctuations among the different biological states is too slight to detect a correlation with a particular biological state using traditional molecular methods. The sensitivity of COB-based molecular detection permits detailed pharmacokinetic analysis of therapeutic and diagnostic agents in small biological samples.

[00109] COB syntheses can be performed by any suitable methods known in the art, including the one described herein.

[00110] In various embodiments, the invention relates to a COB to be synthesized by stepwise addition of assayable polymer subunits (APSs) comprising oligonucleotides. The COB can be attached to the UBA via a common linker (CL). The CL can also be part of an oligonucleotide. In some embodiments, an epitope specific barcode is also provided as an oligonucleotide. In some instances, the epitope specific barcode can be included in the oligonucleotide that comprises the common linker. In some embodiments, CL, ESB and APSs all comprise oligonucleotide sequences. Accordingly, an oligonucleotide CL may be ligated to an oligonucleotide ESB or APS. Substantially complementary or exact complementary annealing regions may be utilized for hybridization. An annealing region may be provided on both ends of an oligonucleotide ESB or APS. In some embodiments, the APSs are added in various steps of a split pool synthesis or any other suitable stepwise synthesis known in the art. An annealing region specific to each step of a stepwise synthesis maybe incorporated into the oligonucleotides. Without being bound by theory, if an oligonucleotide addition is skipped during a step, the step-specific annealing region of the next oligonucleotide will not hybridize efficiently to the step-specific annealing region of the previous oligonucleotide that is available. Thus, the some embodiments of the invention provide methods to stall the synthesis of COBs missing one or more APSs.

[00111] In some embodiments, UBAs are labeled with different CL oligonucleotides, each with a unique ESB sequence specific to the UBA and a common annealing region. In many cases, the APSs are assembled into COBs during the rounds of split pool synthesis. In these cases, in each round, the sample is split into n different containers. A different oligonucleotide APS can be added to each container, totaling m different APSs. In some embodiments n and m are the same. In other embodiments, n is greater than m or m is greater than n. Each APS can be

designed so that it will selectively hybridize an annealing region added during the round before **(Figure 6)**. In various embodiments, a pair of round-specific annealing regions is incorporated in each APS. The annealing regions can be incorporated to each end of the APS. The annealing regions incorporated to each end of an APS can be different. Accordingly, an annealing region of an added APS can be complementary to an available annealing region from a previous round facilitating assembly.

[00112] In some embodiments, the APSs are stitched together and/or to a CL using an annealing primer (Figure 7 and 8). The annealing primer may comprise a first complementary region to the CL or an APS added during the previous round of a stepwise synthesis. The annealing primer may also comprise a second complementary region to the APS being added during a current round. Thus, the annealing primer can hybridize to two oligonucleotide subunits of subsequent rounds stitching them together. In some embodiments, the first complementary regions of annealing primers of each round are different from the first complementary regions of annealing primers of other rounds **(Figure 7)**. In some embodiments, the second complementary regions of annealing primers of each round are different from the second complementary regions of annealing primers of other rounds. In some embodiments, the first or second complementary regions of annealing primers of different rounds are shared between rounds **(Figure 8)**.

[00113] In some embodiments, a CL oligonucleotide comprises pairs of loop annealing regions **(Figures 9 and 10)**. Accordingly, APSs can be designed to hybridize to the CL in a loop geometry, hybridizing on each end to the CL along the loop annealing regions. The loop annealing regions can be specific to the round. The hybridization can populate the APSs along the CL. The APSs can be linked together using any method described herein or other common methods known in the art. The APSs can be designed such that they do not efficiently hybridize to the CL along the loop annealing regions specific to other rounds. Consequently, if an APS from a particular round is missing, the APSs may not be linked together successfully, depending on the linking process. Alternatively, a COB may be synthesized with a missing APS, the location of which is flanked by a pair of loop annealing regions. The resulting COB can then be analyzed accordingly and can either be discarded or the retrieved information can be alternatively processed.

[00114] Each APS in a given round can comprise a unique subcode sequence that is different from the rest of the APSs in that round **(Figures 6-11)**. The subcode may comprise a unique nucleotide sequence.

[00115] A CL, or one or more APSs may further comprise a random tag region allowing for subsequent normalization of the detected COBs **(Figures 6-11)**. Variations of suitable

methods making use of such random tag regions are known in the art, e.g., see Casbon et al. (Nucleic Acids Research, 2011, 39:12, e81). In some cases, the random tag region can function as a molecular counter to estimate the number of template molecules associated with each sequence variant. In some cases, a molecular counter is incorporated into a CL, ESB, APS or an assembled COB prior to an amplification reaction, e.g. PCR. A library of molecular counters comprising degenerate base regions (DBR) may be incorporated into CLs, ESBs, APSs or assembled COBs. The number of unique DBRs in a library generally is limited by the length and base composition of the DBR. For example, a single nucleotide in a DBR would allow for four different possible counters, one for each base. Longer DBRs can produce higher numbers of unique counter sequences. A molecular counter from a library of sequences can each be incorporated in a CL, ESB, APS or an assembled COB. The molecular counter can be used to determine whether a sequence variant is associated with a single template molecule, or alternatively, multiple template molecules. The number of different DBR sequences associated with one sequence variant can serve as a direct measure of the number of initial template molecules. This information can supplement or replace the information provided by read numbers of each sequence variant, including, for example, read numbers after an amplification reaction, e.g. PCR. DBRs can also be used to determine the probability that a sequence variant derives from a polymerase error during an amplification reaction or is a true original variant prior to an amplification reaction, e.g. PCR. In some embodiments, unique binding agents (UBAs) are fixed to their targets prior to or concurrent with the assembly of COBs.

[00116] Various embodiments of the invention relate to the assembly of COBs on the surface of cells. COBs can, for example, be assembled associated with UBAs targeting cell surface components. In some embodiments, UBAs are fixed to cell surface components prior to or concurrent with the assembly of COBs. In some embodiments, UBAs are delivered into cells or into cellular compartments. In some embodiments, COBs are assembled associated with UBAs within cells or cellular compartments. Cells may be fixed prior to the addition of UBAs, ESBs or prior to COB assembly. Suitable cell permeabilization methods are known in the art and can be used to deliver components of the assay into cells and cellular components.

[00117] In some embodiments, the assay is performed on bodies that are not cells. Suitable support materials known in the art, such as beads or surface coatings, can act in the same manner a cell would act to provide an original binding surface. Support materials can be decorated with binding targets. In some embodiments, support materials spatially resolve binding targets from each other.

[00118] In some embodiments, the assay may comprise primary binding targets and one or more secondary binding targets that are capable of binding to the primary target. A support material, can for example, be coated with one or more primary targets. A library of secondary targets can be provided to bind the primary targets. UBAs can be provided to bind epitopes of primary and/or secondary targets. COBs can be assembled associated with these UBAs as described for other types of targets. Inter-dependent binding of secondary targets to primary targets can be monitored by analyzing the COBs.

[00119] In some embodiments, multiple COBs are assembled on the same UBA molecule.

[00120] In some embodiments, the ESBs or assembled COBs encode a derivative sequence. In some embodiments, the ESBs and/or COBs comprise a polynucleotide sequence. In some cases, the ESB and/or COB can encode an RNA sequence. In some cases, the ESB and/or the COB encode a peptide sequence. The ESB and/or the COB can encode for the peptide sequence directly. Alternatively, the COB can encode for a peptide sequence indirectly, for example, through an intermediary RNA sequence. For example, a polynucleotide ESB or COB can encode an open reading frame. In some cases, the peptide sequence is translated after introducing the ESB or COB into a construct enabling peptide expression. In some embodiments, the construct is a vector.

[00121] In some embodiments, the ESBs and COBs are assembled from oligonucleotides. The linking agent can be a ligase. In some embodiments the ligase is T4 DNA ligase, using well known procedures (Maniatis, T. in Molecular Cloning, Cold Spring Harbor Laboratory (1982)). Other DNA ligases may also be used. With regard to ligation, other ligases, such as those derived from thermophilic organisms may be used thus permitting ligation at higher temperatures allowing the use of longer oligonucleotides (with increased specificity) as ESBs, CLs or APSs, which could be annealed and ligated simultaneously under the higher temperatures normally permissible for annealing such oligonucleotides. The ligation, however, need not be by an enzyme and, accordingly, the linking agent may be a chemical agent which will cause the ESB and APSs to link unless there is a nucleotide base pair mismatching at the aneealing region. For simplicity, some embodiments of the invention will be described using T4 DNA ligase as the linking agent. This enzyme requires the presence of a phosphate group on the 5' end that is to be joined to a 3' OH group on a neighboring oligonucleotide.

[00122] When oligos are stacking together to bind to an annealing region with a perfect match at the junction at their ends, it results in a specific binding to the annealing region. The

CLs, ESBs and APSs can be ligated to form an ESB linked COB. The ESB linked COBs, in turn, can be used for detection.

[00123] In some embodiments, the ESB and/or COB assembly comprises the use of CLICK chemistry. Suitable methods to link various molecules using CLICK chemistry are known in the art (for CLICK chemistry linkage of oligonucleotides, see, e.g. El-Sagheer et al. (PNAS, 108:28, 11338-11343, 2011).

[00124] In some embodiments, the ESB and/or COB assembly takes place inside a cell. In some embodiments, the ESBs and/or APSs are first assembled inside a cell. In some embodiments, the ESBs and/or APSs are linked inside a cell. In some embodiments, the ESBs and/or APSs are linked outside a cell.

[00125] In some embodiments, the assembled products are amplified and, optionally, the results are compared with amplification of similar target nucleic acids from a reference sample. In some embodiments, the ligated products of APSs are amplified and, optionally, results are compared with amplification of a similar COB from a reference sample. Amplification can be performed by any means known in the art. In some cases, the ligated products are amplified by polymerase chain reaction (PCR). Examples of PCR techniques that can be used include, but are not limited to, quantitative PCR, quantitative fluorescent PCR (QF-PCR), multiplex fluorescent PCR (MF-PCR), real time PCR (RTPCR), single cell PCR, restriction fragment length polymorphism PCR (PCR-RFLP), PCK-RFLPIRT-PCR-IRFLP, hot start PCR, nested PCR, in situ polonony PCR, in situ rolling circle amplification (RCA), bridge PCR, picotiter PCR and emulsion PCR. Other suitable amplification methods include the ligase chain reaction (LCR), transcription amplification, self-sustained sequence replication, selective amplification of target polynucleotide sequences, consensus sequence primed polymerase chain reaction (CP-PCR), arbitrarily primed polymerase chain reaction (AP-PCR), degenerate oligonucleotide-primed PCR (DOP-PCR) and nucleic acid based sequence amplification (NABSA). Other amplification methods that can be used herein include those described in U.S. Patent Nos. 5,242,794; 5,494,810; 4,988,617; and 6,582,938. In some embodiments, the amplification is performed inside a cell.

[00126] In any of the embodiments, amplification of ligated products may occur on a support, such as a bead or a surface. In any of the embodiments herein, COBs may be assembled on targets from a single cell.

[00127] The uniqueness of each UBA/ESB probe in a population of probes allows for the multiplexed analysis of a plurality of target molecules. Furthermore the uniqueness of each COB

probe in a population of probes allows for the multiplexed analysis of a plurality of target molecules in single cells.

[00128] For example, in some embodiments, each COB contains six APSs. If the APSs are going to be sequenced and there are 20 possible unique sequences for the APS. There will be 3.84×10^8 possible COBs in this example. Thus, 3.84×10^8 cells and their corresponding UBA/ESB probes can be analyzed in this example. Given that in sequencing you don't have the color constraints present in some fluorescent methods, one can analyze multiple UBA/ESB probes per cells. In some embodiments, 10, 15, 20, 25, 30, 40, 50, 60, 65, 70, 90, 100 different UBA/ESBs are analyzed per cell. In some embodiments, up to 100 different UBA/ESBs are analyzed per cell. In some embodiments, up to 1000 different UBA/ESBs are analyzed per cell. In some embodiments, up to 2000 different UBA/ESBs are analyzed per cell.

[00129] In certain embodiments, the methods of detection are performed in multiplex assays, whereby a plurality of target molecules is detected in the same assay (a single reaction mixture). In a preferred embodiment, the assay is a hybridization assay or an affinity binding assay in which the plurality of target molecules is detected simultaneously. In a preferred embodiment, the assay is hybridization assay or an affinity binding assay in which the plurality of target molecules is detected simultaneously in single cells. In certain embodiments, the plurality of target molecules detected in the same assay is, at least 2, at least 5 different target molecules, at least 10 different target molecules, at least 20 different target molecules, at least 50 different target molecules, at least 75 different target molecules, at least 100 different target molecules, at least 200 different target molecules, at least 500 different target molecules, or at least 750 different target molecules, or at least 1000 different target molecules. In other embodiments, the plurality of target molecules detected in the same assay is up to 50 different target molecules, up to 100 different target molecules, up to 150 different target molecules, up to 200 different target molecules, up to 300 different target molecules, up to 500 different target molecules, up to 750 different target molecules, up to 1000 different target molecules, up to 2000 target molecules, or up to 5000 target molecules. In yet other embodiments, the plurality of target molecules detected is any range in between the foregoing numbers of different target molecules, such as, but not limited to, from 20 to 50 different target molecules, from 50 to 200 different target molecules, from 100 to 1000 different target molecules, from 500 to 5000 different target molecules, and so on and so forth.

[00130] In some embodiments, the detection is digital detection. In some embodiments, the detection is direct, i.e. the method acquires a signal that is directly generated by the detected entity. In some embodiments, the detection is indirect, i.e. manipulation of an entity to be

detected takes place before a signal is acquired. In some embodiments, a plurality of components of an entity to be detected give rise to detection signals directly or indirectly. In some embodiments, the order of the plurality of the components can be determined through the detection methods described herein. Such detection methods are also described as ordered detection methods or detection methods with ordered signals.

[00131] In any of the embodiments, the detection or quantification analysis of the COBs can be accomplished by sequencing. The APS subunits or entire COBs can be detected via full sequencing of all DNA tags by any suitable methods known in the art, e.g., Illumina HiSeq 2000, including the sequencing methods described herein.

[00132] Sequencing can be accomplished through classic Sanger sequencing methods which are well known in the art. Sequencing can also be accomplished using high-throughput systems some of which allow detection of a sequenced nucleotide immediately after or upon its incorporation into a growing strand, i.e., detection of sequence in red time or substantially real time. In some cases, high throughput sequencing generates at least 1,000, at least 5,000, at least 10,000, at least 20,000, at least 30,000, at least 40,000, at least 50,000, at least 100,000 or at least 500,000 sequence reads per hour; with each read being at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 120 or at least 150 bases per read.

[00133] In some embodiments, high-throughput sequencing involves the use of technology available by Illumina's HiSeq 2000 machine. This machine uses reversible terminator-based sequencing by synthesis chemistry. This machine can do 200 billion DNA reads in eight days.

[00134] In some embodiments, high-throughput sequencing involves the use of technology available by ABI Solid System. This genetic analysis platform that enables massively parallel sequencing of clonally-amplified DNA fragments linked to beads. The sequencing methodology is based on sequential ligation with dye-labeled oligonucleotides.

[00135] In some embodiments, high-throughput sequencing involves the use of technology available by Ion Torrent Personal Genome Machine (PMG). The PGM can do 10 million reads in two hours.

[00136] In some embodiments, high-throughput sequencing involves the use of technology available by Helicos BioSciences Corporation (Cambridge, Massachusetts) such as the Single Molecule Sequencing by Synthesis (SMSS) method. SMSS is unique because it allows for sequencing the entire human genome in up to 24 hours. This fast sequencing method also allows for detection of a SNP nucleotide in a sequence in substantially real time or real time. Finally, SMSS is powerful because, like the MIP technology, it does not require a pre amplification step

prior to hybridization. In fact, SMSS does not require any amplification. SMSS is described in part in US Publication Application Nos. 2006002471 I; 20060024678; 20060012793; 20060012784; and 20050100932.

[00137] In some embodiments, high-throughput sequencing involves the use of technology available by 454 Lifesciences, Inc. (Branford, Connecticut) such as the Pico Titer Plate device which includes a fiber optic plate that transmits chemiluninescent signal generated by the sequencing reaction to be recorded by a CCD camera in the instrument. This use of fiber optics allows for the detection of a minimum of 20 million base pairs in 4.5 hours.

[00138] Methods for using bead amplification followed by fiber optics detection are described in Marguiles, M., et al. "Genome sequencing in microfabricated high-density pricolitre reactors", Nature, doi: 10.1038/nature03959; and well as in US Publication Application Nos. 20020012930; 20030058629; 20030100102; 20030148344; 20040248161 ; 20050079510, 20050124022; and 20060078909.

[00139] In some embodiments, high-throughput sequencing is performed using Clonal Single Molecule Array (Solexa, Inc.) or sequencing-by-synthesis (SBS) utilizing reversible terminator chemistry. These technologies are described in part in US Patent Nos. 6,969,488; 6,897,023; 6,833,246; 6,787,308; and US Publication Application Nos. 20040106130; 20030064398; 20030022207; and Constans, A., The Scientist 2003, 17(13):36.

[00140] In some embodiments, high-throughput sequencing of RNA or DNA can take place using AnyDot.chjps (Genovoxx, Germany). In particular, the AnyDot-chips allow for 10x - 50x enhancement of nucleotide fluorescence signal detection. AnyDot.chips and methods for using them are described in part in International Publication Application Nos. WO02/088382, WO03/020968, WO03/031947, WO2005/044836, PCT/EP05/105657, PCT/EP05/105655; and German Patent Application Nos. DE 101 49 786, DE 102 14 395, DE 103 56 837, DE 10 2004 009 704, DE 10 2004 025 696, DE 10 2004 025 746, DE 10 2004 025 694, DE 10 2004 025 695, DE 10 2004 025 744, DE 10 2004 025 745, and DE 10 2005 012 301.

[00141] Other high-throughput sequencing systems include those disclosed in Venter, J., et al. Science 16 February 2001; Adams, M. et al, Science 24 March 2000; and M. J, Levene, et al. Science 299:682-686, January 2003; as well as US Publication Application No. 20030044781 and 2006/0078937. Overall such systems involve sequencing a target nucleic acid molecule having a plurality of bases by the temporal addition of bases via a polymerization reaction that is measured on a molecule of nucleic acid, i e., the activity of a nucleic acid polymerizing enzyme on the template nucleic acid molecule to be sequenced is followed in real time. Sequence can then be deduced by identifying which base is being incorporated into the growing

complementary strand of the target nucleic acid by the catalytic activity of the nucleic acid polymerizing enzyme at each step in the sequence of base additions. A polymerase on the target nucleic acid molecule complex is provided in a position suitable to move along the target nucleic acid molecule and extend the oligonucleotide primer at an active site. A plurality of labeled types of nucleotide analogs are provided proximate to the active site, with each distinguishably type of nucleotide analog being complementary to a different nucleotide in the target nucleic acid sequence. The growing nucleic acid strand is extended by using the polymerase to add a nucleotide analog to the nucleic acid strand at the active site, where the nucleotide analog being added is complementary to the nucleotide of the target nucleic acid at the active site. The nucleotide analog added to the oligonucleotide primer as a result of the polymerizing step is identified. The steps of providing labeled nucleotide analogs, polymerizing the growing nucleic acid strand, and identifying the added nucleotide analog are repeated so that the nucleic acid strand is further extended and the sequence of the target nucleic acid is determined.

[00142] In various embodiments, oligonucleotide ESB or COBs are identified directly. The direct identification can comprise nucleic acid sequencing described supra. In some embodiments, the APS sequences encode for derivative sequences that are in turn identified. For example, a polynucleotide ESB and COB can be translated into a peptide sequence. The peptide sequence can then be identified using suitable methods known in the art.

[00143] In some embodiments, a sequence representing a COB or ESB is identified using mass spectrometric analysis. Sequencing methods comprising mass spectrometry are known in the art. In various embodiments, a derivative sequence, such as a peptide, is sequenced using mass spectrometric methods. In some embodiments, the mass spectrometric methods comprise fragmentation. In some embodiments, the mass spectrometric methods comprise N-terminal sequencing. In some embodiments, the mass spectrometric methods comprise C-terminal sequencing. In some embodiments, the mass spectrometric methods comprise Edman degradation. In various embodiments, the derivative sequence is subjected to a separation process prior to identification. In some embodiments, the separation process comprises chromatography. In some embodiments, the separation process comprises HPLC. Suitable separation methods for the separation process comprise, for example, ion-exchange chromatography or hydrophobic interaction chromatography. Ion-exchange chromatography can use any matrix material comprising functional groups that are strongly acidic (typically, sulfonic acid groups, e.g. sodium polystyrene sulfonate or polyAMPS), strongly basic, (quaternary amino groups, for example, trimethylammonium groups, e.g. polyAPTAC), weakly acidic (mostly, carboxylic acid groups), or weakly basic (primary, secondary, and/or ternary amino groups, e.g.

polyethylene amine). The separation method can, for example, use sulfonated polystyrene as a matrix, adding the amino acids in acid solution and passing a buffer of steadily increasing pH through the column. Amino acids are be eluted when the pH reaches their respective isoelectric points. Hydrophobic interaction chromatography may be employed through the use of reversed phase chromatography. Many commercially available C8 and C18 silica columns have demonstrated successful separation of peptides in solution through the use of an elution gradient.

[00144] In some embodiments, peptide sequences representing ESBs, APSs and the resulting COBs are designed to improve the ease of detection. In some cases, peptide sequences can be designed to improve fragmentation patterns in a mass spectrometer. It is well known in the art that the fragmentation efficiency of a bond in a peptide sequence is sequence dependent (see, e.g. Tabb et al. *Anal Chem*. 2003 March 1; 75(5): 1155 and Klammer et al. *Bioinformatics* 2008, 24: 348-356). The sequence-dependent fragmentation efficiencies can be employed to design representative peptide sequences with desired fragmentation patterns.

[00145] In some embodiments, peptide sequences representing ESBs, APSs and the resulting COBs are designed to confer certain physical and chemical characteristics to the peptide molecules. For example, representative peptide sequences can be designed to result in peptide molecules with solubilities in aqueous solutions within a desired range. For another example, representative peptide sequences can be designed to result in peptide molecules with a desired degree of secondary or tertiary structure or lack thereof. For yet another example, representative peptide sequences can be designed to result in peptide molecules with disulfide bonds or lack thereof. For a yet further example, representative peptide sequences can be designed to result in peptide molecules with desired binding characteristics to chosen targets. The sequences of the ESBs, CLs, APSs and the resulting COBs can further be designed to exploit specially loaded tRNAs in a protein expression system. Accordingly, incorporation of non-natural amino acids to the resulting peptide molecules can be accomplished.

[00146] Various embodiments relate to separating the ESB-linked COBs or derivative sequences, such as a peptide sequence, prior to detection. In some embodiments, the separation is based on a suitable physiochemical property of the molecules. These types of separations are particularly useful to direct the molecules to the detectors sequentially thereby increasing their relative abundance and the complexity of the signal at the time of the detection. Various embodiments comprise separation of the molecules in a sequence-targeted manner. For example, all molecules comprising a particular ESB can be isolated by hybridizing the ESB sequence to a sufficiently complementary sequence, affinity purifying using an ESB-specific linked tag, affinity purifying using a derivative sequence encoded by the ESB sequence or any other suitable

method known in the art. Separation methods may also target a cell-specific COB or a portion thereof.

[00147] In some embodiments, constructs comprising ESBs and COBs are subjected to separation. For example, the constructs can be subjected to gradient or affinity purification sorting according to the ESBs. In some embodiments, the separation may comprise multiple dimensions, for example 2, 3, 4, 5, 6, 7, or more dimensions. For example, the constructs can be separated sorting according to the first APSs on one dimensions and sorting the according to the second APSs on another dimension and optionally sorting according to the ESBs on a third dimension.

[00148] In some embodiments, the separation method comprises separation by gradient, on a gel, using an electromagnetic field and/or any other suitable separation method known in the art. In some embodiments, constructs comprising ESBs and/or COBs can be separated prior to detection. For example, the constructs can be separated according to ESBs and the ESBs and/or COBs in the constructs can be detected following the separation. The detection can be molar ratio detection, enzymatic detection, sequencing, differential affinity in a gradient or gel, electromagnetic field, e.g. UV, fluorescence or chemiluminescence, any detection method described herein, or any other suitable detection method known in the art. In some embodiments, the separation comprises immobilizing the constructs. For example, the constructs can be immobilized on an array surface or on a bead. A portion of the ESB and/or COB can be used to immobilize the constructs. The ESB and/or the COB can be detected from immobilization positions. In one example, ESBs and/or COBs comprising oligonucleotide can be immobilized on a bead or microarray that is coated with complementary oligonucleotides.

a. Detectable Molecules or Label Monomers

[00149] The COBs of the present invention can be labeled with any of a variety of label monomers, such as a radioisotope, fluorochrome, dye, enzyme, nanoparticle, chemiluminescent marker, biotin, or other monomer known in the art that can be detected directly (e.g., by light emission) or indirectly (e.g., by binding of a fluorescently-labeled antibody). Generally, one or more of the labeled APSs in the COB is labeled with one or more label monomers, and the signals provided by the label monomers attached to the APS of a COB constitute a detectable code that identifies the target to which the UBA the COB binds. In certain embodiments, the lack of a given signal from the APS (e.g., a dark spot) can also constitute part of the COB's code.

[00150] Example of label monomers that can be used with the COBs described herein and methods to incorporate the labels monomers into the COBs are described in US patent

7,473,767; US applications No. 10/542,458; 12/324,357; 11/645,270 and 12/541,131, incorporated herein by reference in their entirety.

[00151] When adding detectable molecules or label monomers to the COBs, in addition to the qualitative analytical capabilities provided by the COB of the invention and the analytical techniques based thereon, the COB of the invention are uniquely suitable for conducting quantitative analyses. By providing a one to one binding between the COB of the invention and their target molecules in a biomolecular sample, all or a representative portion of the target molecules present in the sample can be identified and counted. This individual counting of the various molecular species provides an accurate and direct method for determining the absolute or relative concentration of the target molecule in the biomolecular sample. Moreover, the ability to address each molecule in a mixture individually leverages benefits of miniaturization including high sensitivity, minimal sample quantity requirements, high reaction rates which are afforded by solution phase kinetics in a small volume, and ultimately very low reagent costs.

Target Molecules

[00152] Target molecules or epitopes are the molecules detected or measured by binding of a UBA whose target-specific region(s) recognize thereto. Examples of target molecules include, but are not limited to, proteins, nucleic acids, lipids, carbohydrates, small molecules, organic monomers, or drugs. Nucleic acids that can be analyzed by the methods herein include: double-stranded DNA, single-stranded DNA, single-stranded DNA hairpins, DNA/RNA hybrids, RNA (e.g. mRNA or miRNA) and RNA hairpins. For convenience only, the methods described herein are explained mostly in the context of analyzing proteins or mRNA. However, the embodiments described herein also can be used to detect non-protein or non-mRNA targets. In some embodiments, the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid.

[00153] A target molecule can be part of a biomolecular sample that contains other components or can be the sole or major component of the sample. A target molecule can be a component of a whole cell or tissue, a cell or tissue extract, a fractionated lysate thereof or a substantially purified molecule. The target molecule can be attached in solution or solid-phase, including, for example, to a solid surface such as a chip, microarray or bead. Also the target molecule can have either a known or unknown structure or sequence.

[00154] The compositions, methods, and kits disclosed herein can also be used in a wide variety of applications to determine the presence of target molecules in a sample. For example but without limitation, the compositions, methods, and kits are useful for, pharmacokinetic studies, including but not limited to, drug metabolism, ADME profiling, and toxicity studies; target validation for drug discovery; gene expression profiling, protein expression profiling; proteome analyses; metabolomic studies; post-translation modification studies, including but not limited to glycosylation, phosphorylation, acetylation, and amino acid modification, such as modification of glutamate to form gamma-carboxy glutamate and hydroxylation of proline to form hydroxylation; analyses of specific serum or mucosal antibody levels; evaluation of non-nucleic acid diagnostic indicators; foreign antigen detection; and the like.

[00155] In certain embodiment, at least one UBA, at least one ESB, or both the UBA and the ESB comprise at least one antibody, aptamer or peptoid that reacts specifically with at least one target molecule. In certain embodiments, at least one UBA, at least ESB, or both the UBA and the ESB comprise binding proteins that specifically interact with at least one target molecule. In some embodiments, the ESB comprise a common linker moiety.

[00156] The skilled artisan will appreciate that the molecular complexes and the at least part of the molecular complexes described herein can be individually detected while tethered or attached to a substrate or while in solution, depending on, among other things, the nature of the specific molecular complex or cleavable component and the SMD technique and detection apparatus employed.

Methods

[00157] The present invention provides methods for detection and quantification of target molecules in biomolecular samples. In particular, the invention provides UBAs that are capable of binding individual target molecules. The invention also provides the use of ESBs and COBs (See Figure 2). Through the ESBs' and COBs' codes, the binding of the UBAs to target molecules results in the identification of the target molecules in single cells. In some embodiments, the ESB/COB complex represents a quantum of information that represents the target molecule and the cell of origin (See Figure 1). Methods of making and using such UBAs and/or ESBs and COBs are also provided.

[00158] In one aspect, the invention provides methods to identify multiple target molecules in every cell of a complex cell population and to retain cell specific information regarding that target molecule. Therefore, for each cell the amount of each target molecule

associated with that cell is assayed. In some embodiments, multiple quanta of information are determined to identify multiple target molecules in every cell of a complex cell population.

[00159] In some embodiments, the invention provides methods for detecting at least one target molecule in a sample comprising the steps:(a) providing: (i) a population of cells potentially comprising at least one target molecule, (ii) a first UBA specific for a first target molecule, (iii) a first epitope specific barcode ESB specific for a region of the first UBA, where the ESB comprises a first common linker moiety, and (iv) a population of COB, where the population of COB comprises a second common linker moiety, where the second linker moiety is complementary to the first common linker moiety is the first ESB; (b) forming at least a first complex comprising the at least one target molecule, the first UBA probe, and the first ESB, where the at least one target molecule is bound to the first UBA and the ESB is bound to the UBA (c) adding the population of COBs, where a second complex is formed with the least one target molecule, the first UBA probe, the first ESB, and a first COB, and where the second common linker moiety from the first COB is bound to the first linker moiety from the first ESB, and where the COBs from the population of COBs is associated with a cell from the population of cells; and (d) detecting the second complex or at least part of the third complex.

[00160] In some embodiments, the invention provides methods for detection and/or quantification of a target molecule by binding a UBA to a target molecule. A UBA comprises at least one reaction portion that allow the UBA to bind to or interact with the target molecule; typically in a sequence-specific, a confirmation-specific manner, or both; for example but not limited to antigen-antibody binding, aptamer-target binding, and the like (See Figures 2 and 3).

[00161] In some embodiments, UBAs can be part of at least one probe set, comprising at least one first probe and at least one second probe. Thus, in some embodiments the invention provides methods for detection and/or quantification of a target molecule by binding a UBA probe set to a target molecule, where the UBA probe set comprises a first UBA probe and a second UBA probe. The first UBA probe and the second UBA probe comprise at least one reaction portion that allow the probes to bind to or interact with different regions of the target molecule, e.g., in a sequence-specific manner, a confirmation-specific manner, or both. In some embodiments, the UBA probe and/or a second UBA probe contain a capture region as described herein.

[00162] In certain embodiments, the UBAs comprise an identity portion or at least part of an identity portion, for example, an ESB, a COB, an ESB and/or a linker oligo. The identity portion allows for the identification of the presence or absence of the UBAs bound to the target molecule in the detection step of the methods described herein. Thus, in some embodiments the

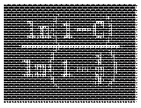
invention provides methods for detection and/or quantification of a target molecule by binding the UBA to a target molecule, wherein UBA contains an identity portion (e.g., ESB, a COB, an ESB and/or a linker oligo).

[00163] In some embodiments, the target molecule is tagged within cells indirectly with an epitope specific barcode (ESB). Each ESB comprises a unique code that can be associated to a specific target molecule. ESBs are molecules or assemblies that are designed to bind with at least one UBA or part of an UBA; and can, under appropriate conditions, form a molecular complex comprising the ESB, the UBA and the target molecule. ESBs comprise at least one identity identification portion that allow them to bind to or interact with at least one UBA; typically in a sequence-specific, a confirmation-specific manner, or both; for example but not limited to UBA-antibody binding, aptamer-target binding, and the like. In some embodiments, the ESB are attached, directly or indirectly, to the UBA. In other embodiments, the ESBs bind to the UBAs in a cell or sample, e.g., as part of the assay procedure. In certain embodiments, the UBAs and/or ESBs comprise a capture region. In some embodiments, the capture region is used for the isolation of the UBA/ESB and/or immobilization of the UBA/ESB into a surface. The capture region can be an affinity tag, a bead, a slide or an array. In certain embodiments, the ESBs comprise common linker moiety, for example, a linker oligo. In certain embodiments, the common linker oligo is complementary to a common linker oligo in the assayable polymer subunits (APSs) that form the cell origination barcode (COB).

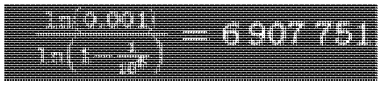
[00164] Figures 3 and 4 show a schematic representation of one of the embodiments of the invention in which a split pool synthesis approach is used to append the COB to the target molecule/UBA/ESB complex. Figure 3 in Step 1 shows the labeling of cells with UBA-ESB-CL reagents. The UBA provides the specificity for the target molecule to be recognized in a cell. In some embodiments, the UBA can be an antibody specific for a surface marker like CD8, or an intracellular epitope like a phospho-epitope on a kinase such as Stat-3. In some embodiments, the UBA could be an antisense DNA for a target mRNA in a fixed cell. The UBA is identified with an ESB that has a CL moiety, the latter for later addition of cell specific tagging information. Figure 4 is Step 2 shows the start of the split pool synthesis. In this example, the cell population is split into 20 tubes. Cell populations can be split into wells, bead or any suitable surfaces known in the art. In Step 3 an APS unit is added to each tube. The APS binds the UBA/ESB complex via the complementary CLs in the APS and the ESB. In Step 4, each cell in a given tube now has appended to each UBA-ESB pair the same subunit as defined by the tube contents (one of 1-20 APS in this example). Each split population from Step 3 now has a DAPS "tag" polymer subunit added to all DBAAs in that cell. In Step 5, the cells from the 20

tubes are pulled into one tube. 1/20 of the cells have the same APS subunit. In Step 6, steps 2-4 are repeated to add a second APS to the prior APS. In this example, cells in one tube will have a mixture of cells all of which have the APS subunit from the round 2 and one of the 20 APSs used in round 1 in a statistically equal distribution. Thus, in round two within each individual tube mixture all polymers are extended by the addition of the same APS. The process is repeated as needed. The epitope/barcode along with linked cell origin signature is read by any methods known in the art including the ones described herein.

[00165] The number of split pool rounds required is defined by the number of cells in an assay and a statistical estimate of what would give an over-representation of the number of tags that ensure unique COB for each cell. This is given by the following equation:

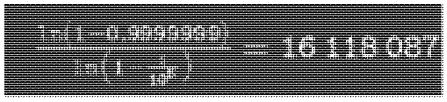
Number of tags required =  where C = certainty of over-representation and where N = number of cells

[00166] Thus, if you have 1 million cells, and you want 99.9% certainty of tag uniqueness, you need:



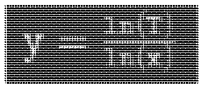
Or approximately 7 million tags. In various embodiments, a high certainty of tag uniqueness ensures a high statistical significance for identification of cells/particles as distinct entities. Without being bound by theory, a high certainty of tag uniqeness provides for a high likelihood that two identical COB labels originated from the same cell/particle.

[00167] However, for 10^6 cells, 7 million tags means 1000 cell pairs could be labeled as the "same" cell. Therefore, to have only a 1 in 10 chance there is a SINGLE pair of duplicate cells, the equation should be set to 99.99999%

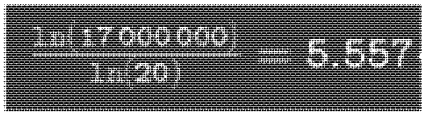


Thus requiring 16-fold more tags than cells.

[00168] To determine the number of rounds given a given number of subunits for barcode creation:

 gives 

If you had 20 subunits, for 1.7×10^7 tags you need the following APS addition cycles:



This can be round up to 6 APS addition cycles.

If you had 100 subunits, you would need only 3.6 rounds, or round-up to 4 APS addition cycles.

[00169] Thus, in some embodiments, the invention provides methods for tagging target molecules within cells indirectly with an ESB. The cell population is treated in a split pool synthesis approach that appends to the epitope specific barcode a second signature that indicates the cell of origin, or a cell origin barcode. UBAs can be antibodies, diabodies, etc. for proteins, or anti-sense DNA tags for RNA and DNA for nucleic acids. ESB can be nucleic acids readable by high throughput sequencing approaches or chemical subunits assayable by mass spectrometry approaches. COBs can be nucleic acids readable by high throughput sequencing approaches or chemical subunits assayable by mass spectrometry approaches.

[00170] The APS can be specific strands of DNA or DNA-mimics. The APSs can be linked via ligase. Example of enzymes that can be used for ligation include but are not limited to DNA ligase, and RNA ligase such as T4 DNA ligase, T4 RNA ligase, Thermus thermophilus (Tth) ligase, Thermus aquaticus (Taq) DNA ligase, or Pyrococcus furiosus (Pfu) ligase. Chemical ligation can be performed using activating and reducing agents such as carbodiimide, cyanogen bromide (BrCN), imidazole, 1-methylimidazole/carbodiimide/cystamine, N-cyanoimidazole, dithiothreitol (DTT) and ultraviolet light. Also within the scope of the invention are ligation techniques such as gap-filling ligation, including, without limitation, gap-filling OLA and LCR, bridging oligonucleotide ligation, and correction ligation. Descriptions of these techniques can be found, among other places, in U.S. Pat. No. 5,185,243, published European Patent Applications EP 320308 and EP 439182, and PCT Publication Nos. WO 90/01069 and WO 01/57268. The APSs can be extended via polymerases.

[00171] These APS subunits can be detected via full sequencing of all DNA tags by any suitable methods known in the art, e.g., Illumina HiSeq 2000, including the sequencing methods described herein.

[00172] The APS can be small molecules as per combinatorial synthesis procedures or buildable complex molecules of deterministic weights. These subunits can be detected via mass spectrometry

[00173] Thus in some embodiments, the cell-specific information is assembled via a UBA (barcode for the epitope to be recognized) as linked to its associated COB (from which cell the information originated) (See Figure 5).

[00174] In some embodiments, the UBA/ESB/COB complexes are isolated via a capture region as described herein. In some embodiments, the capture region is uses for the immobilization of the UBA/ESB/COB complexes into a surface.

[00175] In some embodiments, the information on UBAs can be amplified prior to the then COB procedure (split pool) with any suitable amplification technique known in the art, including the ones described herein such branch chain or rolling circle approaches.

[00176] In some embodiments, error correction & detection can be encoded into subunits. The general idea for achieving error detection and correction is to add some redundancy (i.e., some extra data) to a message, which receivers can use to check consistency of the delivered message, and to recover data determined to be erroneous. Error-detection and correction schemes can be either systematic or non-systematic: In a systematic scheme, the transmitter sends the original data, and attaches a fixed number of check bits (or parity data), which are derived from the data bits by some deterministic algorithm. If only error detection is required, a receiver can simply apply the same algorithm to the received data bits and compare its output with the received check bits; if the values do not match, an error has occurred at some point during the transmission. In a system that uses a non-systematic code, the original message is transformed into an encoded message that has at least as many bits as the original message. Error detection and correction schemes include repetition codes, parity bits, checksums, cyclic redundancy checks (CRCs), cryptographic hash functions, error-correcting codes, automatic repeat request, Hybrid ARQ, error-correcting code, convolutional codes, block codes such as hamming codes, multidimensional parity-check codes, Reed-Solomon codes, turbo codes and low-density parity-check codes (LDPC).

[00177] In some embodiments, the chains at each round are blocked from further addition if polymers did not add. This could be accomplished with DNA as polymer units if each round one uses different overhangs for complementary additions.

[00178] In some embodiments, the epitope/barcode along with linked cell origin signature is read by sequencing using methods known in the art, including the ones described herein. The sensitivity for the sequencing approach assuming that each 100 bp read represents an ESB-COB pair is as follows for target proteins. The protein copy numbers of molecules of interest range from 100 to 100,000. Assuming that one will want to read 100 proteins with the following rough distribution:

	Test 1 proteins	**Test 1 reads**	**Test 2 proteins**	**Test 2 reads**
100 copies	20	2×10^3	20	2×10^3
500 copies	20	1×10^4	20	1×10^4

1000 copies	20	2×10^4	20	2×10^4
10000 copies	20	2×10^5	40	4×10^5
100000 copies	20	2×10^6	0	

[00179] In Test 1 one will need to be able to read 2,232,000 sequences per cell then to access all 100 proteins in a cell. Using a sequencing technique that can do 2×10^9 reads such as Illumina HiSeq 2000, this means the approach can read 100 proteins in ~1000 cells. However, if one limits the high copy number proteins by avoiding them altogether or "capping" their representation by normalization (several approaches can work for this), we can increase the cell number accessible. Say one therefore normalize and cap the 100,000 copies to a 10,000 copy limit (Test 2), then the total number of reads is 432,000. That is one can read 100 proteins in ~5000 cells.

[00180] For mRNA the numbers are different, since RNAs are typically expressed much lower. The RNA copy numbers of molecules of interest range from 5 to 1000 (based on Lewin's Essential Genes numbers); not counting specialized mRNAs for high protein production like actin or Ig. Thus, assuming one wants to read 100 mRNAs with the following rough distribution:

	Test 1 mRNAs	Test 1 reads
5 copies	60	300
50 copies	20	1000
100 copies	10	1000
1000 copies	10	10000

[00181] In Test 1 one needs to be able to read 12,300 sequences per cell to access all 100 mRNAs in a cell. Using a sequencing technique that can do 2×10^9 reads such as Illumina HiSeq 2000, this means the approach can read 100 mRNAs in ~162,000 cells. This is equivalent to a high parameter flow cytometry run. 200 mRNAs with the same distribution could scale linearly to be read on ~80,000 cells.

[00182] As it will be evident to those skilled in the art increasing the number of reads will increase the cell number and parameters accessible (e.g., mRNAs or proteins).

[00183] Any of the embodiments described herein can be used in the detection of multiple target molecules. In some embodiments, the invention provides methods comprising UBA for the analysis of target molecules. In some embodiments, the invention provides a UBA population for use in a multiplexed assay. Each UBA in the population is specific for a target molecule. The binding of the target molecules to the UBAs is then detected using ESB-COB

pairs. Each ESB-COB pair comprises a unique label code that can be associated to a specific target molecule and the cell of origin as described herein.

[00184] In some embodiments, the detection of the ESB-COB as described below is digital in nature in that one molecule at a time is counted. While fluorescence is used to read the code, the signals are high and the spot is either present of not, thus the digital detection. Using digital detection rather than an analogue fluorescent signal used to quantify signal leads to more accurate quantification. Thus the methods described herein allows for multiplexing to levels beyond currently possible, for more accurate quantification, and possibly higher sensitivity.

Biomolecular Samples

[00185] The UBA and ESB/COB systems of the invention can be used to detect target molecules in any biomolecular sample. As will be appreciated by those in the art, the sample may comprise any number of things, including, but not limited to: biological samples, such as cells (including both primary cells and cultured cell lines), cell lysates, or extracts, tissues and tissue extracts; bodily fluids (including, but not limited to, blood, urine, serum, lymph, bile, cerebrospinal fluid, interstitial fluid, aqueous or vitreous humor, colostrum, sputum, amniotic fluid, saliva, anal and vaginal secretions, perspiration and semen, a transudate, an exudate (e.g., fluid obtained from an abscess or any other site of infection or inflammation) or fluid obtained from a joint (e.g., a normal joint or a joint affected by disease such as rheumatoid arthritis, osteoarthritis, gout or septic arthritis) of virtually any organism, with mammalian samples being preferred and human samples being particularly preferred; environmental samples (including, but not limited to, air, agricultural, water and soil samples); biological warfare agent samples; research samples including extracellular fluids, extracellular supernatants from cell cultures, inclusion bodies in bacteria, cellular compartments, cellular periplasm, mitochondria compartment, etc.

[00186] The biomolecular samples can be indirectly derived from biological specimens. For example, where the target molecule of interest is a protein kinase the biomolecular sample of the invention can be a sample containing isolated proteins from a cell lysate. In another example, the biomolecular sample of the invention is generated by subjecting a biological specimen to fractionation, e.g., size fractionation or membrane fractionation.

[00187] Protein isolation techniques are also well known in the art and kits employing at least some of these techniques are commercially available. Protein isolation techniques typically employ one or more of the following: maceration and cell lysis, including physical, chemical and enzymatic methods; centrifugation; separations by molecular weight, such as size exclusion

chromatography and preparative electrophoresis; selective precipitation, for example, salting-in and salting-out procedures; various chromatographic methods; and the like. Detailed descriptions of and relevant protocols for protein purification techniques can be found in, among other places, Marchak et al., Strategies for Protein Purification and Characterization: A Laboratory Course Manual, Cold Spring Harbor Press (1996); Essentials from Cells: A Laboratory Manual, D. Spector and R. Goldman, eds., Cold Spring Harbor Press (2003); R. Simpson, Proteins and Proteomics: A Laboratory Manual, Cold Spring Harbor Press (2003); and D. Liebler, Introduction to Proteomics, Humana Press (2002). Commercially available kits can also be used, for example but not limited to, ProteoExtract.TM. Partial Proteome Extraction Kits (P-PEK) and ProteoExtract.TM. Complete Proteome Extraction Kits (C-PEK), available from CALBIOCHEM.RTM., La Jolla, Calif. The skilled artisan will appreciate that non-nucleic acid analytes for use with the inventive compositions, methods, and kits can be readily obtained without undue experimentation using such purification techniques and commercial kits

[00188] The biomolecular samples of the invention may be either native, e.g., not subject to manipulation or treatment, or treated, which can include any number of treatments, including exposure to candidate agents including drugs, genetic engineering (e.g., the addition or deletion of a gene), etc.

[00189] Biomolecular samples may also include environmental samples, such as those containing bacteria or other organisms, such as diatoms, dinoflagellates, algae, among others, such as in certain marine or earth-based samples.

Detection of COBs

[00190] COBs/ESB complexes are detected by any means available in the art that is capable of detecting the specific sequences or signals on a given COBs/ESB complex.

[00191] In some embodiments, the information on the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof can be amplified. Amplification can be performed by any means known in the art. In some cases, the information on the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof are amplified by polymerase chain reaction (PCR). Examples of PCR techniques that can be used include, but are not limited to, quantitative PCR, quantitative fluorescent PCR (QF-PCR), multiplex fluorescent PCR (MF-PCR), real time PCR (RT-PCR), single cell PCR, restriction fragment length polymorphism PCR (PCR-RFLP), PCR-RFLP/RT-PCR-RFLP, hot start PCR, nested PCR, *in situ* polonony PCR, in situ rolling circle amplification (RCA), bridge PCR, picotiter PCR and emulsion PCR. Other suitable

amplification methods include the ligase chain reaction (LCR), transcription amplification, self-sustained sequence replication, selective amplification of target polynucleotide sequences, consensus sequence primed polymerase chain reaction (CP-PCR), arbitrarily primed polymerase chain reaction (AP-PCR), degenerate oligonucleotide-primed PCR (DOP-PCR) and nucleic acid based sequence amplification (NABSA). Other amplification methods that can be used herein include those described in U.S. Patent Nos. 5,242,794; 5,494,810; 4,988,617; and 6,582,938.

[00192] In any of the embodiments, amplification of the information on the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof may occur on a bead. In any of the embodiments herein, target nucleic acids may be obtained from a single cell.

[00193] In any of the embodiments herein, the information on the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof can be pre-amplified prior to the amplification step (e.g., PCR).

[00194] In some embodiments the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof are quantified. Methods for quantifying nucleic acids are known in the art and include, but are not limited to, gas chromatography, supercritical fluid chromatography, liquid chromatography (including partition chromatography, adsorption chromatography, ion exchange chromatography, size-exclusion chromatography, thin-layer chromatography, and affinity chromatography), electrophoresis (including capillary electrophoresis, capillary zone electrophoresis, capillary isoelectric focusing, capillary electrochromatography, micellar electrokinetic capillary chromatography, isotachophoresis, transient isotachophoresis and capillary gel electrophoresis), comparative genomic hybridization (CGH), microarrays, bead arrays, and high-throughput genotyping such as with the use of molecular inversion probe (MIP).

[00195] Quantification of the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof can be used to determine gene/or allele copy number, gene or exon-level expression, methylation-state analysis, or detect a novel transcript in order to diagnose or condition, i.e. fetal abnormality or cancer.

[00196] In some embodiments the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof are sequenced. Sequencing can be accomplished through classic Sanger sequencing methods which are well known in the art. Sequence can also be accomplished using high-throughput systems some of which allow detection of a sequenced nucleotide immediately after or upon its incorporation into a growing strand, i.e., detection of sequence in real time or substantially real time. In some

cases, high throughput sequencing generates at least 1,000, at least 5,000, at least 10,000, at least 20,000, at least 30,000, at least 40,000, at least 50,000, at least 100,000 or at least 500,000 sequence reads per hour; with each read being at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 120 or at least 150 bases per read.

[00197] In some embodiments, high-throughput sequencing involves the use of technology available by Illumina's HiSeq 2000 machine. This machine uses reversible terminator-based sequencing by synthesis chemistry. This machine can do 200 billion DNA reads in eight days.

[00198] In some embodiments, high-throughput sequencing involves the use of technology available by ABI Solid System. This genetic analysis platform that enables massively parallel sequencing of clonally-amplified DNA fragments linked to beads. The sequencing methodology is based on sequential ligation with dye-labeled oligonucleotides.

[00199] In some embodiments, high-throughput sequencing involves the use of technology available by Ion Torrent Personal Genome Machine (PMG). The PGM can do 10 million reads in two hours.

[00200] In some embodiments, high-throughput sequencing involves the use of technology available by Helicos BioSciences Corporation (Cambridge, Massachusetts) such as the Single Molecule Sequencing by Synthesis (SMSS) method. SMSS is unique because it allows for sequencing the entire human genome in up to 24 hours. Finally, SMSS is described in part in US Publication Application Nos. 20060024711; 20060024678; 20060012793; 20060012784; and 20050100932.

[00201] In some embodiments, high-throughput sequencing involves the use of technology available by 454 Lifesciences, Inc. (Branford, Connecticut) such as the PicoTiterPlate device which includes a fiber optic plate that transmits chemilluminescent signal generated by the sequencing reaction to be recorded by a CCD camera in the instrument. This use of fiber optics allows for the detection of a minimum of 20 million base pairs in 4.5 hours.

[00202] Methods for using bead amplification followed by fiber optics detection are described in Marguiles, M., et al. "Genome sequencing in microfabricated high-density pricolitre reactors", Nature, doi:10.1038/nature03959; and well as in US Publication Application Nos. 20020012930; 20030068629; 20030100102; 20030148344 ; 20040248161; 20050079510, 20050124022; and 20060078909.

[00203] In some embodiments, high-throughput sequencing is performed using Clonal Single Molecule Array (Solexa, Inc.) or sequencing-by-synthesis (SBS) utilizing reversible terminator chemistry. These technologies are described in part in US Patent Nos. 6,969,488;

6,897,023; 6,833,246; 6,787,308; and US Publication Application Nos. 20040106110; 20030064398; 20030022207; and Constans, A., The Scientist 2003, 17(13):36.

[00204] In some embodiments, high-throughput sequencing can take place using AnyDot.chips (Genovoxx, Germany). In particular, the AnyDot.chips allow for 10x – 50x enhancement of nucleotide fluorescence signal detection. AnyDot.chips and methods for using them are described in part in International Publication Application Nos. WO 02088382, WO 03020968, WO 03031947, WO 2005044836, PCT/EP 05/05657, PCT/EP 05/05655; and German Patent Application Nos. DE 101 49 786, DE 102 14 395, DE 103 56 837, DE 10 2004 009 704, DE 10 2004 025 696, DE 10 2004 025 746, DE 10 2004 025 694, DE 10 2004 025 695, DE 10 2004 025 744, DE 10 2004 025 745, and DE 10 2005 012 301.

[00205] Other high-throughput sequencing systems include those disclosed in Venter, J., et al. Science 16 February 2001; Adams, M. et al. Science 24 March 2000; and M. J. Levene, et al. Science 299:682-686, January 2003; as well as US Publication Application No. 20030044781 and 2006/0078937. Overall such system involve sequencing a target nucleic acid molecule having a plurality of bases by the temporal addition of bases via a polymerization reaction that is measured on a molecule of nucleic acid, i.e. the activity of a nucleic acid polymerizing enzyme on the template nucleic acid molecule to be sequenced is followed in real time. Sequence can then be deduced by identifying which base is being incorporated into the growing complementary strand of the target nucleic acid by the catalytic activity of the nucleic acid polymerizing enzyme at each step in the sequence of base additions. A polymerase on the target nucleic acid molecule complex is provided in a position suitable to move along the target nucleic acid molecule and extend the oligonucleotide primer at an active site. A plurality of labeled types of nucleotide analogs are provided proximate to the active site, with each distinguishable type of nucleotide analog being complementary to a different nucleotide in the target nucleic acid sequence. The growing nucleic acid strand is extended by using the polymerase to add a nucleotide analog to the nucleic acid strand at the active site, where the nucleotide analog being added is complementary to the nucleotide of the target nucleic acid at the active site. The nucleotide analog added to the oligonucleotide primer as a result of the polymerizing step is identified. The steps of providing labeled nucleotide analogs, polymerizing the growing nucleic acid strand, and identifying the added nucleotide analog are repeated so that the nucleic acid strand is further extended and the sequence of the target nucleic acid is determined.

[00206] In some embodiments, sequence analysis of the UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof may include a four-color sequencing by ligation scheme (degenerate ligation), which involves

hybridizing an anchor primer to one of four positions. Then an enzymatic ligation reaction of the anchor primer to a population of degenerate nonamers that are labeled with fluorescent dyes is performed. At any given cycle, the population of nonamers that is used is structure such that the identity of one of its positions is correlated with the identity of the fluorophore attached to that nonamer. To the extent that the ligase discriminates for complementarity at that queried position, the fluorescent signal allows the inference of the identity of the base. After performing the ligation and four-color imaging, the anchor primer:nonamer complexes are stripped and a new cycle begins. Methods to image sequence information after performing ligation are known in the art.

[00207] One or more UBA, ESB, COB, UBA/ESB complexes, UBA/ESB/COB complexes COB/ESB complexes and/or a combination thereof complexes can be detected and/or quantified by any method that detects and/or quantifies the presence of the assembled detection complex of interest. Such methods may include radioimmunoassay (RIA) or enzyme linked immunoabsorbance assay (ELISA), immunohistochemistry, immunofluorescent histochemistry with or without confocal microscopy, Raman spectroscopy, X-ray autoradiography, X-ray radiography, luminescence spectrometry, reversed phase assays, homogeneous enzyme immunoassays, and related non-enzymatic techniques, Western blots, whole cell staining, immunoelectron microscopy, nucleic acid amplification, gene array, protein array, mass spectrometry, patch clamp, 2-dimensional gel electrophoresis, differential display gel electrophoresis, microsphere-based multiplex protein assays, label-free cellular assays and flow cytometry, etc. U.S. Pat. No. 4,568,649 describes ligand detection systems, which employ scintillation counting. These techniques are particularly useful for modified protein parameters. Cell readouts for proteins and other cell determinants can be obtained using fluorescent or otherwise tagged reporter molecules. Microscopy methods are useful for measuring parameters in a morphological context. Flow cytometry methods are useful for measuring intracellular parameters.

[00208] When using fluorescent labeled components in the methods and compositions of the present invention, it will recognized that different types of fluorescent monitoring systems, e.g., Cytometric measurement device systems, can be used to practice the invention. In some embodiments, flow cytometric systems are used or systems dedicated to high throughput screening, e.g. 96 well or greater microtiter plates, e.g., Lakowicz, J. R., Principles of Fluorescence Spectroscopy, New York: Plenum Press (1983); Herman, B., Resonance energy transfer microscopy, in: Fluorescence Microscopy of Living Cells in Culture, Part B, Methods in Cell Biology, vol. 30, ed. Taylor, D. L. & Wang, Y.-L., San Diego: Academic Press (1989), pp.

219-243; Turro, N. J., Modern Molecular Photochemistry, Menlo Park: Benjamin/Cummings Publishing Col, Inc. (1978), pp. 296-361. Where the COBs/ESB complex is fluorescently labeled, suitable consideration of appropriate excitation sources may be investigated. Possible sources may include but are not limited to arc lamp, xenon lamp, lasers, light emitting diodes or some combination thereof. The appropriate excitation source is used in conjunction with an appropriate optical detection system, for example an inverted fluorescent microscope, an epi-fluorescent microscope or a confocal microscope. Preferably, a microscope is used that can allow for detection with enough spatial resolution to determine the sequence of the spots on the COB/ESB complexes. If for example, the COB/ESB complexes are labeled with three different colors, Alexa 488, Cy3 and Alexa 647 (labeled 1, 2 and 3, respectively). Colors 1, 2 and 3 are each acquired in different channels and the first and second registers, which can be seen as rows of spots, are shifted up by several pixels to be able to show each register individually. Examples of methods for detection of multiple colors that can be used in the methods of the invention are described in US patent 7,473,767, US patent publication no. 2007/0166708, US application number 11/645,270, and PCT application no US06/049274, incorporated by reference herein in its entirety.

[00209] Fluorescence in a sample can be measured using a fluorimeter. Other methods of detecting fluorescence may also be used, e.g., Quantum dot methods (see, e.g., Goldman et al., J. Am. Chem. Soc. (2002) 124:6378-82; Pathak et al. J. Am. Chem. Soc. (2001) 123:4103-4; and Remade et al., Proc. Natl. Sci. USA (2000) 18:553-8, each expressly incorporated herein by reference) as well as confocal microscopy.

[00210] In some embodiments, a FACS cell sorter (e.g. a FACSVantageTM, LSRII, or Canto Cell Sorter, Becton Dickinson Immunocytometry Systems, San Jose, Calif.) is used to sort and collect cells based on the presence or absence of an COB/ESB complex. By imparting an electromagnetic charge to droplets containing the positive cells, the cells can be separated from other cells. The positively selected cells can then be harvested in sterile collection vessels. These cell-sorting procedures are described in detail, for example, in the FACSVantageTM. Training Manual, with particular reference to sections 3-11 to 3-28 and 10-1 to 10-17, which is hereby incorporated by reference in its entirety for the above instruments.

[00211] In another embodiment, positive cells can be sorted using magnetic separation of cells based on the presence of a COB/ESB complex. In such separation techniques, cells to be positively selected are first contacted with a COB/ESB complex comprising retrievable particles (e.g., magnetically responsive particles). The cell can then be physically separated from non-positive or non-labeled cells, for example, using a magnetic field. When using magnetically

responsive particles, the positive or labeled cells can be retained in a container using a magnetic field while the negative cells are removed. These and similar separation procedures are described, for example, in the Baxter Immunotherapy Isolex training manual which is hereby incorporated in its entirety.

[00212] In some embodiments, one or more cells are contained in a well of a 96 well plate or other commercially available multi-well plate. In an alternate embodiment, the reaction mixture or cells are in a cytometric measurement device. Other multi-well plates useful in the present invention include, but are not limited to 384 well plates and 1536 well plates. Still other vessels for containing the reaction mixture or cells and useful in the present invention will be apparent to the skilled artisan.

[00213] In some embodiments, the abundance of a COB/ESB complex is measured using Inductively Coupled Plasma Mass Spectrometer (ICP-MS). A UBA that has been labeled with a specific element binds to the COB/ESB complex. When the cell is introduced into the ICP, it is atomized and ionized. The elemental composition of the cell, including the COB/ESB complex, is measured. The presence and intensity of the signals corresponding to the labels on the COB/ESB complex indicates the abundance of the COB/ESB complexes on that cell (Tanner et al. Spectrochimica Acta Part B: Atomic Spectroscopy, (2007), 62(3):188-195.).

[00214] In some embodiments the 'flow cytometer' is a microfluidic device where the cell measurements, or some of the measurements of the cell's contents, are carried out in channels devised to direct cells past detection devices in parallel sets of multiple channels. See U. S. Patent Nos. 7,378,280; 7,294,503; 7,294,298; and 6,830,936.

[00215] In some embodiments the cells, or some portion of their contents, are sonically encapsulated within individual droplets of liquid and interrogated with detection devices designed to measure each individual droplet's characteristics and the materials within such droplets.

[00216] Flexible hardware and software allows instrument adaptability for multiple applications. The software program modules allow creation, modification, and running of methods. The system diagnostic modules allow instrument alignment, correct connections, and motor operations. Customized tools, labware, and liquid, particle, cell and organism transfer patterns allow different applications to be performed. Databases allow method and parameter storage. Robotic and computer interfaces allow communication between instruments.

[00217] In some embodiments, the methods of the invention include the use of liquid handling components. The liquid handling systems can include robotic systems comprising any

number of components. In addition, any or all of the steps outlined herein may be automated; thus, for example, the systems may be completely or partially automated.

[00218] As will be appreciated by those in the art, there are a wide variety of components which can be used, including, but not limited to, one or more robotic arms; plate handlers for the positioning of microplates; automated lid or cap handlers to remove and replace lids for wells on non-cross contamination plates; tip assemblies for sample distribution with disposable tips; washable tip assemblies for sample distribution; 96 well loading blocks; cooled reagent racks; microtiter plate pipette positions (optionally cooled); stacking towers for plates and tips; and computer systems.

[00219] Fully robotic or microfluidic systems include automated liquid-, particle-, cell- and organism-handling including high throughput pipetting to perform all steps of screening applications. This includes liquid, particle, cell, and organism manipulations such as aspiration, dispensing, mixing, diluting, washing, accurate volumetric transfers; retrieving, and discarding of pipet tips; and repetitive pipetting of identical volumes for multiple deliveries from a single sample aspiration. These manipulations are cross-contamination-free liquid, particle, cell, and organism transfers. This instrument performs automated replication of microplate samples to filters, membranes, and/or daughter plates, high-density transfers, full-plate serial dilutions, and high capacity operation.

[00220] In some embodiments, chemically derivatized particles, plates, cartridges, tubes, magnetic particles, or other solid phase matrix with specificity to the assay components are used. The binding surfaces of microplates, tubes or any solid phase matrices include non-polar surfaces, highly polar surfaces, modified dextran coating to promote covalent binding, antibody coating, affinity media to bind fusion proteins or peptides, surface-fixed proteins such as recombinant protein A or G, nucleotide resins or coatings, and other affinity matrix are useful in this invention.

[00221] In some embodiments, platforms for multi-well plates, multi-tubes, holders, cartridges, minitubes, sonic levitation and encapsulation, deep-well plates, microfuge tubes, cryovials, square well plates, filters, chips, microchannel chips, microfluidics chips, optic fibers, beads, and other solid-phase matrices or platform with various volumes are accommodated on an upgradeable modular platform for additional capacity. This modular platform includes a variable speed orbital shaker, and multi-position work decks for source samples, sample and reagent dilution, assay plates, sample and reagent reservoirs, pipette tips, and an active wash station. In some embodiments, the methods of the invention include the use of a plate reader.

[00222] In some embodiments, thermocycler and thermoregulating systems are used for stabilizing the temperature of heat exchangers such as controlled blocks or platforms to provide accurate temperature control of incubating samples from 0° C. to 100° C.

[00223] In some embodiments, interchangeable pipet heads (single or multi-channel) with single or multiple magnetic probes, affinity probes, or pipetters robotically manipulate the liquid, particles, cells, and organisms. Multi-well or multi-tube magnetic separators or platforms manipulate liquid, particles, cells, and organisms in single or multiple sample formats.

[00224] In some embodiments, the instrumentation will include a detector, which can be a wide variety of different detectors, depending on the labels and assay. In some embodiments, useful detectors include a microscope(s) with multiple channels of fluorescence; plate readers to provide fluorescent, ultraviolet and visible spectrophotometric detection with single and dual wavelength endpoint and kinetics capability, fluorescence resonance energy transfer (FRET), luminescence, quenching, two-photon excitation, and intensity redistribution; CCD cameras to capture and transform data and images into quantifiable formats; and a computer workstation.

[00225] In some embodiments, the robotic apparatus includes a central processing unit which communicates with a memory and a set of input/output devices (e.g., keyboard, mouse, monitor, printer, etc.) through a bus. Again, as outlined below, this may be in addition to or in place of the CPU for the multiplexing devices of the invention. The general interaction between a central processing unit, a memory, input/output devices, and a bus is known in the art. Thus, a variety of different procedures, depending on the experiments to be run, are stored in the CPU memory.

[00226] These robotic fluid handling systems can utilize any number of different reagents, including buffers, reagents, samples, washes, assay components such as label probes, etc.

Applications for Target Molecule Detection

[00227] The compositions and methods of the invention can be used for diagnostic, prognostic, therapeutic, patient stratification, drug development, treatment selection, and screening purposes. The present invention provides the advantage that many different target molecules can be analyzed at one time from a single biomolecular sample using the methods of the invention. This allows, for example, for several diagnostic tests to be performed on one sample.

[00228] The composition and methods of the invention can be used in proteomics. The methods described herein will typically provide an answer rapidly which is very desirable for

this application. The methods and composition described herein can be used in the process of finding biomarkers that may be used for diagnostics or prognostics and as indicators of health and disease. The methods and composition described herein can be used to screen for drugs, e.g., drug development, selection of treatment, determination of treatment efficacy and/or identify targets for pharmaceutical development. The ability to test protein expression on screening assays involving drugs is very important because proteins are the final gene product in the body. In some embodiments, the methods and compositions described herein will measure both protein and gene expression simultaneously which will provide the most information regarding the particular screening being performed.

[00229] The composition and methods of the invention can be used in gene expression analysis. The methods described herein discriminate between nucleotide sequences. The difference between the target nucleotide sequences can be, for example, a single nucleic acid base difference, a nucleic acid deletion, a nucleic acid insertion, or rearrangement. Such sequence differences involving more than one base can also be detected. In some embodiments, the UBAs, e.g., oligonucleotide probe, have substantially the same length so that they hybridize to target nucleotide sequences at substantially similar hybridization conditions. As a result, the process of the present invention is able to detect infectious diseases, genetic diseases, and cancer. It is also useful in environmental monitoring, forensics, and food science. Examples of genetic analyses that can be performed on nucleic acids include e-g., SNP detection, STR detection, RNA expression analysis, promoter methylation, gene expression, virus detection, viral subtyping and drug resistance.

[00230] The present methods can be applied to the analysis of biomolecular samples obtained or derived from a patient so as to determine whether a diseased cell type is present in the sample, the stage of the disease, the prognosis for the patient, the ability to the patient to respond to a particular treatment, or the best treatment for the patient. The present methods can also be applied to identify biomarkers for a particular disease.

[00231] In some embodiments, the methods described herein are used in the diagnosis of a condition. As used herein the term "diagnose" or "diagnosis" of a condition includes predicting or diagnosing the condition, determining predisposition to the condition, monitoring treatment of the condition, diagnosing a therapeutic response of the disease, and prognosis of the condition, condition progression, and response to particular treatment of the condition. For example, a blood sample can be assayed according to any of the methods described herein to determine the presence and/or quantity of markers of a disease or malignant cell type in the sample, thereby diagnosing or staging the a disease or a cancer.

[00232] In some embodiments, the methods and composition described herein are used for the diagnosis and prognosis of a condition.

[00233] Numerous immunologic, proliferative and malignant diseases and disorders are especially amenable to the methods described herein. Immunologic diseases and disorders include allergic diseases and disorders, disorders of immune function, and autoimmune diseases and conditions. Allergic diseases and disorders include but are not limited to allergic rhinitis, allergic conjunctivitis, allergic asthma, atopic eczema, atopic dermatitis, and food allergy. Immunodeficiencies include but are not limited to severe combined immunodeficiency (SCID), hypereosinophilic syndrome, chronic granulomatous disease, leukocyte adhesion deficiency I and II, hyper IgE syndrome, Chediak Higashi, neutrophilias, neutropenias, aplasias, Agammaglobulinemia, hyper-IgM syndromes, DiGeorge/Velocardial-facial syndromes and Interferon gamma-TH1 pathway defects. Autoimmune and immune dysregulation disorders include but are not limited to rheumatoid arthritis, diabetes, systemic lupus erythematosus, Graves' disease, Graves ophthalmopathy, Crohn's disease, multiple sclerosis, psoriasis, systemic sclerosis, goiter and struma lymphomatosa (Hashimoto's thyroiditis, lymphadenoid goiter), alopecia aerata, autoimmune myocarditis, lichen sclerosis, autoimmune uveitis, Addison's disease, atrophic gastritis, myasthenia gravis, idiopathic thrombocytopenic purpura, hemolytic anemia, primary biliary cirrhosis, Wegener's granulomatosis, polyarteritis nodosa, and inflammatory bowel disease, allograft rejection and tissue destructive from allergic reactions to infectious microorganisms or to environmental antigens.

[00234] Proliferative diseases and disorders that may be evaluated by the methods of the invention include, but are not limited to, hemangiomatosis in newborns; secondary progressive multiple sclerosis; chronic progressive myelodegenerative disease; neurofibromatosis; ganglioneuromatosis; keloid formation; Paget's Disease of the bone; fibrocystic disease (e.g., of the breast or uterus); sarcoidosis; Peronies and Duputren's fibrosis, cirrhosis, atherosclerosis and vascular restenosis.

[00235] Malignant diseases and disorders that may be evaluated by the methods of the invention include both hematologic malignancies and solid tumors.

[00236] Hematologic malignancies are especially amenable to the methods of the invention when the sample is a blood sample, because such malignancies involve changes in blood-borne cells. Such malignancies include non-Hodgkin's lymphoma, Hodgkin's lymphoma, non-B cell lymphomas, and other lymphomas, acute or chronic leukemias, polycythemias, thrombocythemias, multiple myeloma, myelodysplastic disorders, myeloproliferative disorders, myelofibroses, atypical immune lymphoproliferations and plasma cell disorders.

[00237] Plasma cell disorders that may be evaluated by the methods of the invention include multiple myeloma, amyloidosis and Waldenstrom's macroglobulinemia.

[00238] Example of solid tumors include, but are not limited to, colon cancer, breast cancer, lung cancer, prostate cancer, brain tumors, central nervous system tumors, bladder tumors, melanomas, liver cancer, osteosarcoma and other bone cancers, testicular and ovarian carcinomas, head and neck tumors, and cervical neoplasms.

[00239] Genetic diseases can also be detected by the process of the present invention. This can be carried out by prenatal or post-natal screening for chromosomal and genetic aberrations or for genetic diseases. Examples of detectable genetic diseases include: 21 hydroxylase deficiency, cystic fibrosis, Fragile X Syndrome, Turner Syndrome, Duchenne Muscular Dystrophy, Down Syndrome or other trisomies, heart disease, single gene diseases, HLA typing, phenylketonuria, sickle cell anemia, Tay-Sachs Disease, thalassemia, Klinefelter Syndrome, Huntington Disease, autoimmune diseases, lipidosis, obesity defects, hemophilia, inborn errors of metabolism, and diabetes.

[00240] The methods described herein can be used to diagnose pathogen infections, for example infections by intracellular bacteria and viruses, by determining the presence and/or quantity of markers of bacterium or virus, respectively, in the sample.

[00241] A wide variety of infectious diseases can be detected by the process of the present invention. Typically, these are caused by bacterial, viral, parasite, and fungal infectious agents. The resistance of various infectious agents to drugs can also be determined using the present invention.

[00242] Bacterial infectious agents which can be detected by the present invention include Escherichia coli, Salmonella, Shigella, KlESBiella, Pseudomonas, Listeria monocytogenes, Mycobacterium tuberculosis, Mycobacterium aviumintracellulare, Yersinia, Francisella, Pasteurella, Brucella, Clostridia, Bordetella pertussis, Bacteroides, Staphylococcus aureus, Streptococcus pneumonia, B-Hemolytic strep., Corynebacteria, Legionella, Mycoplasma, Ureaplasma, Chlamydia, Neisseria gonorrhea, Neisseria meningitides, Hemophilus influenza, Enterococcus faecalis, Proteus vulgaris, Proteus mirabilis, Helicobacter pylori, Treponema palladium, Borrelia burgdorferi, Borrelia recurrentis, Rickettsial pathogens, Nocardia, and Acitnomycetes.

[00243] Fungal infectious agents which can be detected by the present invention include Cryptococcus neoformans, Blastomyces dermatitidis, Histoplasma capsulatum, Coccidioides immitis, Paracoccidioides brasiliensis, Candida albicans, Aspergillus fumigautus, Phycomycetes (Rhizopus), Sporothrix schenckii, Chromomycosis, and Maduromycosis.

[00244] Viral infectious agents which can be detected by the present invention include human immunodeficiency virus, human T-cell lymphocytotrophic virus, hepatitis viruses (e.g., Hepatitis B Virus and Hepatitis C Virus), Epstein-Barr Virus, cytomegalovirus, human papillomaviruses, orthomyxo viruses, paramyxo viruses, adenoviruses, corona viruses, rhabdo viruses, polio viruses, toga viruses, bunya viruses, arena viruses, rubella viruses, and reo viruses.

[00245] Parasitic agents which can be detected by the present invention include Plasmodium falciparum, Plasmodium malaria, Plasmodium vivax, Plasmodium ovale, Onchoverva volvulus, Leishmania, Trypanosoma spp., Schistosoma spp., Entamoeba histolytica, Cryptosporidum, Giardia spp., Trichimonas spp., Balatidium coli, Wuchereria bancrofti, Toxoplasma spp., Enterobius vermicularis, Ascaris lumbricoides, Trichuris trichiura, Dracunculus medinesis, trematodes, Diphyllobothrium latum, Taenia spp., Pneumocystis carinii, and Necator americanis.

[00246] The present invention is also useful for detection of drug resistance by infectious agents. For example, vancomycin-resistant Enterococcus faecium, methicillin-resistant Staphylococcus aureus, penicillin-resistant Streptococcus pneumoniae, multi-drug resistant Mycobacterium tuberculosis, and AZT-resistant human immunodeficiency virus can all be identified with the present invention

[00247] Thus, the target molecules detected using the compositions and methods of the invention can be either patient markers (such as a cancer marker) or markers of infection with a foreign agent, such as bacterial or viral markers.

[00248] Because of the quantitative nature of UBA/ESB/COBs, the compositions and methods of the invention can be used to quantify target molecule whose abundance is indicative of a biological state or disease condition, for example, blood markers that are upregulated or downregulated as a result of a disease state.

[00249] In some embodiments, the methods and compositions of the present invention can be used for cytokine detection. The low sensitivity of the methods described herein would be helpful for early detection of cytokines, e.g., as biomarkers of a condition, diagnosis or prognosis of a disease such as cancer, and the identification of subclinical conditions.

Kits

[00250] The invention further provides kits comprising one or more components of the invention. The kits can comprise, for example, one or more UBAs, one or more ESBs, and/or one or more APSs. The kits can be used for any purpose apparent to those of skill in the art, including those described above.

[00251] In certain embodiments, the present invention also provides kits useful for the extension and selective immobilization of COBs, UBAs, ESBs and/or a combination thereof. The kits can comprise a substrate for immobilization and one or more binding partners to facilitate extension or immobilization of a COB, UBA, ESB and/or a combination thereof. The binding partners could, in certain embodiments, comprise a moiety useful for extension of the COB, UBA, ESB and/or a combination thereof, in an appropriate force. In certain embodiments, the binding partners could facilitate immobilization or selective immobilization of the COB, UBA, ESB and/or a combination thereof, to the surface. In further embodiments, the kits could comprise a device capable of extending the COB, UBA, ESB and/or a combination thereof.

[00252] The kits can contain a population of COBs, APSs, UBAs, ESBs and/or a combination thereof as described herein.

[00253] The kits can contain pre-labeled APSs, or unlabeled APSs with one or more components for labeling the APSs. Moreover, the ESBs and/or APSs provided in a kit may or may not have UBAs pre-attached. In one embodiment, the UBAs are provided in the kit unattached to the ESBs and/or APSs.

[00254] The kits can comprise other reagents such as linker oligos and bridging oligos. In some embodiments, the kits can separate the UBAs into different premixes.

[00255] The kits can include other reagents as well, for example, buffers for performing hybridization reactions, linkers, restriction endonucleases, and DNA I ligases.

[00256] The kits also will include instructions for using the components of the kit, and/or for making and/or using the APSs, COBs, UBAs, and/or ESBs.

EXAMPLES

Prophetic Example 1 – Oligonucleotide preparation

[00257] Oligonucleotides can be synthesized according to standard techniques known in the art. For instance, oligonucleotides can be synthesized on a 394A DNA Synthesizer (Applied Biosystems Division of Perkin-Elmer Corp., Foster City, Calif.).

[00258] The oligonucleotides are purified by ethanol precipitation after overnight deprotection at 55° C. The primer-specific portions of the oligonucleotides used for PCR amplification are purified by polyacrylamide gel electrophoresis on 10% acrylamide/7M urea gels. Oligonucleotides are visualized after electrophoresis by UV shadowing against a lightening screen and excised from the gel (Applied Biosystems Inc., 1992). They are then eluted overnight at 64° C. in TNE (i.e. Tris-sodium EDTA) buffer (100 mM Tris/HCl pH 8.0 containing 500 mM NaCl and 5 mM EDTA) and recovered from the eluate using Sep Pak cartridges (Millipore Corp, Milford, Mass.) following the manufacturer's instructions.

[00259] Oligonucleotides are resuspended in 100 µl TE (i.e. 10 mM Tri-HCl pH 8.0 containing 1 mM EDTA). Typical concentrations of these original oligonucleotide solutions are about 1 µg/µl or approximately 74 pm/µl.

[00260] As a prerequisite for ligation reactions, the oligonucleotides are phosphorylated with T4 polynucleotide kinase at the 5'-end. Aliquots of the oligonucleotides equivalent to 200 pm are combined with 10 µl of 10x kinase buffer (500 mM Tris/HCl pH 8.0, 100 mM $MgCl_2$), 10 µl of 10 mM ATP, 20 U T4 kinase, and sufficient water-ME to give a final volume of 100 µl. Phosphorylation is carried out at 37° C. for 30 min followed by incubation for 10 min at 85° C to inactivate the T4 enzyme.

[00261] The solutions of the oligonucleotides are adjusted to convenient concentrations. The kinased oligonucleotide solution is diluted fourfold in water to yield a concentration of 1000 fm/µl. A solution of the oligonucleotides is made by combining volumes of the oligonucleotides equivalent to 200 pm with sufficient water to give a final volume of 400 µl. This created a solution 1000 fm/µl in each of the oligonucleotides. Aliquots (20 µl) of the kinased and unkinased oligonucleotides are frozen for subsequent use.

General method for oligonucleotide synthesis and purification for Click chemistry

[00262] Oligonucleotides are synthesized as described in El-Sagheer et al. (PNAS, 108:28, 11338-11343, 2011). Briefly, standard DNA phosphoramidites, solid supports and additional reagents are purchased from Link Technologies and Applied Biosystems. Oligonucleotides are synthesized on an Applied Biosystems 394 automated DNA/ RNA synthesizer using a standard 0.2 or 1.0 μmole phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. All β-cyanoethyl phosphoramidite monomers are dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers is 35 sec, whereas the coupling time for the reverse amidites is 180 sec. Alkyne phosphoramidite monomer (**2c** in **Fig. 2,** El-Sagheer et al., PNAS, 108:28, 11338-11343, 2011) and other non-standard monomers are coupled for 360 sec. Cleavage of oligonucleotides from the solid support and deprotection is achieved by exposure to concentrated aqueous ammonia solution for 60 min at room temperature followed by heating in a sealed tube for 5 hr at 55 °C. The oligonucleotides are purified by reversed-phase HPLC on a Gilson system using an XBridgeTM BEH300 Prep C18 10 µM 10x250 mm column (Waters) with a gradient of acetonitrile in ammonium acetate (0% to 50% buffer B over 30 min, flow rate 4 mL/min), buffer A: 0.1 M ammonium acetate, pH 7.0, buffer B: 0.1 M ammonium acetate, pH 7.0, with 50% acetonitrile. Elution is monitored by UV absorption at 305 or 295 nm. After HPLC purification, oligonucleotides are desalted using NAP-10 columns and analyzed by gel electrophoresis.

i) Synthesis of 3'-alkyne oligonucleotides

[00263] Synthesis of 3'-alkyne oligonucleotides is performed as described in El-Sagheer et al. (PNAS, 108:28, 11338-11343, 2011). Briefly, 3'-Alkyne oligonucleotides are synthesized using the 3'-propargylthymidine phosphoramidite monomer **2c** and assembling the required sequence in the 5' to 3'-direction using the 3'-*O*-(4,4'-dimethoxytrityl)deoxyribonucleoside-5'-phosphor-amidites of A, G, C and T (reverse phosphoramidites, Link Technologies) or by the attachment of 5'-*O*-(4,4'-dimethoxytrityl)-3'-*O*-propargyl-5-methyl-deoxycytidine to solid support (33 μmol/g loading, AM polystyrene, Applied Biosystems) according to El-Sagheer et al. (*Proc. Natl. Acad. Sci. U. S. A.* 107(35):15329-15334.). The resin is packed into a twist column (Glen Research), then used to assemble the required sequence in the 3'- to 5'-direction by standard phosphoramidite oligonucleotide synthesis. The oligonucleotides are then cleaved, deprotected and purified as described above.

ii) Synthesis of 5'-azide oligonucleotides

[00264] Synthesis of 5'-azide oligonucleotides is performed as described in El-Sagheer et al. (PNAS, 108:28, 11338-11343, 2011). Briefly, oligonucleotides are assembled on the 0.2 or 1.0 μmol scale (trityl-off) as described in the general method (above) with normal 5'-HO-dC, 5'-HO-dT (or with 5'-iodo-dT using the commercially available 5'-iodo dT monomer from Glen Research). To convert the 5'-hydroxyl group to 5'-iodo, the protected oligomers attached to the synthesis column are treated with a 0.5 M solution of methyltriphenoxyphosphonium iodide in DMF (1.0 mL), which is periodically passed through the column *via* two 1 mL syringes over 15 min at room temperature. The column is then washed several times with dry DMF. To convert the 5'-iodo (dT or dC) to 5'-azido (dT or dC), sodium azide (50 mg) is suspended in dry DMF (1 mL), heated for 10 min at 70 °C then cooled down and the supernatant taken up into a 1 mL syringe, passed back and forth through the column then left at room temperature overnight (or for 5 hr at 55 °C). The column is then washed with DMF and acetonitrile and dried by the passage of a stream of argon gas. The resultant 5'-azide oligonucleotide is cleaved from the solid support, deprotected and purified as described above.

iii) Synthesis of 3'-alkyne-5'-azide oligonucleotides

[00265] Synthesis of 3'-alkyne-5'azide oligonucleotides is performed as described in El-Sagheer et al. (PNAS, 108:28, 11338-11343, 2011). Briefly, 5'-*O*-(4,4'-Dimethoxytrityl)-3'-*O*-propargyl-5-methyldeoxycytidine on polystyrene solid support is packed into a twist column (Glen Research) and used to assemble the required sequence in the 3'- to 5'-direction (standard

phosphoramidite oligonucleotide synthesis) with 5′-iodo dT, 5′-HO-dT or 5′-HO-dC at the 5′-end. The 5′-hydroxyl or iodo groups are then converted to azide using the conditions described above for the synthesis of the 5′-azide oligonucleotides.

Prophetic Example 2. Click chemistry ligation

[00266] Oligonucleotide APSs are annealed to a template and kept overnight at 4° C. A solution of Cu^{I} click catalyst is prepared from *tris*-hydroxypropyltriazole ligand as described in Chan et al. (Chan TR, Hilgraf R, Sharpless KB, & Fokin VV (2004) Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Org. Lett.* 6(17):2853-2855; 2.8 μmol in 0.2 M NaCl, 38.0 μL), sodium ascorbate (4.0 μmol in 0.2 M NaCl, 8.0 μL) and CuSO4.5H2O (0.4 μmol in 0.2 M NaCl, 4.0 μL). This solution is added to the annealed oligonucleotides and the reaction mixture is kept at 0 °C for 1 hr, then at room temperature for a further 1 hr. Reagents are removed using a NAP-25 gel-filtration column.

Prophetic Example 3. Split-pool synthesis of COBs on beads

[00267] In this Example COBs are synthesized attached to beads. 4 different methods are used for the assembly of APSs into COBs:

[00268] **Patchwork COB (Figure 6)**

[00269] Aminomethyl macroporous polystyrene (MPPS) beads are labeled with ten different CL oligonucleotides, each with an optional first amplification primer complementary region, one of 10 different ESB sequences and a common annealing region. Six rounds of split pool synthesis are performed. In each round, beads are split into 20 different containers. A different oligonucleotide APS is added to each container, totaling 20 different APSs. Each APS in a given round further comprises a unique subcode sequence that is different from the rest of the APSs in that round.

[00270] In the first round, each APS comprises an annealing region 1 that is complementary to the annealing region of the CL oligonucleotide on one end and an annealing region 2 on the other end. Upon addition, the oligonucleotide APS hybridizes to the CL oligonucleotide along the complementary annealing region 1. The annealing region 2 remains single stranded and available to hybridization with an APS added in the subsequent round. In subsequent rounds, each APS comprises an annealing region complementary to the available annealing region of the APS from the previous round on one end and an additional annealing region on the other end. The added APS hybridizes to the APS added in the previous round along the complementary annealing region.

[00271] The last subunit optionally comprises a second amplification primer complimentary region for hybridization of PCR or sequencing primers.

[00272] A CL, or one or more APSs further comprise a random tag region, which acts as a molecular counter as described supra, allowing for subsequent normalization of the detected COBs.

[00273] Upon the addition of an APS in each round, the beads are pooled and divided into new 20 pools initiating the subsequent round. A new set of 20 APSs are added in each round with a pair of round specific annealing region as described above. After the addition of 6 APSs, the hybridized APSs on the beads are patched together using a polymerase/ligase. The COBs are optionally PCR amplified for sequencing using primers targeting the amplification primer complementary regions on the CL and the last APS subunit

[00274] **Stitch COB using specific annealing of primers (Figure 7)**

[00275] Aminomethyl macroporous polystyrene (MPPS) beads are labeled with ten different CL oligonucleotides, each with an optional first amplification primer complementary region, a one of 10 different ESB sequences and a common annealing region. Six rounds of split pool synthesis are performed. In each round, beads are split into 20 different containers. A different oligonucleotide APS is added to each container, totaling 20 different APSs. Each APS in a given round further comprises a unique subcode sequence that is different from the rest of the APSs in that round.

[00276] An annealing primer is also added. In the first round, the annealing primer has a complementary region to the CL oligonucleotide and a complementary region to the APS. The annealing primer hybridizes to both, stitching them together. In subsequent rounds, the annealing primer has a complementary region to the APS added during the previous round and a complementary region to the APS being added in the current round. Similarly, the annealing primer hybridizes to APSs from subsequent rounds stitching them together. The complementary regions of the annealing primer are specific to each round allowing efficient hybridization of subunits only from the previous and current rounds. Accordingly, the annealing primer does not hybridize to subunits of earlier rounds, which would not have complementary regions to the annealing primer of a current round, thus blocking the further synthesis of COBs missing subunits of particular rounds.

[00277] The last subunit optionally comprises a second amplification primer complimentary region for hybridization of PCR or sequencing primers.

[00278] A CL, or one or more APSs further comprise a random tag region, which acts as a molecular counter as described supra, allowing for subsequent normalization of the detected COBs.

[00279] Upon the addition of an APS and an annealing primer in each round, the beads are pooled and divided into new 20 pools initiating the subsequent round. A new set of 20 APSs are added in each round with a pair of round specific annealing region as described above. After the addition of 6 APSs, the hybridized APSs on the beads are permanently stitched together using a polymerase/ligase or using Click chemistry as described in *Example 2*. The COBs are optionally PCR amplified for sequencing using primers targeting the amplification primer complementary regions on the CL and the last APS subunit.

[00280] Stitch COB using annealing of primers with common complementary regions (Figure 8)

[00281] Aminomethyl macroporous polystyrene (MPPS) beads are labeled with ten different CL oligonucleotides, each with an optional first amplification primer complementary region, a one of 10 different ESB sequences and a common annealing region. Six rounds of split pool synthesis are performed. In each round, beads are split into 20 different containers. A different oligonucleotide APS is added to each container, totaling 20 different APSs. Each APS in a given round further comprises a unique subcode sequence that is different from the rest of the APSs in that round.

[00282] An annealing primer is also added. In the first round, the annealing primer has a first complementary region to the CL oligonucleotide and a second complementary region to the APS being added in the current round. The annealing primer hybridizes to both, stitching them together. In subsequent rounds, the annealing primer has a first complementary region to the APS added during the previous round and a second complementary region to the APS being added in the current round. Similarly, the annealing primer hybridizes to APSs from subsequent rounds stitching them together. Of the two complementary regions of the annealing primer, the first complementary regions are specific to each round allowing efficient hybridization to subunits only from the previous round. Accordingly, the annealing primer does not hybridize to subunits of earlier rounds, which would not have complementary regions to the annealing primer of a current round, thus blocking the further synthesis of COBs missing subunits of particular rounds.

[00283] The last subunit optionally comprises a second amplification primer complimentary region for hybridization of PCR or sequencing primers.

[00284] A CL, or one or more APSs further comprise a random tag region, which acts as a molecular counter as described supra, allowing for subsequent normalization of the detected COBs.

[00285] Upon the addition of an APS and an annealing primer in each round, the beads are pooled and divided into new 20 pools initiating the subsequent round. A new set of 20 APSs are added in each round with a pair of round specific annealing region as described above. After the addition of 6 APSs, the hybridized APSs on the beads are permanently stitched together using a polymerase/ligase or using Click chemistry as described in *Example 2*. The COBs are optionally PCR amplified for sequencing using primers targeting the amplification primer complementary regions on the CL and the last APS subunit.

[00286] **Loop COB (Figure 9)**

[00287] Aminomethyl macroporous polystyrene (MPPS) beads are labeled with ten different CL oligonucleotides, each with an optional first amplification primer complementary region, one of 10 different ESB sequences, six pairs of APS-specific loop annealing regions and an optional second amplification primer complementary region. Six rounds of split pool synthesis are performed. In each round, beads are split into 20 different containers. A different oligonucleotide APS is added to each container, totaling 20 different APSs. Each APS in a given round further comprises a unique subcode sequence that is different from the rest of the APSs in that round.

[00288] The APSs are designed to hybridize to the CL in a loop geometry, hybridizing on each end to the CL along the loop annealing regions specific to the round. The hybridization populates the APSs along the CL, which are then linked together. The APSs are designed such that they do not efficiently hybridize to the CL along the loop annealing regions specific to other rounds. Consequently, if an APS from a particular round is missing, the APSs may not be linked together successfully, depending on the linking process. Alternatively, a COB is synthesized with a missing APS, the location of which is flanked by a pair of loop annealing regions. The resulting COB can then be analyzed accordingly and can either be discarded or the retrieved information can be alternatively processed.

[00289] A CL, or one or more APSs further comprise a random tag region, which acts as a molecular counter as described supra, allowing for subsequent normalization of the detected COBs.

[00290] Upon the addition of an APS and an annealing primer in each round, the beads are pooled and divided into new 20 pools initiating the subsequent round. A new set of 20 APSs

are added in each round with a pair of round specific annealing region as described above. After the addition of 6 APSs, the hybridized APSs on the beads are permanently stitched together using a polymerase/ligase or using Click chemistry as described in *Example 1*. The COBs are optionally PCR amplified for sequencing using primers targeting the amplification primer complementary regions on the CL.

[00291] Polymerase free COB-ESB linkage (Figure 10)

[00292] Aminomethyl macroporous polystyrene (MPPS) beads are labeled with ten different CL oligonucleotides, each with a pair of loop annealing region specific for one of the 10 different loop ESB sequences, and six pairs of APS-specific loop annealing regions. All 10 loop ESB sequences are added to anneal to the loop ESB specific portion of the CL in a loop geometry. Loop ESB sequences are designed to minimize non-specific annealing to the remainder of the loop ESB specific regions of the CLs. Loop ESB sequences comprise an optional first amplification primer complementary region, an ESB sequence, and a pair of annealing regions sufficiently complementary to the loop ESB-specific loop annealing regions in the CL. Six rounds of split pool synthesis are performed. In each round, beads are split into 20 different containers. A different oligonucleotide APS is added to each container, totaling 20 different APSs. Each APS in a given round further comprises a unique subcode sequence that is different from the rest of the APSs in that round. The APSs in the final round optionally further comprise a second amplification primer complementary region.

[00293] The APSs are designed to hybridize to the CL in a loop geometry, hybridizing on each end to the CL along the loop annealing regions specific to the round. The hybridization populates the APSs along the CL, which are then linked together. The APSs are designed such that they do not efficiently hybridize to the CL along the loop annealing regions specific to other rounds. Consequently, if an APS from a particular round is missing, the APSs may not be linked together successfully, depending on the linking process. Alternatively, a COB is synthesized with a missing APS, the location of which is flanked by a pair of loop annealing regions. The resulting COB can then be analyzed accordingly and can either be discarded or the retrieved information can be alternatively processed.

[00294] A CL, or one or more APSs further comprise a random tag region, which acts as a molecular counter as described supra, allowing for subsequent normalization of the detected COBs.

[00295] Upon the addition of an APS and an annealing primer in each round, the beads are pooled and divided into new 20 pools initiating the subsequent round. A new set of 20 APSs

are added in each round with a pair of round specific annealing region as described above. After the addition of 6 APSs, the hybridized APSs on the beads are permanently stitched together using a using Click chemistry as described in *Example 2*. The COBs are optionally PCR amplified for sequencing using primers targeting the amplification primer complementary regions on the CL and the last APS subunit.

Prophetic Example 4. Detection by nucleic acid sequencing

[00296] The assembled ESB-linked COBs resulting from any of the methods in *Example 3* are sequenced by Illumina's HiSeq 2000 machine. The resulting sequences comprise at least one of 10 different ESB sequences, a random tag region, and a combination of 6 subcodes originating from APSs added to that particular bead during the 6 rounds of split pool synthesis.

Prophetic Example 5. Detection by peptide sequencing (Figure 11)

[00297] ESB-linked COBs are synthesized using any of the methods in *Example 3*. The resulting sequences comprise a T7 promoter an SP6 start site, a start codon, an ESB, a COB and optionally a region encoding a His(6) tag **(Figure 11)**. The T7 promoter and SP6 start site can be incorporated into the sequence linked to the ESB, using the same method that is used to incorporated the ESB. Alternatively, these sequences can be incorporated withing the last APS. Optionally, a His(6) tag encoding region can be incorporated linked to the final APS or to the ESB.

[00298] The assembled ESB-linked COBs are transcribed and translated into peptide sequences using the Expressway™ Maxi Cell-Free E. coli Expression System (Invitrogen). The peptide sequences are isolated using affinity chromatography and/or HPLC prior to being sequenced using a tandem mass spectrometer.

Prophetic Example 6. Split-pool synthesis of COBs on cell surfaces

[00299] Cell surface receptors on leukocyte cell lines (HL60, JY and U937) are detected and quantified using split-pool synthesis of COBs on cell surfaces. Using Antibody-Oligonucleotide All-in-One Conjugation Kit (Solulink), antibodies against CD1, CD3, CD8 and CD4 are conjugated with amine modified CL oligonucleotides described in *Example 3*. The singly-labeled antibodies are isolated using affinity chromatography using complementary oligonucleotides targeting a sequence in CL oligonucleotides and the number of labels on each antibody is verified using mass spectrometry. A cell suspension of 10^7 cells are incubated with the combination of the antibodies under suitable conditions followed by 6 rounds of split-pool synthesis. The resulting ESB-linked COBs are detected as described in *Example 3* or *Example 4*.

The detected signals related to COB-linked ESBs are quantified for each COB combination. Co-expression of each of the CD1, CD3, CD8, and CD4 antigens on the cells are plotted pairwise. Principle component analysis is used to identify the strongest correlations in expression profiles.

Prophetic Example 7. Split-pool synthesis of COBs in cells

[00300] Methanol is cooled to -20 °C. A cell culture comprising 10^7 HeLa cells is grown using suitable tissue culture conditions known in the art. The growth medium is removed by aspiration.The cells are immediately fixed and permeabilized by adding 50 mL cold methanol. The cells are allowed to incubate at ambient temperature for 10 minutes with gentle shaking. Methanol is carefully remived by aspiration. The cells are rinsed with 100 mL of 1X PBS, three times.

[00301] The cells are blocked with 150 ml of 0.1% Casein solution in 0.2% PBS for 1.5 hrs at room temperature by gentle shaking. Rabbit anti-cleaved caspase-3, rabbit anti phos-p38, rabbit anti-phos-ERK2, mouse anti-ERK2, and mouse anti-β-tubulin (clone AA2) are CL-conjugated as described in *Example 5*. Cells are incubated with CL-conjugated antibodies overnight at 4 °C, gently shaking. The cells are washed 5 times with 1X PBS + 0.1% Tween-20 for 5 minutes at room temperature, followed by 6 rounds of split-pool synthesis.

[00302] The resulting ESB-linked COBs are detected as described in *Example 3* or *Example 4*. The detected signals related to COB-linked ESBs are quantified for each COB combination. Co-expression of each of the Phospho-p53, ERK1 in the cells are plotted pairwise. Principle component analysis is used to identify the strongest correlations in expression profiles.

[00303] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for identifying whether a plurality of targets are in a plurality of cells comprising: binding to the targets a plurality of tags, wherein a tag comprises a code that represents a) the target identity and b) the identity of the cell in which tag is binding.

2. The method of claim 1, wherein individual cell separation or isolation is unnecessary for the binding step.

3. A method for identifying a single cell associated with a target comprising:
binding to the target a tag, wherein the tag comprises a code that represents

a) the target, and

b) the single cell;

wherein the during the binding the single cell is not isolated from a population of cells,

and

wherein the code that represents the single cell is unknown before the binding.

4. The method of claims 1-3, further comprising detecting the code, wherein individual cell separation or isolation is unnecessary for the detecting step.

5. The method of claims 1-4, wherein each target is a protein or a nucleic acid.

6. The method of claims 1-5, wherein the tag is a nucleic acid.

7. The method of claims 1-6, wherein the tag comprises a UBA.

8. The method of claims 1-7, wherein the UBA is specific for one of the targets.

9. The method of claims 1-8, wherein the tag comprises a UBA.

10. The method of claims 1-9, wherein the UBA comprises an antibody.

11. The method of claims 1-10, wherein the tag comprises a ESB.

12. The method of claims 1-11, wherein the ESB comprises a common linker (CL).

13. The method of claims 1-12, wherein the ESB codes the target identity.

14. The method of claims 1-13, wherein the ESB comprises a nucleic acid.

15. The method of claims 1-14, wherein the tag comprises an APS.

16. The method of claims 1-15, wherein the APS is detectable a detectably distinct coding unit.

17. The method of claims 1-16, wherein during the binding step multiple APSs are added to the tag in an ordered manner during successive rounds of split pool synthesis.

18. The method of claims 1-17, wherein the tag comprises at least 10 APSs.

19. The method of claims 1-18, wherein the APS comprises a nucleic acid.

20. The method of claims 1-19, wherein the tag comprises multiple APSs, an ESB, and a UBA linked by ligation.

21. The method of claims 1-20, wherein multiple APSs, the ESB, and/or the UBA is capable of being linked Click chemistry.

22. The method of claims 1-21, wherein the APS or the ESB comprises an amplification primer binding region.

23. The method of claims 1-22, wherein the UBA, ESB, or APS is templatable.

24. A composition comprising:

a) a first target molecule,

b) a first unique binding agent (UBA) specific for the first target molecule,

c) a first linkable UBA-dependent epitope specific barcode (ESB), and

d) a plurality of ordered assayable polymer subunit (APS), wherein the order of APSs is detectable.

25. The composition of claim 24, wherein the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid.

26. A composition comprising a population of particles each comprising at least a first target molecule, wherein the first target molecule is associated with:

(a) a first unique binding agent (UBA) specific for the first target molecule,

(b) a first linkable UBA-dependent epitope specific barcode (ESB), and

(c) a first plurality of ordered assayable polymer subunits (APS),
wherein the plurality of ordered APSs associated with the first target molecule of a first particle in the population is detectably different than the plurality of ordered APSs associated with the first target molecule of a second particle in the population.

27. The composition of claims 24-26, wherein the plurality of ordered APSs comprises 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 APSs.

28. The composition of claims 24-27, wherein the plurality of ordered APSs comprises more than 20 APSs.

29. The composition of claims 24-28, wherein the APSs are templatable.

30. The composition of claims 26-29, wherein at least one particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead.

31. The composition of claims 26-30, wherein the target molecule is selected from the group consisting of of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid

32. The composition of claims 24-31, wherein the first ESB comprises a first common linker (CL).

33. The composition of claims 24-32, wherein said first target molecule is directly bound to said first UBA and said first ESB is directly bound to said first UBA.

34. The composition of claims 24-33, wherein the plurality of ordered APS is formed by stepwise addition of APSs in separate rounds.

35. The composition of 34, wherein the APS added on each round is linked to the first complex.

36. The composition of claim 34-35, wherein the linking is in order of rounds.

37. The composition of claims 24-36, wherein the linking of an APS, an ESB, or a UBA is performed using chemical methods.

38. The composition of claim 37, wherein the chemical method comprises Click chemistry.

39. The composition of claims 35-38, wherein the linking is performed in the presence of Cu^{I}.

40. The composition of claims 24-39, wherein a UBA, an APS, or an ESB comprises nucleic acids.

41. The composition of 40, further comprising a first linking oligonucleotide comprising a first and a second complementary region to two components selected from a UBA, an APS, and an ESB.

42. The composition of claims 40-41, wherein a UBA, an APS, or an ESB is linked using a linking oligonucleotide comprising the first and the second complementary region to two components selected from a UBA, an APS, and an ESB.

43. The composition of claims 40-42, further comprising a second linking oligonucleotide comprising a third and a fourth complementary region to two components selected from a UBA, an APS, and an ESB.

44. The composition of claims 40-43, wherein a UBA, an APS, or an ESB is linked using a linking oligonucleotide comprising the third and the fourth complementary region to two components selected from a UBA, an APS, and an ESB.

45. The composition of claims 43-44, wherein the second and fourth complementary regions are identical.

46. The composition of claims 43-45, wherein the second and fourth complementary regions are identical.

47. The composition of claims 41-46, wherein the first or second complementary region is shared between two APSs within the plurality of APSs.

48. The composition of claim 42-47, wherein the linking is performed by ligation.

49. The composition of claims 41-48, wherein the linking oligonucleotide comprises a subcode encoding the origin of the APS or the ESB.

50. The composition of claims 24-49, wherein the APS has a subcode encoding the origin of the APS.

51. The composition of claims 24-50, wherein the ESB has a subcode encoding the origin of the ESB.

52. The composition of claims 24-51, wherein an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag.

53. The composition of claim 52, wherein the unique counter tag is detectable.

54. The composition of claims 41-53, wherein the ESB is covalently linked to the linking oligonucleotide.

55. The composition of claims 24-54, wherein an APS or an ESB comprises an amplification primer binding region.

56. The composition of claims 24-55, wherein the APSs and ESB, when linked, are capable of encoding a secondary product.

57. The composition of claim 56, wherein the secondary product is an RNA or a peptide.

58. The composition of claims 24-57, wherein the APSs and ESB, when linked, comprises a polymerase start site.

59. The composition of claims 57-58, wherein the peptide comprises an affinity tag.

60. The composition of claim 59, wherein the affinity tag is a His-tag.

61. The composition of claims 24-60, wherein the UBA, ESB, or APS is templatable.

62. The composition of claims 24-61, further comprising a probe.

63. The composition of claims 24-62, wherein the probe is attached to a surface.

64. The composition of claims 24-63, wherein the surface comprises an array.

65. The composition of claims 24-64, wherein the surface comprises a bead.

66. The composition of claims 24-65, wherein the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid.

67. The composition of claims 24-66, wherein the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits.

68. The composition of claims 24-67, wherein said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight.

69. A kit for labeling a target molecule of a cell in a population of cells with a cell origination barcode, comprising

(a) n sets of m assayable polymer subunits (APSs) each comprising a distinct package of information; wherein the packages of information are capable of being linked in an ordered fashion;

(b) a target molecule-specific unique binding agent (UBA).

70. A kit for labeling a target molecule of a cell in a population of cells with a cell origination barcode, comprising

(a) n sets of m assayable polymer subunits (APSs) each comprising a distinct package of information; wherein the packages of information are capable of being linked in an ordered fashion;

(b) a plurality of target molecule-specific unique binding agents (UBA) each linked with a UBA-specific epitope specific barcode (ESB).

71. A kit for labeling a target molecule of a cell in a population of cells with a cell origination barcode, comprising

(a) n sets of m assayable polymer subunits (APSs) each comprising a distinct package of information; wherein the packages of information are capable of being linked in an ordered fashion;

(b) a plurality of target molecule-specific unique binding agents (UBA);

(c) a plurality of UBA-specific epitope specific barcode (ESB), wherein each ESB is capable of linking with a designated UBA.

72. The kit of claims 69-71, wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

73. The kit of claims 69-72, wherein m is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

74. The kit of claims 69-73, wherein n is greater than 10.

75. The kit of claims 69-74, wherein m is greater than 20.

76. The kit of claim 70-75, wherein the first ESB comprises a first common linker (CL).

77. The kit of claim 70-76, wherein said ESB is capable of directly binding to said UBA.

78. The kit of claims 69-77, wherein said UBA is capable of directly binding to the target molecule.

79. The kit of claims 69-78, wherein at least two of the assayable polymer subunit (APS) sets are identical.

80. The kit of claims 69-79, wherein the APSs in a first set is linkable to the APSs in a second set.

81. The kit of claim 80, the APSs in a first set is further linkable to the APSs in a second set in an ordered fashion.

82. The kit of claims 79-81, wherein an APS, an ESB, or a UBA is capable of being linked using chemical methods.

83. The kit of claim 82, wherein the chemical method comprises Click chemistry.

84. The kit of claims 82-83, wherein the presence of Cu^{I} is required for linkage.

85. The kit of claims 69-84, wherein a UBA, an APS, or an ESB comprises nucleic acids.

86. The kit of claim 85, further comprising a first linking oligonucleotide comprising a first and a second complementary region to two components selected from a UBA, an APS, and an ESB.

87. The kit of claim 86, further comprising a second linking oligonucleotide comprising a third and a fourth complementary region to two components selected from a UBA, an APS, and an ESB.

88. The kit of claim 87, wherein the first and third complementary regions are identical.

89. The kit of claims 86-87, wherein the second and fourth complementary regions are identical.

90. The kit of claims 85-89, an APS, an ESB, or a UBA is capable of being linked by ligation.

91. The kit of claims 86-90, wherein the linking oligonucleotide comprises a subcode encoding the original set of the APS or the ESB.

92. The kit of claims 69-91, wherein the APS has a subcode encoding the origin population of the APS.

93. The kit of claims 69-92, wherein the ESB has a subcode encoding the origin population of the ESB.

94. The kit of claims 69-93, wherein an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag.

95. The kit of claim 94, wherein the unique counter tag is detectable.

96. The kit of claims 86-95, wherein the ESB is covalently linked to the linking oligonucleotide.

97. The kit of claims 69-96, wherein an APS or an ESB comprises an amplification primer binding region.

98. The kit of claims 69-97, wherein the APSs and ESB, when linked, are capable of encoding a secondary product.

99. The kit of claim 98, wherein the secondary product is an RNA or a peptide.

100. The kit of claims 69-99, wherein the APSs and ESB, when linked, comprises a polymerase start site.

101. The kit of claims 99-100, wherein the peptide comprises an affinity tag.

102. The kit of claim 101, wherein the affinity tag is a His-tag.

103. The kit of claims 69-102, wherein the UBA, ESB, or APS is templatable.

104. The kit of claims 69-103, further comprising a probe.

105. The kit of claims 69-104, wherein the probe is attached to a surface.

106. The kit of claims 69-105, wherein the surface comprises an array.

107. The kit of claims 69-106, wherein the surface comprises a bead.

108. The kit of claims 69-107, wherein the plurality of UBAs comprises 2, 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600, 700, 800, 900, 1000 or more than 1000 UBAs.

109. The kit of claims 69-108, wherein the plurality of UBAs comprises up to 2000 UBAs.

110. The kit of claims 69-109, wherein the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid.

111. The kit of claims 69-110, wherein the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits.

112. The kit of claims 69-111, wherein said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight.

113. A method for identifying target molecules sharing a common particle origin, comprising

labeling a first plurality of targets of a first particle in a population of x particles with a first origination barcode; and

labeling a second plurality of targets of a second particle in a population of x particles with a second origination barcode;

wherein each origination barcode comprises a set of n assayable polymer subunits (APS);

wherein each of the n APSs in the first and second set of APSs is selected from a group comprising m different APSs; and

wherein the first and second origination barcodes are detectably different from each other with a certainty of $c = 1- [(1 – 1/x)^(m^n)]$.

114. The method of claim 113, wherein x is greater than 1,000,000.

115. The method of claims 113-114, wherein c is greater than 99.9%.

116. The method of claims 113-115, wherein c is greater than 99.99%.

117. The method of claims 113-116, wherein c is greater than 99.999%.

118. The method of claims 113-117, wherein c is greater than 99.9999%.

119. The method of claims 113-118, wherein c is greater than 99.99999%.

120. The method of claims 113-119, wherein n is 2, 3, 4, 5, 6, 7, 8, 9, or 10.

121. The method of claims 113-120, wherein m is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

122. The method of claims 113-121, wherein n is greater than 10.

123. The method of claims 113-122, wherein m is greater than 20.

124. The method of claims 113-123, wherein at least one discrete particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead.

125. The method of claims 113-124, wherein the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid.

126. The method of claims 113-125, wherein at least two groups comprising m different APSs are identical.

127. The method of claims 113-126, wherein the n APSs are added in separate rounds.

128. The method of claims 113-127, wherein the APSs of separate rounds are linked.

129. The method of claim 128, wherein the linking is in order of rounds.

130. A method of imparting a particle specific code to a component of a particle of a population of particles, the method comprising:

linking a first ordered set of assayable polymer subunits (APS) to a first component of a first particle of a population of particles, wherein the order of the APSs is detectable by a detection method.

131. The method of claim 130, further comprising detecting the first ordered set of APSs linked with the first component, thereby determining a particle origin of the first component.

132. The method of claims 130-131, further comprising linking a second ordered set of assayable polymer subunits (APS) to a second component of the first particle of a population of particles, wherein the order of the APSs is detectable.

133. The method of claim 132, further comprising detecting the second ordered set of APSs linked with the second component, thereby determining the particle origin of the second component.

134. The method of claims 130-133, wherein the first and the second ordered sets of APSs linked with the first and second components of the first particle are the same.

135. The method of claims 130-134, further comprising linking a third ordered set of assayable polymer subunits (APS) to a first component of second particle of a population of particles, wherein the order of the APSs is detectable.

136. The method of claim 135, wherein the first ordered set of APSs linked with the first component of the first particle is different than the third ordered set of assayable polymer subunits linked with the first component of the second particle.

137. The method of claims 130-136, further comprising linking a component specific epitope specific barcode (ESB) to the first component.

138. The method of claims 130-137, further comprising linking a component specific ESB to the second component.

139. The method of claims 130-138, wherein at least the particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead.

140. The method of claims 130-139, wherein said at least one target molecule is directly bound to said first UBA and said ESB is directly bound to said UBA.

141. The method of claims 130-140, wherein at least two of the assayable polymer subunit (APS) sets are identical.

142. The method of claims 130-141, wherein each APS from the ordered set of APSs is linked to the first complex.

143. The method of claim 142, wherein the linking is in order of rounds.

144. The method of claims 130-143, wherein the UBA, ESB, or APS encodes a secondary product.

145. The method of claim 144, wherein the secondary product is an RNA or a peptide.

146. The method claims 130-145, wherein the UBA, ESB, or APS is templatable.

147. The method of claims 137-146, wherein the ESB further comprises a unique counter tag.

148. The method of claim 147, wherein the quantity of the target molecule of the molecule is estimated using the counter tag.

149. The method of any preceding claims 143-148, wherein the APS further comprises a round-specific subcode.

150. The method of claim 149, wherein the detection further comprises determining the presence of an APS from a designated round.

151. The method of claims 130-150, wherein detection is digital.

152. The method of claims 130-151, wherein detection is indirect.

153. The method of claims 130-152, wherein detecting comprises mass spectrometry.

154. The method of claims 130-153, wherein detecting comprises nucleic acid sequencing.

155. The method of claims 130-154, wherein detecting comprises peptide sequencing.

156. The method of claims 130-155, wherein detecting comprises mass gel electrophoresis.

157. The method of claims 130-156, wherein detecting comprises HPLC or other chromatographic separation.

158. The method of claims 130-157, wherein detecting comprises detecting one or more signals associated with one or more individual APSs.

159. The method of claim 158, wherein the signals are ordered.

160. The method of claims 130-159, wherein detecting comprises using one or more probes.

161. The method of claim 160, wherein the probe is attached to a surface.

162. The method of claim 161, wherein the surface comprises an array.

163. The method of claim 161, wherein the surface comprises a bead.

164. The method of claims 130-163, wherein detecting comprises a separation.

165. The method of claim 164, wherein the separation is multi-dimensional.

166. The method of claims 164-165, wherein the separation resolves the first linkable UBA-dependent epitope specific barcode (ESB) from a second linkable UBA-dependent epitope specific barcode (ESB).

167. The method of claims 130-166, wherein 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600, 700, 800, 900, 1000 or more than 1000 different target molecules are detected.

168. The method of claims 130-167, wherein up to 2000 different target molecules are detected.

169. The method of claims 130-168, wherein the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid.

170. The method of claims 130-169, wherein the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits.

171. The method of claims 130-170, wherein said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight.

172. The method of claims 130-171, wherein the APSs are linked to through ligation or extension via polymerization.

173. The method of claims 130-172, wherein a cell origination barcode (COB) is generated with the APSs from the ordered set of APSs.

174. The method of claim 173, wherein each COB in said plurality of complexes has a detectable signal or sequence that distinguishes it from other COBs in said population of cells.

175. The method of claims 130-174, wherein an APS, an ESB, or a UBA is linked using chemical methods.

176. The method of claim 175, wherein the chemical method comprises Click chemistry.

177. The method of claims 175-176, wherein the linking is performed in the presence of Cu^{I}.

178. The method of claims 130-177, wherein a UBA, an APS, or an ESB comprises nucleic acids.

179. The method of claim 178, wherein the linking of a UBA, an APS, or an ESB is performed using a linking oligonucleotide that comprises a first and a second complementary region to two components to be linked.

180. The method of claim 179, wherein the first or second complementary region is shared between APSs within a population of APSs.

181. The method of claim 179-180, wherein the first or second complementary region is distinct for two different round-specific sets of APSs.

182. The method of claim 179-181, further comprising ligation.

183. The method of claims 130-182, wherein the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid..

184. The method of claims 130-183, wherein the first ESB comprises a first common linker (CL).

185. The method of claims 130-184, wherein the first ESB comprises a first common linker (CL).

186. The method of claims 130-185, wherein an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag.

187. The method of claim 186, wherein detection comprises detecting the unique counter tag.

188. The method of claim 187, wherein the number of unique counter tags associated with a specific ESB is determined.

189. The method of claim 188, wherein the number of detected unique counter tags relate to the initial quantity of the specific ESB.

190. The method of claims 179-189, wherein the ESB is covalently linked to the linking oligonucleotide.

191. The method of claims 130-190, wherein an APS or an ESB comprises an amplification primer binding region.

192. The method of claims 173-191, wherein a COB encodes a peptide sequence.

193. The method of claims 173-192, wherein the COB comprises a polymerase start site.

194. The method of claims 192-193, wherein the peptide comprises an affinity tag.

195. The method of claim 194, wherein the affinity tag is a His-tag.

196. A method for detecting a plurality of properties originating from a plurality of discrete particles, the method comprising:

a) providing:

i) a population of particles comprising at least a first target molecule;

ii) a first unique binding agent (UBA) specific for the first target molecule;

iii) a first linkable UBA-dependent epitope specific barcode (ESB);

iv) a plurality of round-specific assayable polymer subunit (APS) sets, each set containing a plurality of APSs that are detectably distinct from each other;

b) forming at least a first complex comprising said at least first target molecule, said first UBA probe, and said first ESB;

c) performing n rounds of split pool synthesis, each round comprising;

i) splitting the population of particles into m reaction volumes;

ii) contacting one or more reaction volumes with an APS from the APS set specific for the round;

iii) pooling two or more reaction volumes;

d) detecting a plurality of properties from at least one particle from the population of particles;

wherein at least one of the properties relate to a quantity or an identity for a target molecule associated with the particle.

197. A method for detecting a plurality of properties originating from a plurality of discrete particles, the method comprising:

a) providing:

i) a population of particles comprising at least a first target molecule;

ii) a first unique binding agent (UBA) specific for the first target molecule;

iii) a first linkable UBA-dependent epitope specific barcode (ESB); and

iv) a plurality of round-specific assayable polymer subunit (APS) sets, each set containing a plurality of APSs that are detectably distinct from each other;

b) forming at least a first complex comprising said at least first target molecule, said first UBA probe, and said first ESB;

c) performing n rounds of split pool synthesis, each round comprising:

i) splitting the population of particles into m reaction volumes;

ii) contacting one or more reaction volumes with an APS from the APS set specific for the round; and

iii) pooling two or more reaction volumes;

d) performing another round of split pool synthesis comprising steps c) i) and c) ii);

e) detecting a plurality of properties from at least one particle from the population of particles;

wherein at least one of the properties relate to a quantity or an identity for a target molecule associated with the particle.

198. The method of claims 196-197, wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

199. The method of claims 196-198, wherein n is more than 20.

200. The method of claims 196-199, wherein m is different between at least two rounds.

201. The method of claims 196-200, wherein m is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

202. The method of claims 196-201, wherein m is more than 20.

203. The method of claims 196-202, wherein at least one discrete particle is selected from the group consisting of a cell, a liposome, an organelle, a micelle, a droplet and a bead.

204. The method of claims 196-203, wherein the target molecule is selected from the group consisting of a peptide, a polypeptide, an oligopeptide, a protein, a phosphoprotein, an antibody, a nucleic acid, a peptide nucleic acid, a synthetic small molecule, a disaccharide, a trisaccharide, an oligosaccharide, a polysaccharide, a lipid, a steroid, and a phospholipid.

205. The method of claims 196-204, wherein the first ESB comprises a first common linker (CL).

206. The method of claims 196-205, wherein said at least one target molecule is directly bound to said first UBA and said ESB is directly bound to said UBA.

207. The method of claims 196-206, wherein at least two of the assayable polymer subunit (APS) sets are identical.

208. The method of claims 196-207, wherein the APS added on each round is linked to the first complex.

209. The method of claim 208, wherein the linking is in order of rounds.

210. The method of claims 196-209, wherein the UBA, ESB, or APS encodes a secondary product.

211. The method of claim 210, wherein the secondary product is an RNA or a peptide.

212. The method of claims 196-211, wherein the UBA, ESB, or APS is templatable.

213. The method of claims 196-212, wherein the ESB further comprises a unique counter tag.

214. The method of claim 213, wherein the quantity of the target molecule of the molecule is estimated using the counter tag.

215. The method of claims 196-214, wherein the APS further comprises a round-specific subcode.

216. The method of claim 215, wherein the detection further comprises determining the presence of an APS from a designated round.

217. The method of claims 196-216, wherein detection is digital.

218. The method of claims 196-217, wherein detection is indirect.

219. The method of claims 196-218, wherein detecting comprises mass spectrometry.

220. The method of claims 196-219, wherein detecting comprises nucleic acid sequencing.

221. The method of claims 196-220, wherein detecting comprises peptide sequencing.

222. The method of claims 196-221, wherein detecting comprises detecting one or more signals associated with one or more individual APSs.

223. The method of claim 222, wherein the signals are ordered.

224. The method of claims 196-223, wherein detecting comprises using one or more probes.

225. The method of claim 224, wherein the probe is attached to a surface.

226. The method of claim 225, wherein the surface comprises an array.

227. The method of claim 225, wherein the surface comprises a bead.

228. The method of claims 196-227, wherein detecting comprises a separation.

229. The method of claim 228, wherein the separation is multi-dimensional.

230. The method of claims 228-229, wherein the separation resolves the first linkable UBA-dependent epitope specific barcode (ESB) from a second linkable UBA-dependent epitope specific barcode (ESB).

231. The method of claims 196-230, wherein 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600, 700, 800, 900, 1000 or more than 1000 different target molecules are detected.

232. The method of claims 196-231, wherein up to 2000 different target molecules are detected.

233. The method of claims 196-232, wherein the UBA is selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid.

234. The method of claims 196-233, wherein the ESB is selected from the group consisting of nucleic acids, beads and chemical subunits.

235. The method of claims 196-234, wherein said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight.

236. The method of claims 196-235, wherein the APSs are linked to through ligation or extension via polymerization.

237. The method of claims 196-236, wherein a cell origination barcode (COB) is generated from APSs of round specific APS sets.

238. The method of claim 237, wherein each COB in said plurality of complexes has a detectable signal or sequence that distinguishes it from other COBs in said population of cells.

239. The method of claims 196-238, wherein an APS, an ESB, or a UBA is linked using chemical methods.

240. The method of claim 239, wherein the chemical method comprises Click chemistry.

241. The method of claims 239-240, wherein the linking is performed in the presence of Cu^{I}.

242. The method of claims 196-241, wherein a UBA, an APS, or an ESB comprises nucleic acids.

243. The method of claim 242, wherein the linking of a UBA, an APS, or an ESB is performed using a linking oligonucleotide that comprises a first and a second complementary region to two components to be linked.

244. The method of claim 243, wherein the first or second complementary region is shared between APSs within a population of APSs.

245. The method of claim 243-244, wherein the first or second complementary region is distinct for two different round-specific sets of APSs.

246. The method of claim 243-245, further comprising ligation.

247. The method of claims 243-246, wherein the linking oligonucleotide comprises a subcode encoding the origin population of the APS or the ESB.

248. The method of claims 196-247, wherein the APS has a subcode encoding the round-specific set of the APS.

249. The method of claims 196-248, wherein the ESB has a subcode encoding the origin presence of the ESB.

250. The method of claims 196-249, wherein an individual APS, ESB, or linking oligonucleotide molecule comprises a unique counter tag.

251. The method of claim 250, wherein detection comprises detecting the unique counter tag.

252. The method of claim 251, wherein the number of unique counter tags associated with a specific ESB is determined.

253. The method of claim 252, wherein the number of detected unique counter tags relate to the initial quantity of the specific ESB.

254. The method of claims 243-253, wherein the ESB is covalently linked to the linking oligonucleotide.

255. The method of claims 196-254, wherein an APS or an ESB comprises an amplification primer binding region.

256. The method of claims 237-255, wherein a COB encodes a peptide sequence.

257. The method of claims 237-256, wherein the COB comprises a polymerase start site.

258. The method of claims 256-257, wherein the peptide comprises an affinity tag.

259. The method of claim 258, wherein the affinity tag is a His-tag.

260. The method of claims 196-259, wherein each of the reaction volumes created by the most recent splitting receives a different APS from the APS set.

261. A method for detecting at least one target molecule in a sample comprising the steps:

(a) providing:

(i) a population of cells potentially comprising at least one target molecule,

(ii) a first unique binding agent (UBA) specific for a first target molecule,

(iii) a first epitope specific barcode (ESB) specific for a region of said first UBA, wherein said ESB comprises a first common linker moiety, and

(iv) a population of assayable polymer subunits (APSs), wherein said APSs comprises a second common linker moiety and a third common linker moiety, wherein said second linker moiety is complementary to said first common linker moiety is said first ESB;

(b) forming at least a first complex comprising said at least one target molecule, said first UBA probe, and said first ESB, wherein said at least one target molecule is bound to said first UBA and said ESB is bound to said UBA;

(c) splitting said population into two or more samples;

(d) adding one APS from said population of APSs per sample to said two or more samples from step (c), wherein a second complex is formed with said least one target molecule, said first UBA probe, said first ESB, and a first APS, and wherein said second common linker moiety from said first APS is bound to said first linker moiety from said first ESB;

(e) pooling said two or more samples from step (c) into one sample;

(f) splitting said sample from step (e) into two or more samples;

(g) adding one APS from said population of APSs per sample to said two or more samples from step (e), wherein a third complex is formed with said least one target molecule, said first UBA probe, said first ESB, said first APS, and said second APS, wherein said second common linker

moiety from said second APS is bound to said third linker moiety from said first APS, and wherein said first APS and said second APS form a cell origination barcode (COB); and
(c) detecting said third complex or at least part of said third complex.

262. The method of claim 261 further comprising repeating steps (e), through (g).

263. The method of claim 261 further comprising detecting of a plurality of target molecules by a method comprising forming a plurality of complexes in step (b), each complex comprising (i) at least one target molecule (ii) a first UBA and (iii) a first epitope specific barcode (ESB) specific for a region of said first UBA, wherein said ESB comprises a first common linker moiety, wherein said at least one target molecule is bound to said first UBA and said ESB is bound to said UBA.

264. The method of claim 263, wherein each COB in said plurality of complexes has a detectable signal or sequence that distinguishes it from other COB in said population of cells.

265. The method of claims 261-264, wherein said complex is detected by sequencing or mass spectrometry.

266. The method of claim 261, wherein said third complex is detected by a method comprising individually counting the presence of one or more molecules of said third complex wherein the presence of said one or more molecules of said third complex is indicative of the concentration or presence of said target molecule in a cell.

267. The method of claim 266, wherein said individually detecting further comprises detecting a digital signal.

268. The method of claim 263, wherein 3, 4, 5, 10, 20, 30, 50, 100, 200, 300, 500, 600, 700, 800, 900, 1000 or more than 1000 different target molecules are detected.

269. The method of claim 263, wherein up to 2000 different target molecules are detected.

270. The method of claim 261, wherein said UBA comprises a structure selected from the group consisting of antibody, peptide, aptamer, peptoid and nucleic acid or any combination.

271. The method of claim 261, wherein said ESB comprises a structure selected from the group consisting of nucleic acids, beads and chemical subunits or any combination.

272. The method of claim 261, wherein said APS comprises a nucleic acid, a small molecule, or buildable complex molecules of deterministic weight.

273. The method of claim 261, wherein said APS comprises a single-stranded nucleic acid hybridized to a complementary polynucleotide sequence having attached thereto a detectable label.

274. The method of claim 261, wherein said first APS is attached to said first ESB through ligation or extension via polymerization.

275. The method of claim 261, wherein said second APS is attached to said first APS through ligation or extension via polymerization.

276. The method of claim 261, wherein said common linker moiety is a nucleic acid.

277. The method of claims 261-276, wherein said ESB is attached to said USB.

278. A method for detecting at least one target molecule in a sample comprising the steps:
(a) providing:
(i) a population of cells potentially comprising at least one target molecule
(ii) a first UBA specific for a first target molecule,
(iii) a first epitope specific barcode ESB specific for a region of said first UBA, wherein said ESB comprises a first common linker moiety, and
(iv) a population of COB, wherein said population of COB comprises a second common linker moiety, wherein said second linker moiety is complementary to said first common linker moiety is said first ESB;
(b) forming at least a first complex comprising said at least one target molecule, said first UBA probe, and said first ESB, wherein said at least one target molecule is bound to said first UBA and said ESB is bound to said UBA;
(c) adding said population of COBs, wherein a second complex is formed with said least one target molecule, said first UBA probe, said first ESB, and a first COB, and wherein said second common linker moiety from said first COB is bound to said first linker moiety from said first ESB, and wherein the COBs from said population of COBs is associated with a cell from said population of cells; and
(d) detecting said second complex or at least part of said third complex.

279. The method of claim 278, wherein said first COB comprises a plurality of APS.

280. The method of claim 278, further comprising detecting of a plurality of target molecules by a method comprising:
forming a plurality of complexes in step (b), each complex comprising (i) at least one target molecule (ii) a first UBA and (iii) a first epitope specific barcode (ESB) specific for a region of said first UBA, wherein said ESB comprises a first common linker moiety, wherein said at least one target molecule is associated with said first UBA and said ESB is associated with said UBA.

281. The method of claim 263, wherein each COB in said plurality of complexes has a detectable signal or sequence that distinguishes it from other COBs in said population of cells.

282. The method of claims 261-281, wherein the linking of an APS, an ESB, or a UBA is performed using chemical methods.

283. The method of claim 282, wherein the chemical method comprises Click chemistry.

284. The method of claims 282-283, wherein the linking is performed in the presence of Cu^{I}.

285. The method of claims 264-284, wherein linking of an ABSm an ESB, or a UBA is performed using binding affinity.

286. The method of claims 261-285, wherein a UBA, an APS, or an ESB comprises nucleic acids.

287. The method of claim 286, wherein the linking of a UBA, an APS, or an ESB is performed using a linking oligonucleotide that comprises a first and a second complementary region to two components to be linked.

288. The method of claim 287, wherein the first or second complementary region is shared between APSs within a population of APSs.

289. The method of claim 287-288, wherein the first or second complementary region is distinct for different populations of APSs.

290. The method of claims 287-289, further comprising ligation.

291. The method of claims 287-290, wherein the linking oligonucleotide comprises a subcode encoding the origin population of the APS or the ESB.

292. The method of claims 261-290, wherein the APS has a subcode encoding the origin population of the APS.

293. The method of claims 261-292, wherein the ESB has a subcode encoding the origin population of the ESB.

294. The method of claims 261-293, wherein an individual APS, ESB, or linking oligonucleotide molecule comprises a unique tag.

295. The method of claim 294, wherein detection comprises detecting the unique tag.

296. The method of claim 295, wherein the number of unique tags associated with a specific ESB is determined.

297. The method of claim 296, wherein the number of detected unique tags relate to the initial quantity of the specific ESB.

298. The method of claims 287-297, wherein the ESB is covalently linked to the linking oligonucleotide.

299. The method of claims 261-298, wherein an APS or an ESB comprises an amplification primer binding region.

300. The method of claims 261-299, wherein a COB encodes a peptide sequence.

301. The method of claims 261-300, wherein the COB comprises a polymerase start site.

302. The method of claims 300-301, wherein the peptide comprises an affinity tag.

303. The method of claim 302, wherein the affinity tag is a His-tag.

304. The method of claim 261, wherein the two or more samples comprise at least 5 samples.

305. The method of claim 261, wherein the two or more samples comprise at least 10 samples.

306. The method of claim 261, wherein the two or more samples comprise at least 20 samples.

307. The method of claim 261, wherein each of the samples created by the most recent splitting receives a different APS.

308. A method for labeling an ESB linked target molecule of a cell in a population of cells with a cell origination barcode (COB), comprising:

separating each cell into an individual reaction volume; and

adding the COB to the ESB via chemical or affinity means.

309. The method of claim 308, wherein the reaction volume is selected from the group consisting of a microbubble, a microdroplet, a well, a microwell, and an enclosure in a microfluidics device.

310. A method comprising disassociating a variety of types of components originating from a cell and placing the components on a particle, wherein the components are labeled on said particle.

311. The method of claim 310, wherein the labeling comprises labeling according to cell origin.

312. The method of claims 310-311, wherein the labeling comprises labeling according to component type.

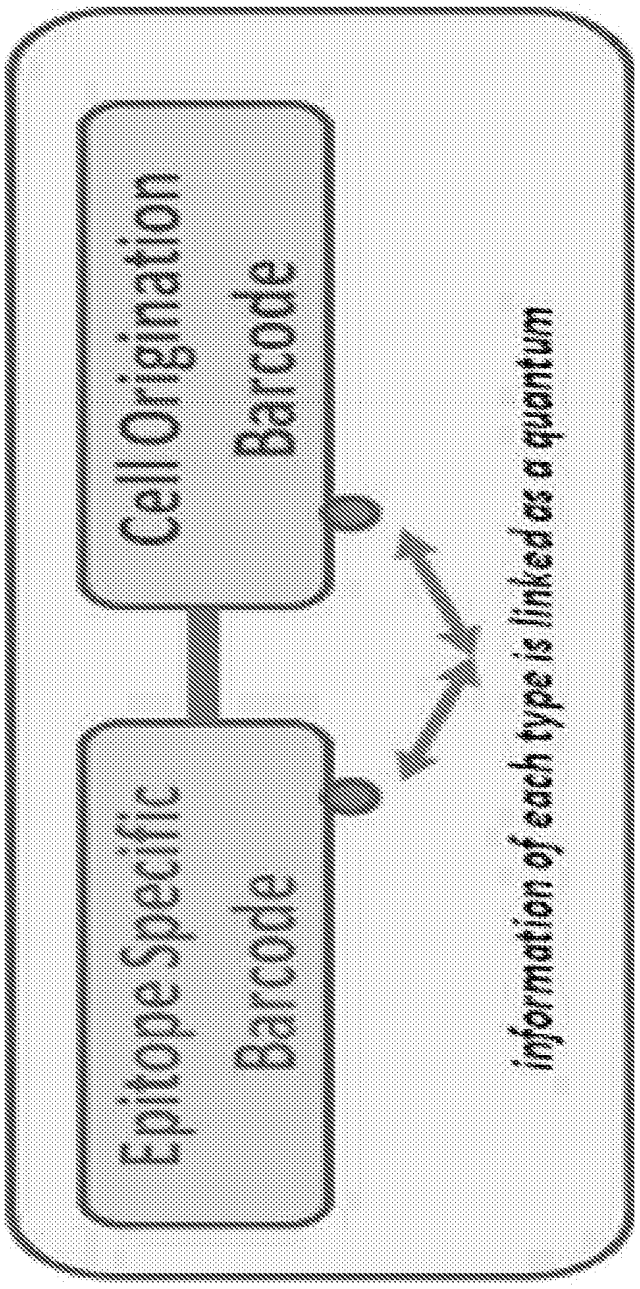

Cell Origination Barcode
Epitope Specific barcode
information of each type is linked as a quantum

Figure 1

ESB: Epitope Specific Barcodes
UBA: Unique Binding Agents
APS: Assayable Polymer Subunit (used to create cell origination barcode)
Common Linker (CL)
A
B
C
n
Epitope Specific Barcodes (ESB)
Above components are assembled in cells to make these
ESB
epitope specific barcode
Single APS
Assayable Polymer Subunit
COB
Cell Origination barcode
UBA
Unique Binding Agent

Figure 2

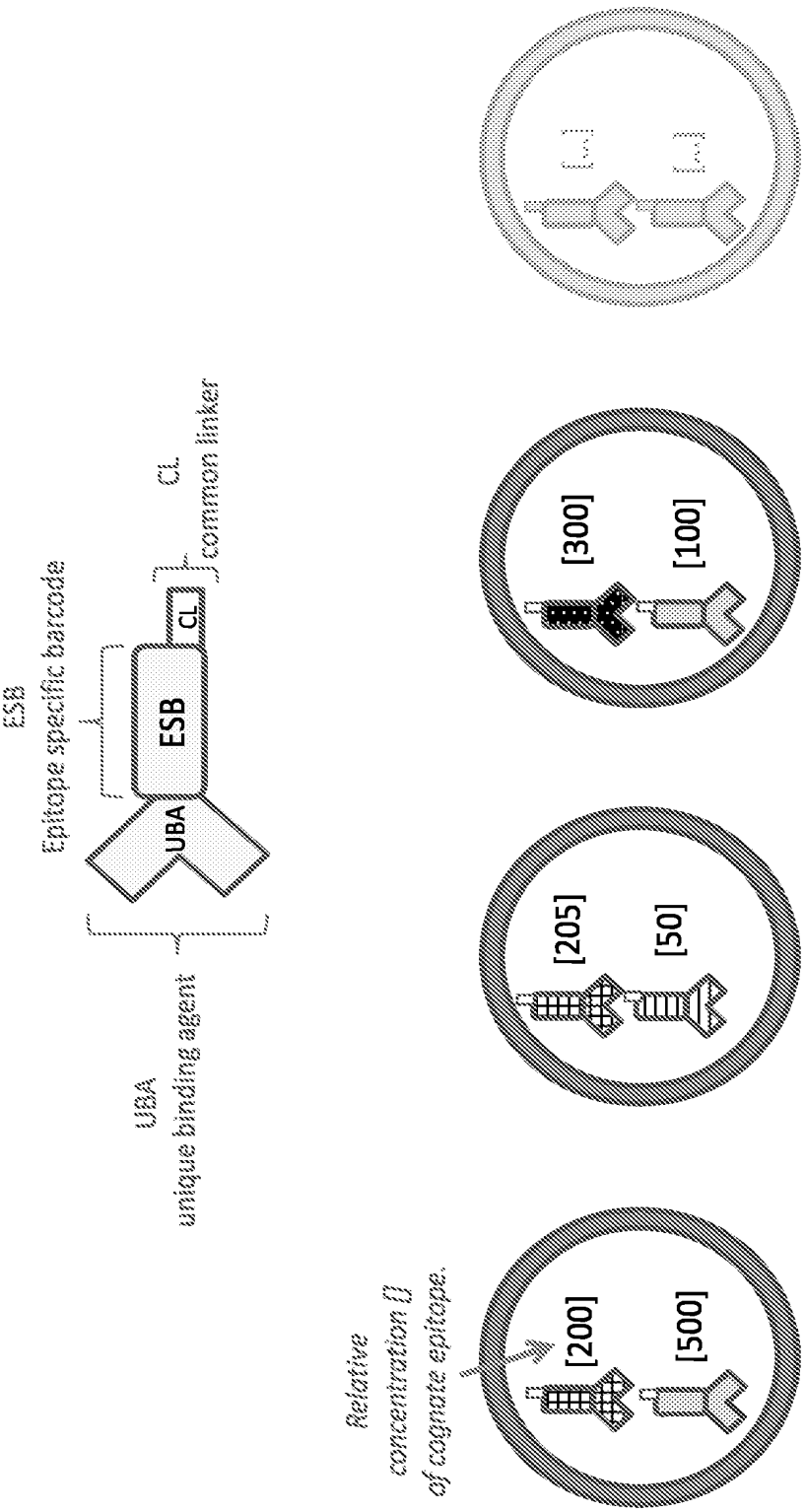

ESB
Epitope specific barcode
UBA
unique binding agent
CL
common linker
UBA
ESB
CL
Relative concentration [] of cognate epitope.
[200]
[500]
Cell α
[205]
[50]
Cell β
[300]
[100]
Cell γ
...
Cell...

Figure 3

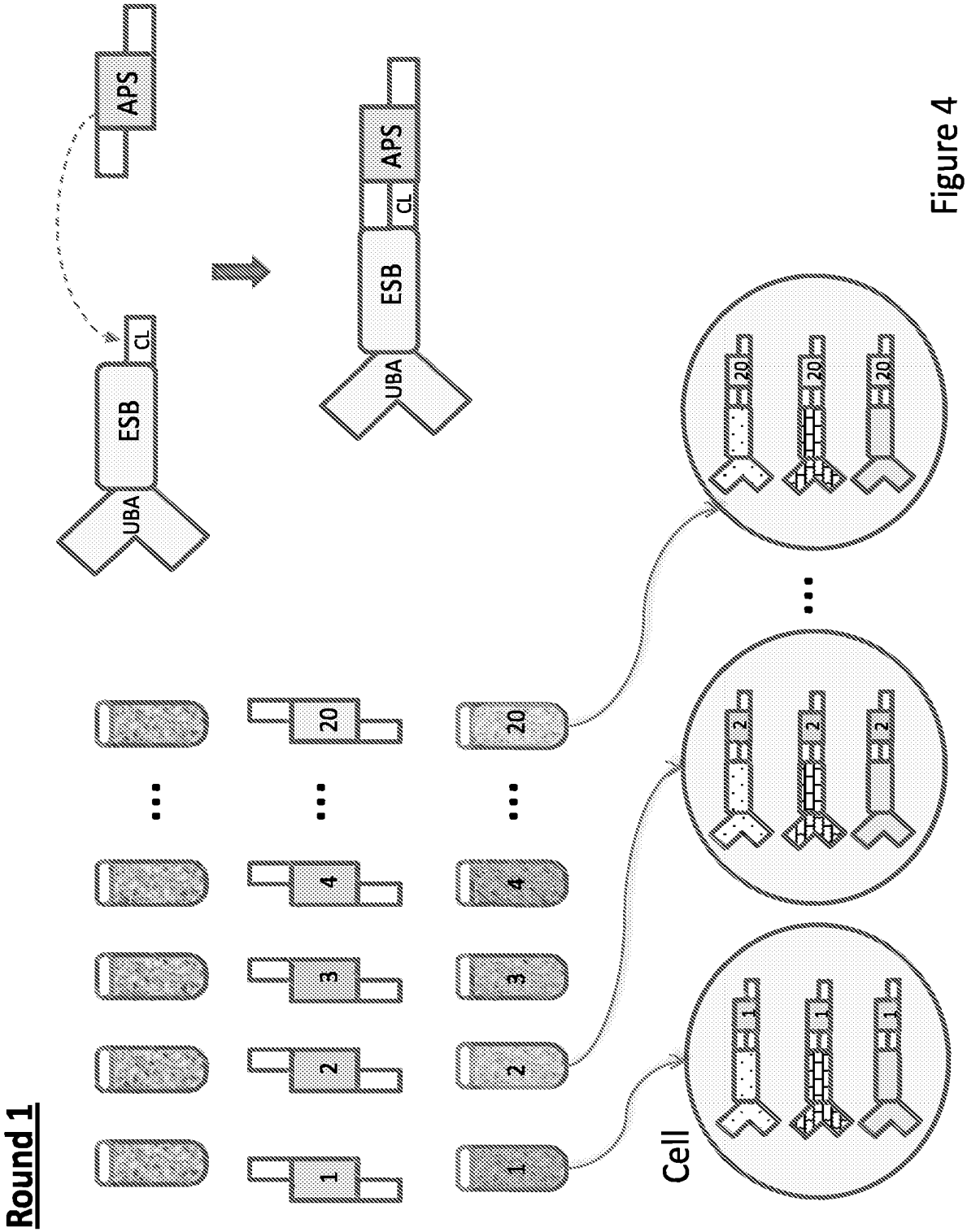
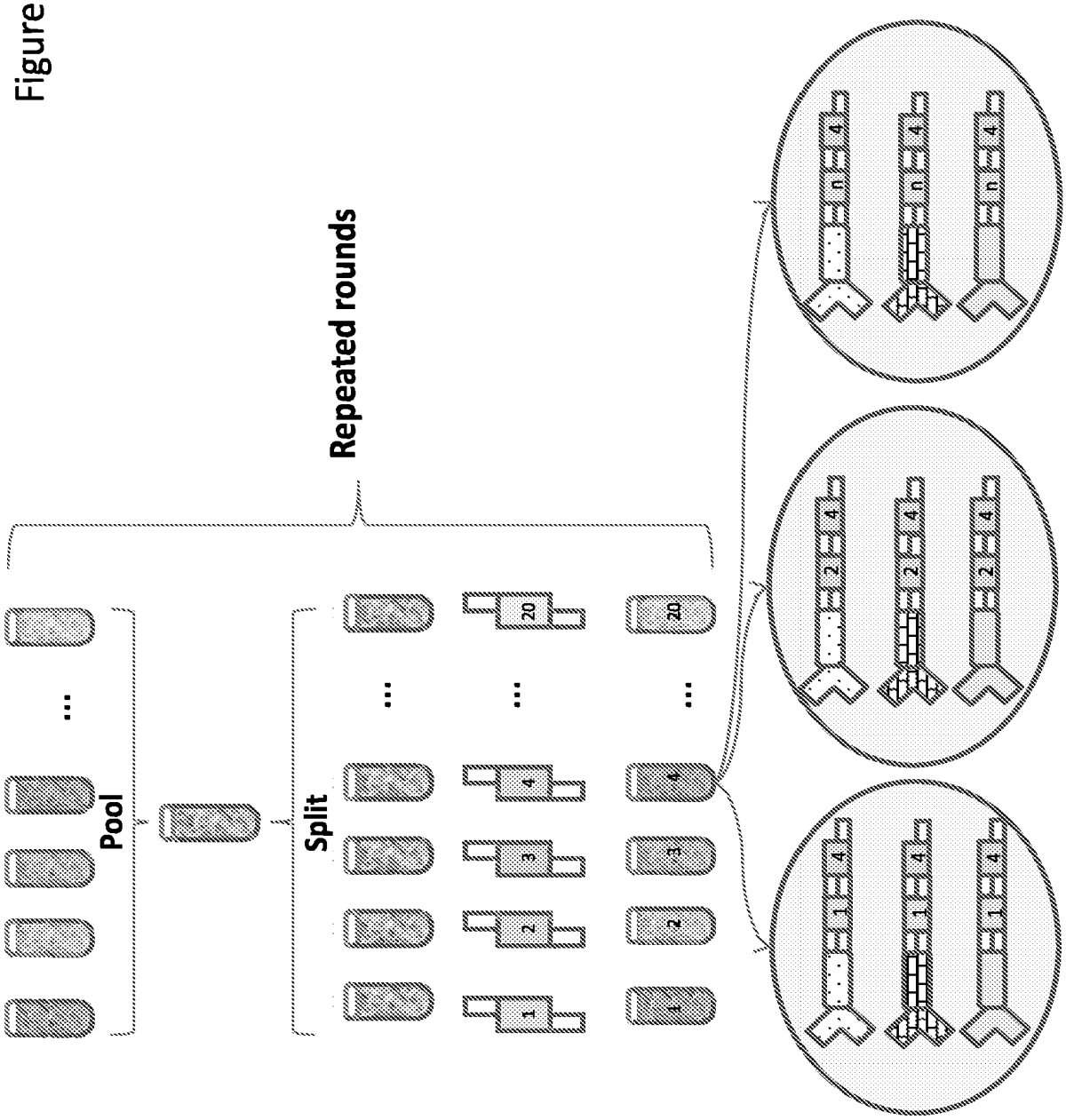


Figure 4

Figure 4 (cont.)



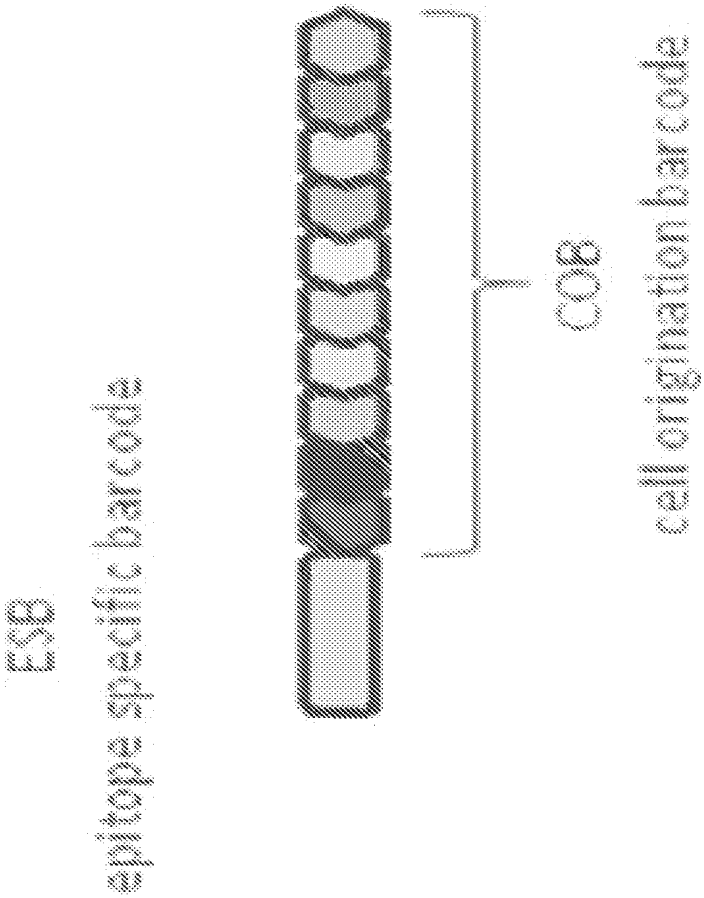
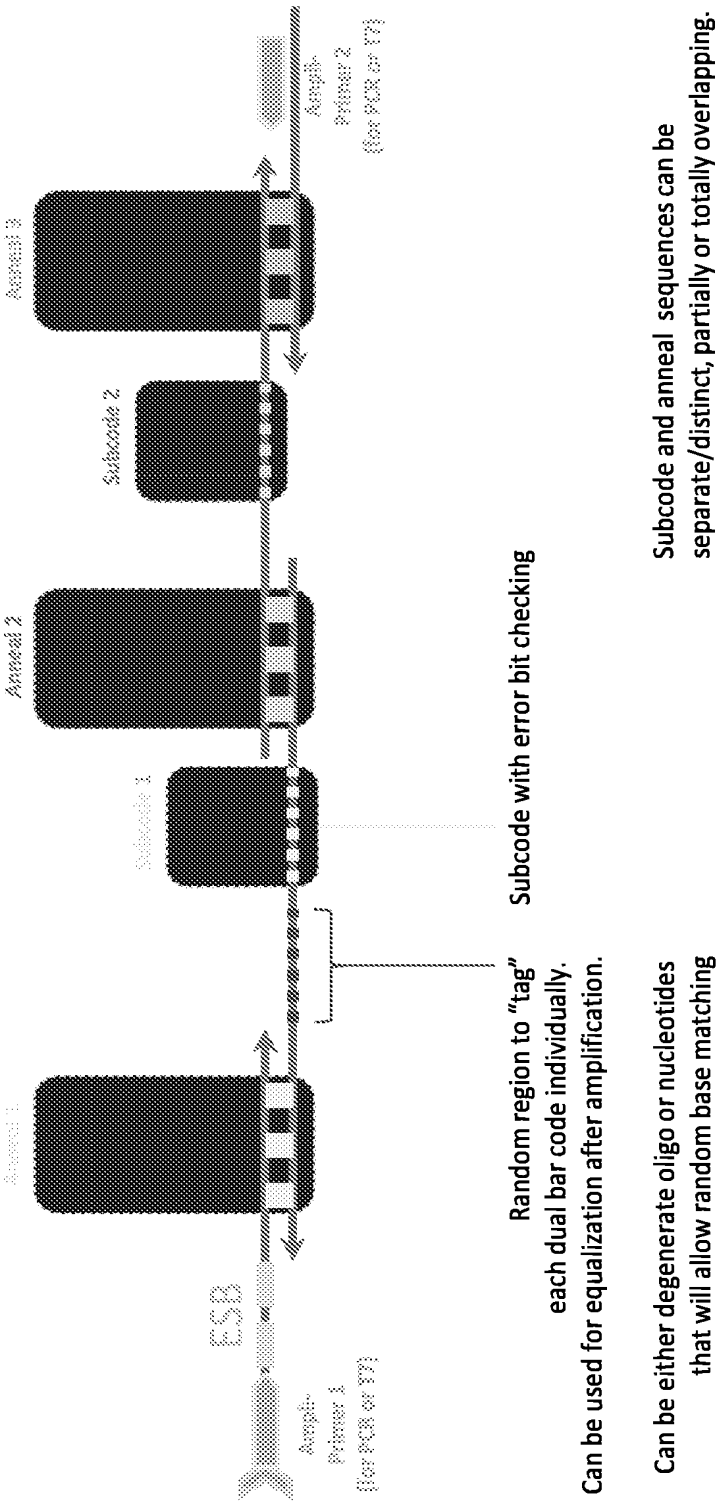
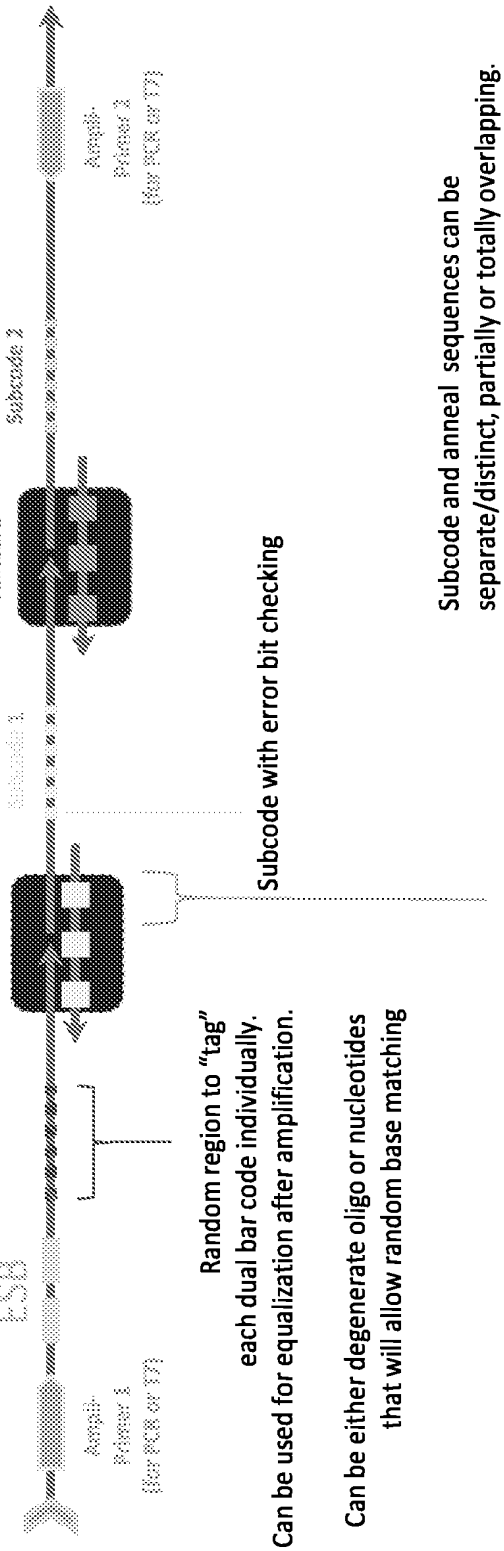

ESB
epitope specific barcode
COB
cell origination barcode

Figure 5



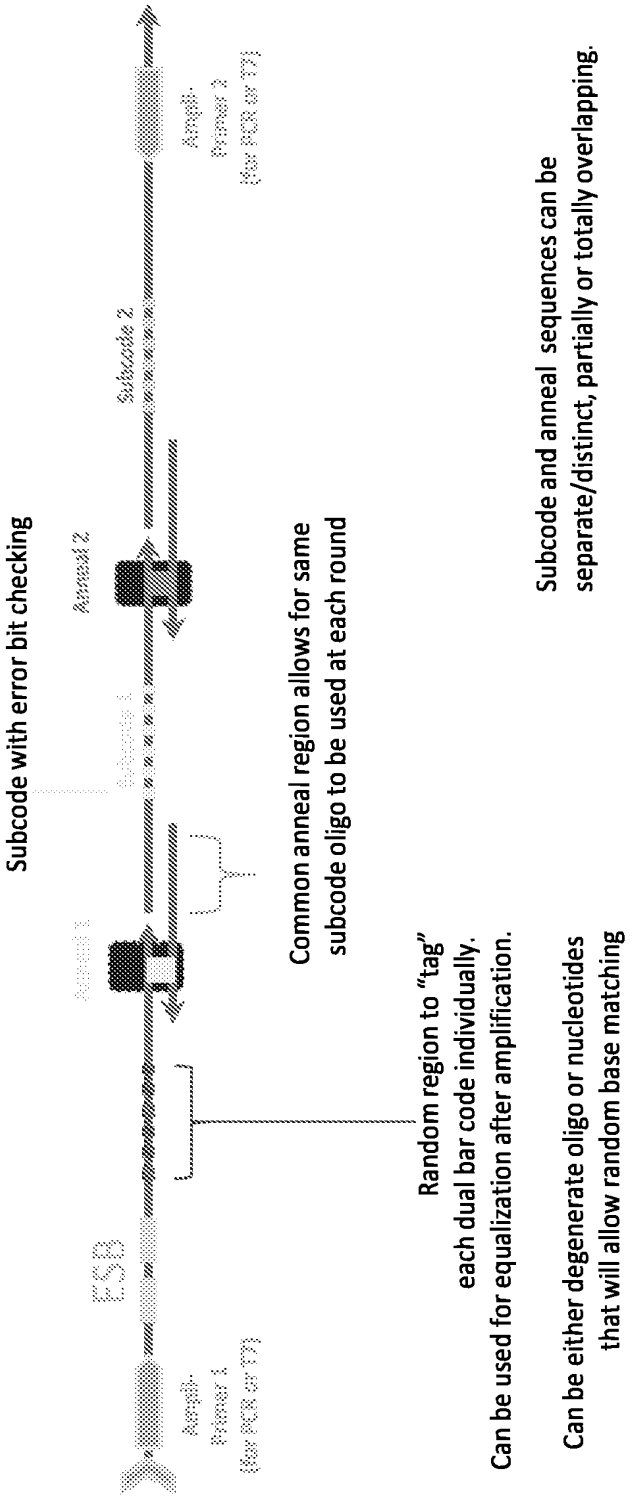
Patchwork COB

Figure 6



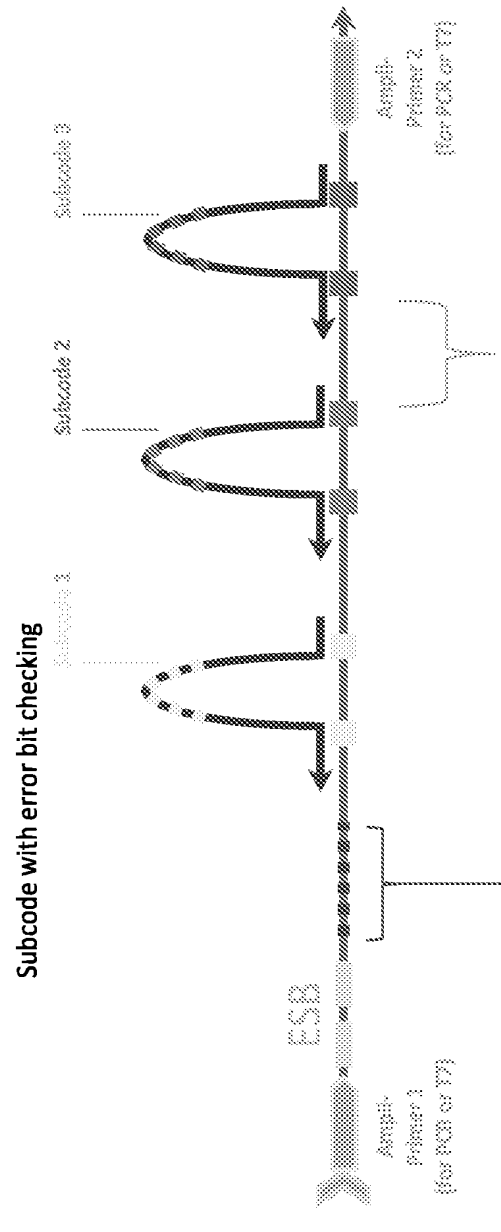
Stitch COB_SPA (specific primer anneal)

Figure 7



Stitch COB_CPA (common primer anneal)

Figure 8



Loop COB

Figure 9

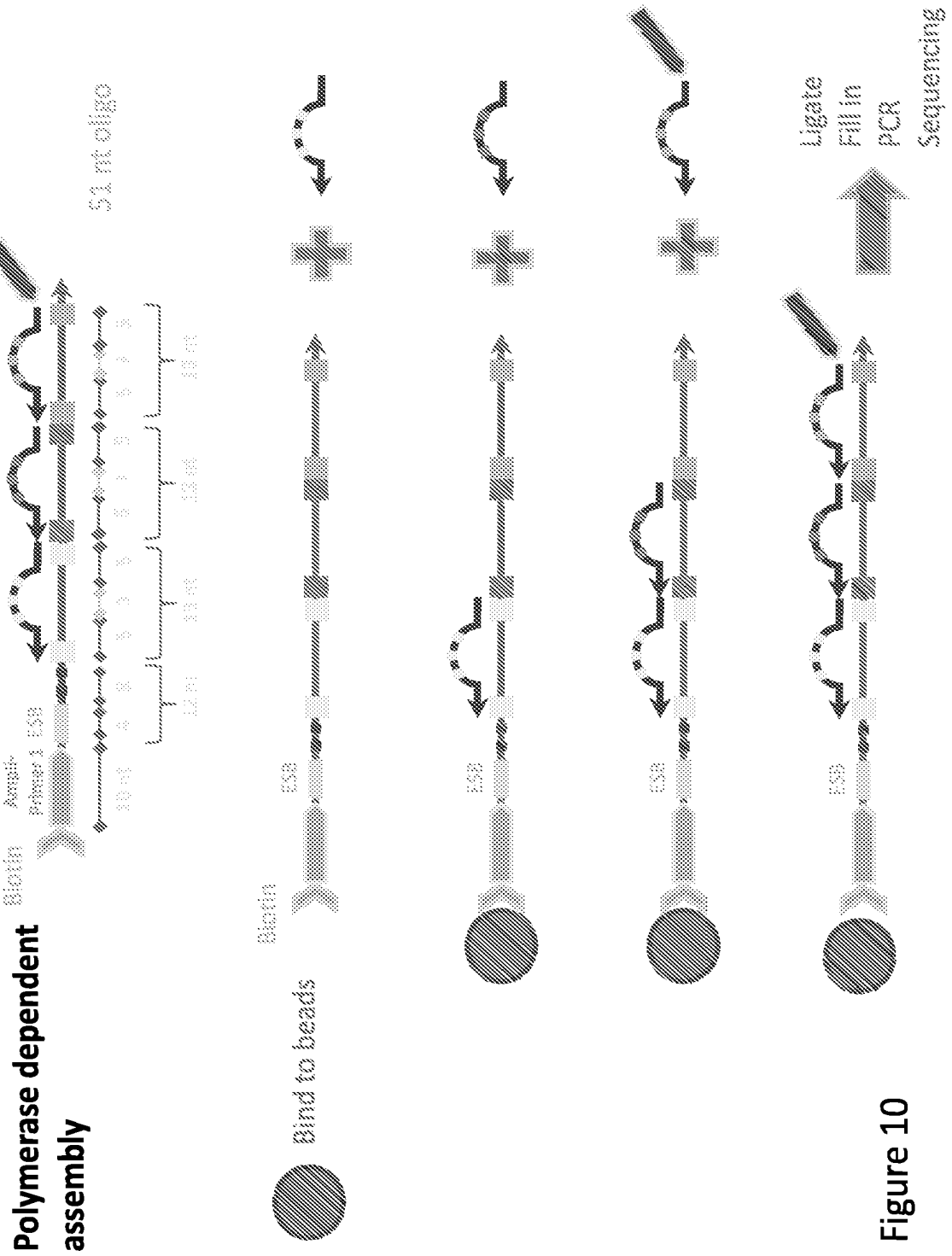
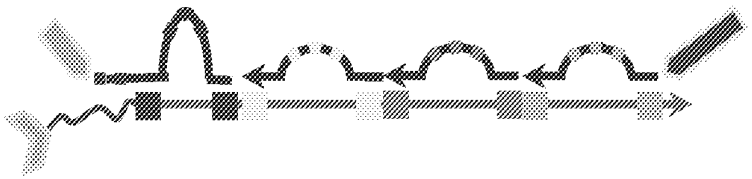

Polymerase dependent assembly
Biotin
Acrylic
Primer 1
ESB
51 nt oligo
Bind to beads
Ligate
Fill in
PCR
Sequencing

Figure 10

Figure 10 (cont)

606

Peptide readout



Can insert cleavage site for post purification MS

Can insert MS Fragile Site(s) to enable sequencing

SP6 Start	ATG-XXX	ESB- 3AA	COB- 6 * (2+1+2) = 30AA	3' His(6) tag
~ 30 bp	~ 15bp	9 bp	18 bp (6-3-6) * 6 = 90 bp	18 bp

5' SP6 Start: Acts as primer and SP6 polymerase start site

ESB-3AA: Epitope Specific Barcode of 3 amino acides covers 8000 potential independent reads

COB-6AA: 64 million distinct tags covered with 6 AA. Each is flanked with 2AA. Total length is 30AA

3' His tag: For purification. Doubles as 3' primer.

Total length of protein is 33 aa. If do 100 aa SC tags, it's 31.

For 1 x 10^6 cells according to equation
we need 16 million barcodes to cover. 6 AA will cover (see above)

Polymer Units	Barcodes Needed	# AA positions needed	
20	2.00E+01	1	
20	4.00E+02	2	
20	8.00E+03	3	
20	1.60E+05	4	
20	3.20E+06	5	
20	6.40E+07	6	
100	1.00E+02	1	
100	1.00E+04	2	
100	1.00E+06	3	
100	1.00E+08	4	

Figure 11

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 August 2013 (22.08.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/122996 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *G01N 33/58* (2006.01)
C12N 15/11 (2006.01) *G01N 33/50* (2006.01)

(21) International Application Number:
PCT/US2013/025860

(22) International Filing Date:
13 February 2013 (13.02.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/598,513 14 February 2012 (14.02.2012) US

(71) Applicant: **THE JOHNS HOPKINS UNIVERSITY** [US/US]; 3400 North Charles Street, Baltimore, Maryland 21218 (US).

(72) Inventors; and
(71) Applicants : **WANG, Tza-Huei** [US/US]; 2133 East Ridge Road, Timonium, Maryland 21093 (US). **LIU, Kelvin J.** [US/US]; 105 W. 39th St., #1002, Baltimore, Maryland 21218 (US). **SONG, Yunke** [CN/US]; 222 E. University Pkwy., Baltimore, Maryland 21218 (US).

(74) Agent: **CONTRERA, Joseph**; 100 N. Charles Street, 5th Floor, Baltimore, Maryland 21201 (US).

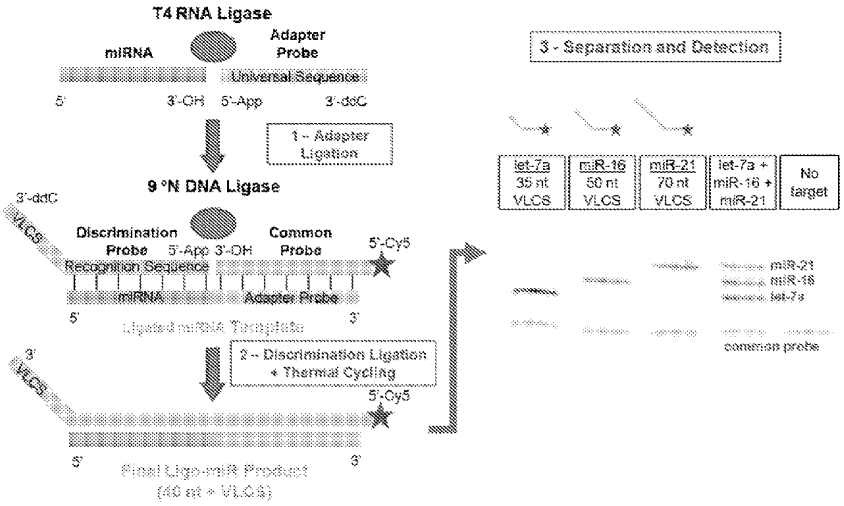
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) **Title**: MIRNA ANALYSIS METHODS

FIGURE 1



(57) **Abstract**: The present invention provides a PCR-free, multiplexed ligation assay for miRNA expression analysis that produces highly quantitative, 10-100 plex miRNA profiling in a single reaction. The inventive methods use a 2-step ligation assay to generate an array of miRNA specific ligation products that can be decoded and quantified by a size discrimination method such as gel electrophoresis or single molecule separation. One embodiment is a low-cost assay that can be performed using standard tools available in nearly all molecular biology laboratories. This assay requires nothing more than a gel apparatus and reader for detection. Other embodiments include use of magnetic beads and other size exclusion apparatus which give increasingly higher sensitivity, lower sample consumption, and reduced processing steps.

WO 2013/122996 A1

MIRNA ANALYSIS METHODS

REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 61/598,513, filed on February 14, 2012, which is hereby incorporated by reference for all purposes as if fully set forth herein.

STATEMENT OF GOVERNMENTAL INTEREST

[0002] This invention was made with U.S. government support under grant nos. 1R01CA155305, 1R43GM103360, and 1U54CA151838. The U.S. government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0003] MicroRNAs (miRNA) are short, noncoding RNAs with pervasive roles throughout gene expression in cellular processes such as differentiation and disease states such as cancer. Uncovering the roles of these molecules in development and tumorigenesis are key steps to the discovery of robust, new biomarkers and potential disease cures.

[0004] Existing methods to analyze miRNA expression consist of single-target PCR based methods and highly multiplexed array methods. PCR is very sensitive but can only analyze a single target. Thus, detection of miRNA panels requires large numbers of parallel reactions, greatly increasing cost, sample consumption, and complexity. Existing array based methods can analyze hundreds of miRNA in a single reaction but have high cost per sample, low sample throughput, and limited sensitivity. Successful clinical validation and translation of these promising miRNA panels will require that 10-100 miRNA be accurately quantified in a cost-effective, high-throughput, and robust manner. No existing technology can bridge this gap between the highly multiplexed but expensive array-based methods and the sensitive but singleplex qPCR methods. Currently, no method is able to achieve the high sensitivity, multiplex detection necessary for low cost miRNA profiling in rare clinical samples.

[0005] Thus, there exists a need for a low cost, PCR-free, multiplexed assay for miRNA expression analysis that performs highly quantitative, 10-100 plex miRNA profiling in a single reaction in less time than currently existing methods.

SUMMARY OF THE INVENTION

[0006] To address this need, the inventors have developed an inexpensive and facile multiplex, ligation assay targeting short RNA sequences, such as miRNA for RNA profiling.

[0007] In accordance with one or more embodiments the present invention provides novel compositions and methods for miRNA analysis. These inventive compositions and methods are collectively named the "Ligo-miR TM" assays.

[0008] In an embodiment termed the "Ligo-miR EZ assay", the embodiment leverages the high multiplex capability and high specificity of the ligation mechanism for PCR-free and RT-free analysis of miRNAs in rare clinical samples. The methods of the present invention use a 2-step ligation process to create a fast (~4 hours), sensitive ($<10^{-18}$ moles), and highly specific (>1000:1) miRNA profiling tool. When compared to the dominant microarray technologies, assay cost will be 20-fold less, assay times will be 5-fold shorter, and sensitivity will be 100-fold higher.

[0009] The compositions and methods of the present invention use a novel 2-step ligation process to generate size encoded miRNA ligation products that can be individually identified and quantified by gel separation (Ligo-miR EZ), single molecule analysis (termed "Ligo-miR HD"), capillary electrophoresis, chromatography, or any other nucleic acid sizing method. A 50 cycle linear amplification (non-PCR) can be integrated into the second ligation step to boost sensitivity while maintaining excellent linearity.

[0010] In one or more embodiments, the present invention provides methods for detecting miRNA biomarkers in the blood, serum, sputum, urine, stool, or other body fluid of subjects, which comprise a noninvasive diagnostic or prognostic technology that is sufficiently sensitive to detect or predict the outcome or response of disease states such as oncogenic, cancerous, premalignant or metaplastic changes in a biological sample of a mammalian subject.

[0011] In one or more embodiments, the present invention provides methods for detecting miRNA biomarkers in the tissue of subjects, which comprise a diagnostic or prognostic technology that is sufficiently sensitive to detect or predict the outcome or response of disease states such as oncogenic, cancerous, premalignant or metaplastic changes in a biological sample of a mammalian subject.

[0012] In accordance with an embodiment, the present invention provides a method for detection of one or more target RNA of interest in a sample comprising: a) obtaining a

sample containing one or more target RNA of interest; b) adding to the sample of a) a sufficient amount of adapter probes and a sufficient amount of a first ligase c) performing a first ligation step by incubation and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more target RNA of interest in the sample; d) adding to the sample of b) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase; e) performing a second ligation step by incubating the sample of c) for a sufficient amount of time to effect hybridization of probes of d) with ligated target RNA of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; and f) analysis of the products of by one or more size discrimination methods.

[0013] In accordance with another embodiment, the present invention provides a method for detection of one or more a target RNA of interest in a sample comprising: a) obtaining a sample containing one or more target RNA of interest; b) adding to the sample of a) a sufficient amount of adapter probes and a sufficient amount of a first ligase and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more target RNA of interest in the sample; c) adding to the sample of b) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase; d) thermocycling the sample of c) for a sufficient amount of time to effect hybridization of probes of c) with ligated target RNAs of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; e) separation of the products of d) by gel electrophoresis; f) detecting the products of e) through imaging of the gel.

[0014] In accordance with a further embodiment, the present invention provides a method for detection of one or more miRNA of interest in a sample comprising: a) obtaining a sample containing one or more miRNA of interest; b) adding to the sample of a) a sufficient amount of adapter probes and a sufficient amount of a first ligase and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more miRNA of interest in the sample; c) adding to the sample of b) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase; d) thermocycling the sample of c) for a sufficient amount of time to effect hybridization of probes of c) with ligated miRNA of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; e) analysis of the products by size discrimination.

[0015] In accordance with an embodiment, the present invention provides a method for detection of one or more miRNA of interest in a sample comprising: a) obtaining a sample

containing one or more miRNA of interest; b) adding to the sample of a) a sufficient amount of adapter probes and a sufficient amount of a first ligase and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more miRNA of interest in the sample; c) adding to the sample of b) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase; d) thermocycling the sample of c) for a sufficient amount of time to effect hybridization of probes of c) with ligated miRNAs of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; e) separation of the products of d) using single molecule free solution hydrodynamic separation (SML-FSHS); f) detecting the products of e) through analysis of peaks in the resulting chromatograph.

[0016] In accordance with a further embodiment, the present invention provides a method for detection of one or more miRNA of interest in a sample comprising: a) obtaining a sample containing one or more miRNA of interest; b) adding to the sample of a) a sufficient amount of adapter probes bound to a plurality of magnetic particles, and a sufficient amount of a first ligase and allowing the adapter probes bound to the magnetic particles to be ligated to the 3'-OH ends of the one or more miRNA of interest in the sample; c) magnetically isolating the particles having the miRNA of interest ligated to the bound adapter probes; d) washing the particles of c) and resuspending the particles in buffer; e) adding to a sample of the particles of d)) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase, and water; f) thermocycling the sample of e) for a sufficient amount of time to effect hybridization of probes of e) with ligated miRNAs of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; g) separation of the products of f) by gel electrophoresis; h) detecting the products of g) through imaging of the gel.

[0017] In accordance with another embodiment, the present invention provides a method for high efficiency and low bias ligation of an adapter probe to a short RNA sequence comprising: a) obtaining a sample containing one or more target RNA of interest; b) adding to the sample a saturating amount of adapter probe; c) adding to the sample a saturating amount of ligase; d) adding to the sample a saturating amount of PEG; and e) performing a ligation step by incubation and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more target RNA of interest in the sample.

[0018] In accordance with an embodiment, the present invention provides methods of diagnosis of the presence of a disease in a subject comprising obtaining a biological sample

from the subject and measuring the expression levels of one or more miRNAs of interest which are associated with the presence of the disease or condition using the methods of the present invention.

[0019] In accordance with an embodiment, the present invention provides methods of generating a miRNA profile of at least one or more tumors in a subject comprising obtaining a biological sample from the subject and measuring the expression levels of one or more miRNAs of interest which are associated with a particular cancer profile using the methods of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] Figure 1 shows a schematic of the overall Ligo-miR TM analysis methods of the present invention.

[0021] Figure 2 shows a gel using the PCR-free, 6-plex Ligo-miR TM assay of the present invention to perform multiplex detection of *let-7a*, *miR-16*, *miR-21, miR-26a, miR-29b, and miR-34a* using a single reaction. The respective length encoded ligation products are only formed when the correct miRNA target is present, enabling identification by size. No cross-talk is seen between the probe sets.

[0022] Figure 3 shows a typical gel image of 6-plex analysis of titrated miRNA where all 6 miRNA targets are serially diluted simultaneously.

[0023] Figure 4 depicts a titration curve of miR-21 obtained using Ligo-miR EZ. High sensitivity ($<2.5 \times 10^{-18}$ moles), excellent linearity, and high repeatability are achieved. Shown are the average and standard deviation from 3 independent experiments. The microfluidic Ligo-miR HD will enable even higher yoctomole level sensitivity (10^{-23} moles).

[0024] Figure 5 depicts an illustration of a SML-FSHS microfluidic device used for Ligo-miR HD analysis which is fabricated from anodic bonding of silicon to glass. Separation is performed in the buffer-filled serpentine separation channel after which cylindrical illumination confocal spectroscopy (CICS) based single molecule detection is used to analyze the separated peaks. The CICS laser sheet spans the entire channel cross-section for 100% mass detection efficiency (inset). The dimensions of the channels are controlled by balancing silicon etch depth with the thickness of the silicon dioxide (right).

[0025] Figure 6 depicts a typical chromatogram of 20-plex miRNA analysis performed using Ligo-miR HD. SML-FSHS is used to identify the Ligo-miR TM products by size and quantify them by single molecule counting. Each Ligo-miR TM product is identified as a size specific peak based on the length of its variable length coding sequence. Quantity is determined by single molecule counting of the Alexa-647 labels.

[0026] Figure 7 shows a schematic of the Ligo-miR Bead embodiment of the present invention. This embodiment uses magnetic beads and a 2-step ligation to integrate miRNA capture and multiplex detection into a single streamlined process. The first ligation captures miRNA from serum and forms miRNA templates concentrated on the bead surface. A magnetic isolation removes all background impurities. The second ligation includes a 50-fold linear amplification and generates single stranded miRNA specific ligation products which are identified and quantified by single molecule analysis.

[0027] Figure 8 shows a gel image and capture efficiency analysis from adapter probe ligation to miRNA using the high efficiency and low bias first step ligation only. (a) Cy3-labeled miRNAs were captured using an optimized protocol. All miRNAs are uniformly captured with high efficiency. (c) Quantified miRNA capture efficiency of (a).

DETAILED DESCRIPTION OF THE INVENTION

[0028] By "nucleic acid" as used herein includes "polynucleotide," "oligonucleotide," and "nucleic acid molecule," and generally means a polymer of DNA or RNA, which can be single-stranded or double-stranded, synthesized or obtained (e.g., isolated and/or purified) from natural sources, which can contain natural, non-natural or altered nucleotides, and which can contain a natural, non-natural or altered internucleotide linkage, such as a phosphoroamidate linkage or a phosphorothioate linkage, instead of the phosphodiester found between the nucleotides of an unmodified oligonucleotide. It is generally preferred that the nucleic acid does not comprise any insertions, deletions, inversions, and/or substitutions. However, it may be suitable in some instances, as discussed herein, for the nucleic acid to comprise one or more insertions, deletions, inversions, and/or substitutions.

[0029] In an embodiment, the nucleic acids of the invention are recombinant. As used herein, the term "recombinant" refers to (i) molecules that are constructed outside living cells by joining natural or synthetic nucleic acid segments to nucleic acid molecules that can replicate in a living cell, or (ii) molecules that result from the replication of those described in

(i) above. For purposes herein, the replication can be in vitro replication or in vivo replication.

[0030] As used herein, the term "target RNA of interest" can be any short ribonucleic acid sequence. The length of such a target RNA sequence is from about 8 to about 100 nucleotides. In some embodiments, the target RNA could be longer that 100 nucleotides. Examples of target RNAs include miRNA, siRNA and other RNA sequences that are useful in diagnosis and treatment of disease and other uses. In accordance with an embodiment, the target RNA of interest is miRNA.

[0031] The nucleic acids used as primers in embodiments of the present invention can be constructed based on chemical synthesis and/or enzymatic ligation reactions using procedures known in the art. See, for example, Sambrook et al. (eds.), *Molecular Cloning, A Laboratory Manual*, 3rd Edition, Cold Spring Harbor Laboratory Press, New York (2001) and Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing Associates and John Wiley & Sons, NY (1994). For example, a nucleic acid can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed upon hybridization (e.g., phosphorothioate derivatives and acridine substituted nucleotides). Examples of modified nucleotides that can be used to generate the nucleic acids include, but are not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N^6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N^6-substituted adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N^6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, 3-(3-amino-3-N-2-carboxypropyl) uracil, and 2,6-diaminopurine. Alternatively, one or more of the nucleic acids of the invention can be purchased from companies, such as Integrated DNA Technologies (Coralville, IA) and Genscript (Piscataway, NJ).

[0032] The term "isolated and purified" as used herein means a protein that is essentially free of association with other proteins or polypeptides, e.g., as a naturally occurring protein

that has been separated from cellular and other contaminants by the use of antibodies or other methods or as a purification product of a recombinant host cell culture.

[0033] The term “biologically active” as used herein means an enzyme or protein having structural, regulatory, or biochemical functions of a naturally occurring molecule.

[0034] As used herein, the term "subject" refers to any mammal, including, but not limited to, mammals of the order Rodentia, such as mice and hamsters, and mammals of the order Logomorpha, such as rabbits. It is preferred that the mammals are from the order Carnivora, including Felines (cats) and Canines (dogs). It is more preferred that the mammals are from the order Artiodactyla, including Bovines (cows) and Swines (pigs) or of the order Perssodactyla, including Equines (horses). It is most preferred that the mammals are of the order Primates, Ceboids, or Simoids (monkeys) or of the order Anthropoids (humans and apes). An especially preferred mammal is the human.

[0035] In accordance with an embodiment, the present invention provides A method for detection of one or more target RNA of interest in a sample comprising: a) obtaining a sample containing one or more target RNA of interest; b) adding to the sample of a) a sufficient amount of adapter probes and a sufficient amount of a first ligase c) performing a first ligation step by incubation and allowing the adapter probes to be ligated to the 3’-OH ends of the one or more target RNA of interest in the sample; d) adding to the sample of b) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase; e) performing a second ligation step by incubating the sample of c) for a sufficient amount of time to effect hybridization of probes of c) with ligated target RNA of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; f) analyzing the products of d) by a size discrimination method.

[0036] Size discrimination methods used in accordance with the inventive methods can be any techniques known in the art, such as, for example, gel electrophoresis, capillary electrophoresis, high pressure liquid chromatography, mass spectrometry, fluorescence spectroscopy, and other methods.

[0037] It will be understood by those of ordinary skill in the art that the inventive methods disclosed herein are capable of both detecting the target RNA of interest as well as quantifying the amount of target RNA in a sample. The quantification of target RNA of interest allows elucidation of expression levels of the target RNA in the sample.

[0038] In some embodiments of the above method, the adapter probe is bound to a magnetic bead at the 3' end.

[0039] In some alternate embodiments of the above method, a magnetic separation, wash, and resuspension are performed after c) first ligation step. It will be understood by those of ordinary skill, that the magnetic beads and methods using them in separations are those commonly used in the art.

[0040] In accordance with one or more embodiments, the inventive methods can include heat denaturing steps are performed post c) the first ligation step and e) the second ligation step. Heat denaturing, as used herein, can include direct heat and incubation of the sample for a set period of time, or can be accomplished through the use of a thermocycling apparatus commonly used in the nucleic acid arts. In some embodiments, the incubation step comprises thermocycling between 1-100 cycles of between about 0.25 to about 5 minutes @ 45 ºC to about 95 ºC denaturation followed by between about 0.25 to about 100 minutes @ 4 ºC to about 80 ºC ligation. In a preferred embodiment, the incubation step comprises thermocycling between 1-100 cycles of about 30 s @ 95 ºC denaturation followed by about 30 s @ 45 ºC ligation.

[0041] In accordance with one or more embodiments, the inventive methods can also include PCR amplification of the reaction products is performed after e) the second ligation step but before f) size discrimination.

[0042] In alternative embodiments which include PCR amplification, one PCR primer is anti-sense to the VLCS sequence and the second primer is anti-sense to the adapter probe sequence.

[0043] In accordance with one or more embodiments, the inventive methods can be used to identify target RNA of interest, including miRNA, siRNA, or other RNA sequences comprising between 8 to 150 nucleotides in length.

[0044] It will be understood by those of ordinary skill in the art that the methods of the present invention comprise two general steps. The first step being the first ligation step wherein the target RNA of interest is incubated with the adapter probes and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more target RNA of interest in the sample. The second general step being the second ligation step wherein the addition to the sample, a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase, and incubating the sample for a sufficient amount of time to effect hybridization of probes with ligated target RNA of interest and adapter probes and to allow ligation between the discrimination probes and common probes. In accordance with some embodiments, the two steps of these methods can be separated and performed at substantially

different times. For example, the first step can be applied to a sample and stored for later analysis using the second step.

[0045] In alternative embodiments, only the first step is applied to the sample and then alternative analytical methods can be applied, such as PCR, qPCR, microarray, sequencing, or other methods known in the art. Due to their short length, the first step of nearly all microRNA analysis techniques is to lengthen the microRNA target through either poly(A)-tailing or, more commonly, ligation. MicroRNA capture through adapter probe ligation is the pervasive first step in the majority of PCR-, microarray-, multiplexing-, and sequencing-based assays. However, this ligation can introduce substantial bias to the resultant to microRNA expression profile analysis. This inventive embodiment enables the downstream assessment of high fidelity expression profiles, using any number of techniques, that are largely free from the blurring effects of ligation bias due to the high efficiency and low bias adapter ligation. Key assay parameters such as PEG%, adapter probe concentration, ligase type, ligase amount, incubation time, incubation temperature, and adapter probe design are used to reduce ligation preference.

[0046] In an alternative embodiment, 100 - 1000 units of ligase, 1 - 12 pmol adenylated adapter probe, 10 - 40% PEG8000 or PEG4000, and T4 RNA ligase reaction buffer (New England Biolabs) are added to 1 - 10,000 ng of total RNA to form a 5 - 100 μL reaction mixture. The reaction is incubated at 4 - 95 °C for 0.25 - 24 hours. Potential ligases comprise T4 RNA ligase 1, T4 RNA ligase 2 truncated, T4 RNA ligase 2 K227Q, T4 RNA ligase 2 truncated KQ, and *Mth.* RNA ligase.

[0047] In an alternative embodiment, 200 - 400 units of T4 RNA Ligase 2 truncated K227Q, 4 - 8 pmol adenylated adapter probe, 20-30% PEG8000 or PEG4000, and T4 RNA ligase reaction buffer (New England Biolabs) are added to 100 - 1000 ng of total RNA to form a 10 - 30 μL reaction mixture. The reaction is incubated at 4 - 37 °C for 1 - 8 hours.

[0048] In yet another alternative embodiment, 300 units of T4 RNA Ligase 2 truncated K227Q, 6 pmol adenylated adapter probe, 25% PEG8000, and T4 RNA ligase reaction buffer (New England Biolabs) are added to 500 ng of total RNA to form a 20 μL reaction mixture. The reaction is incubated at 25°C for 4 hours.

[0049] In an alternative embodiment, the adapter probe is comprised of DNA bases. In another embodiment, the adapter probe is comprised of a mixture of DNA and RNA bases. In yet another embodiment, the adapter probe is comprised of RNA bases at or near the ligation site. As shown in Figure 8, this method is able to achieve high efficiency (85% average

capture efficiency) and low bias (11% CV) across a 20-plex miRNA panel. Such high efficiency and low bias has not previously been achieved. By using a PEG8000 level that is substantially higher than commonly recommended (25% vs. 15%), saturating adapter probe levels (≥ 6 pmol for 500 ng of total RNA), and a saturating enzyme amount (≥300 units for 500 ng of total RNA), ligation efficiency is greatly enhanced and bias is significantly suppressed. Examples of adapter probe sequences are given in Table 1.

[0050] In alternative embodiments, only the second step is applied to the sample and alternative analytical methods are applied to the microRNA beforehand such as reverse transcription or poly (A) tailing.

[0051] In accordance with one or more embodiments of the present invention, it will be understood that the types of cancer diagnosis which may be made, using the methods provided herein, is not necessarily limited. For purposes herein, the cancer can be any cancer. As used herein, the term "cancer" is meant any malignant growth or tumor caused by abnormal and uncontrolled cell division that may spread to other parts of the body through the lymphatic system or the blood stream.

[0052] The cancer can be a metastatic cancer or a non-metastatic (e.g., localized) cancer. As used herein, the term "metastatic cancer" refers to a cancer in which cells of the cancer have metastasized, e.g., the cancer is characterized by metastasis of a cancer cells. The metastasis can be regional metastasis or distant metastasis, as described herein.

[0053] Other diseases or conditions where analysis of miRNA and other RNA using the methods of the present invention can be applied include cardiovascular disease, liver disease, neurological disorders, psychiatric disorders, diabetes, sepsis, arthritis, viral infection, Alzheimer disease, and immune disorders.

[0054] The terms "treat," and "prevent" as well as words stemming therefrom, as used herein, do not necessarily imply 100% or complete treatment or prevention. Rather, there are varying degrees of treatment or prevention of which one of ordinary skill in the art recognizes as having a potential benefit or therapeutic effect. In this respect, the inventive methods can provide any amount of any level of diagnosis, staging, screening, or other patient management, including treatment or prevention of cancer in a mammal. Furthermore, the treatment or prevention provided by the inventive method can include treatment or prevention of one or more conditions or symptoms of the disease, e.g., cancer, being treated or prevented. Also, for purposes herein, "prevention" can encompass delaying the onset of the disease, or a symptom or condition thereof.

[0055] "Complement" or "complementary" as used herein to refer to a nucleic acid may mean Watson-Crick (e.g., A-T/U and C-G) or Hoogsteen base pairing between nucleotides or nucleotide analogs of nucleic acid molecules.

[0056] "Identical" or "identity" as used herein in the context of two or more nucleic acids or polypeptide sequences may mean that the sequences have a specified percentage of residues that are the same over a specified region. The percentage may be calculated by optimally aligning the two sequences, comparing the two sequences over the specified region, determining the number of positions at which the identical residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the specified region, and multiplying the result by 100 to yield the percentage of sequence identity. In cases where the two sequences are of different lengths or the alignment produces one or more staggered ends and the specified region of comparison includes only a single sequence, the residues of single sequence are included in the denominator but not the numerator of the calculation. When comparing DNA and RNA, thymine (T) and uracil (U) may be considered equivalent. Identity may be performed manually or by using a computer sequence algorithm such as BLAST or BLAST 2.0.

[0057] "Probe" as used herein may mean an oligonucleotide capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. Probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. There may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids described herein. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. A probe may be single stranded or partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. Probes may be directly labeled or indirectly labeled such as with biotin to which a streptavidin complex may later bind.

[0058] "Substantially complementary" used herein may mean that a first sequence is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% or 99% identical to the complement of a second sequence over a region of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,

20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more nucleotides, or that the two sequences hybridize under stringent hybridization conditions.

[0059] "Substantially identical" used herein may mean that a first and second sequence are at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% or 99% identical over a region of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more nucleotides or amino acids, or with respect to nucleic acids, if the first sequence is substantially complementary to the complement of the second sequence.

[0060] "Target" as used herein can mean an oligonucleotide or portions or fragments thereof, which may be bound by one or more probes under stringent hybridization conditions. "Target" as used herein may also mean a specific miRNA or portions or fragments thereof, which may be bound by one or more probes under stringent hybridization conditions.

[0061] The nucleic acids of the present invention may also comprise a sequence of a miRNA or a variant thereof. The miRNA sequence may comprise from 13-33, 18-24 or 21-23 nucleotides. The miRNA may also comprise a total of at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 50, 60, 70, 80, 90 and up to 100 nucleotides. The sequence of the miRNA may be the first 13-33 nucleotides of the pre- miRNA. The sequence of the miRNA may also be the last 13-33 nucleotides of the pre- miRNA.

[0062] As used herein, the term "a first ligase and/or second ligase" means many different types of ligase can be used in accordance with the methods of the present invention. Examples of such ligases include E. Coli DNA ligase, Taq DNA ligase, 9 °N DNA ligase, T4 DNA ligase, T4 RNA ligase 1, T4 RNA ligase 2, Ampligase, genetically modified variants of natural DNA / RNA ligases, and others known in the art.

[0063] In accordance with one or more embodiments, the concentrations of reagents such as ligases and probes can be in "saturating amounts." As used herein, a saturating amount of reagent (enzyme, probe, reaction time, etc.) is the point at which increasing beyond that point causes no further increase in reaction efficiency. That point is found by spiking each miRNA into a background of total RNA and slowly increasing the parameter (enzyme, probe, reaction time) until no further increase in miRNA capture efficiency is seen. It is critical that this is performed while spiking into a total RNA background as opposed to pure buffer because the optimum conditions obtained using a total RNA background are quite different than those obtained using a buffer background.

[0064] In accordance with one or more embodiments, additives can be used in the methods of the present invention. For example, additives such as PEG or DMSO are added to increase ligation efficiency and reduce ligation bias.

[0065] In accordance with one or more embodiments, internal controls are added to account for variations in ligation efficiency.

[0066] A method of identifying a nucleic acid associated with a disease or a pathological condition is also provided. The method comprises measuring a level of the nucleic acid in a sample that is different than the level of a control. In accordance with an embodiment, the nucleic acid is a miRNA and the detection may be performed by contacting the sample with a probe or biochip described herein and detecting the amount of hybridized product.

[0067] The level of the nucleic acid in the sample may also be compared to a control cell (e.g., a normal cell) to determine whether the nucleic acid is differentially expressed (e.g., overexpressed or underexpressed). The ability to identify miRNAs that are differentially expressed in pathological cells compared to a control can provide high-resolution, high-sensitivity datasets which may be used in the areas of diagnostics, prognostics, therapeutics, drug development, pharmacogenetics, biosensor development, and other related areas.

[0068] The expression level of a disease-associated nucleic acid or miRNA provides information in a number of ways. For example, a differential expression of a disease-associated nucleic acid compared to a control may be used as a diagnostic that a patient suffers from the disease. Expression levels of a disease-associated nucleic acid may also be used to monitor the treatment and disease state of a patient. Furthermore, expression levels of a disease-associated miRNA may allow the screening of drug candidates for altering a particular expression profile or suppressing an expression profile associated with disease.

[0069] A target nucleic acid or portions or fragments thereof, may be detected and levels of the target nucleic acid measured by contacting a sample comprising the target nucleic acid with a biochip comprising an attached probe sufficiently complementary to the target nucleic acid and detecting hybridization to the probe above control levels.

[0070] In accordance with another embodiment of the present invention, it will be understood that the term "biological sample" or "biological fluid" includes, but is not limited to, any quantity of a substance from a living or formerly living patient or mammal. Such substances include, but are not limited to, blood, serum, plasma, urine, cells, organs, tissues, bone, bone marrow, lymph, lymph nodes, synovial tissue, chondrocytes, synovial

macrophages, endothelial cells, and skin. In a preferred embodiment, the fluid is blood or serum.

[0071] A method of diagnosis is also provided. The method comprises detecting a differential expression level of two or more disease-associated miRNAs in a biological sample. The sample may be derived from a subject. Diagnosis of a disease state in a subject may allow for prognosis and selection of therapeutic strategy. Further, the developmental stage of cells may be classified by determining temporarily expressed disease-associated miRNAs.

[0072] In accordance with an embodiment, the present invention provides methods for diagnosis of a disease or condition in a subject comprising: a) obtaining a biological sample from the subject; b) analyzing the sample of a) for one or more target RNA of interest using the methods described above; c) comparing the expression levels of the one or more target RNA of interest in the sample to the expression levels of the one or more target RNA of interest in a control sample; d) diagnosing the subject as having the disease or condition when detecting the expression levels of one or more target RNA of interest which are associated with the presence of the disease or condition in the sample of the subject when compared to the control sample.

[0073] In some embodiments of the methods for diagnosis of a disease or condition in a subject, the target RNA is miRNA, siRNA or other RNA sequences comprising between 8 to 150 nucleotides in length.

[0074] The identity and relative quantity of miRNAs in a sample can be used to provide a miRNA profiles for a particular sample. A miRNA profile for a sample includes information about the identities of miRNAs contained in the sample, quantitative levels of miRNAs contained in the sample, and/or changes in quantitative levels of miRNAs relative to another sample. For example, a miRNA profile for a sample includes information about the identities, quantitative levels, and/or changes in quantitative levels of miRNAs associated a particular cellular type, process, condition of interest, or other cellular state. Such information can be used, for diagnostic purposes, drug development, drug screening and/or drug efficacy testing.

[0075] In another example, with regard to diagnostics, if it is known that the presence or absence of a particular miRNA or group of miRNAs is associated with the presence or absence of a particular condition of interest, then a diagnosis of the condition can be made by obtaining the miRNA profile of a sample taken from a patient being diagnosed.

EXAMPLES

[0076] Fundamental Ligo-miR TM Assay Principles. In accordance with one or more embodiments, the inventive Ligo-miR TM methods provided herein comprise a PCR-free, 2-step ligation assay for multiplex miRNA detection that encodes and detects miRNA of interest based on probe length. In a first ligation step, a universal adapter probe is ligated to the 3'-OH of every miRNA in a sample using a first ligase to form a ligated oligonucleotide template. In a second ligation step, miRNA of interest specific discrimination probes and a universal common probe are then added to the solution and allowed to hybridize to the miRNA-adapter templates and ligated using a second ligase to form the final Ligo-miR TM reaction products.

[0077] In accordance with an embodiment, the common probe and adapter probe are complementary to each other and shared across all miRNA. Each discrimination probe is specifically complimentary to each miRNA of interest at the 5' end, and which also comprises a variable length coding sequence portion (VLCS) at its 3' end. Each discrimination probe comprises a corresponding 5 – 150 nt VLCS that is used to identify the specific miRNA probe by its unique length using size discrimination methods such as electrophoresis. Thus, in a set of probes to be used for multiplex miRNA detection there will be 1 discrimination probe for each miRNA of interest being detected, a single adapter probe and a single common probe. Only in the presence of the specific miRNA of interest, will the specific discrimination probe and common probe be ligated together in the second step. This 2-step ligation process ensures that only mature miRNA are detected with high specificity. In the first ligation step, the adapter probe must be ligated immediately adjacent to the mature miRNA sequence for the second ligation step to even occur.

[0078] Detection of Ligo-miR TM Reaction Products. In accordance with another embodiment, the analysis of the reaction products is performed through detection of the fluorescent label attached to the 5' end of the common probe. Each specific miRNA leads to a corresponding reaction product which is then identified by length. In accordance with an alternative embodiment, the discrimination probes are additionally fluorescently labeled at the 3' end. Analysis of the products is performed through coincidence analysis of the 3' discrimination probe signal and the 5' common probe signal that occurs only when ligation has occurred.

[0079] In accordance with a further embodiment, the discrimination probes and common probe are labeled with fluorescent dyes such that fluorescent resonance energy transfer (FRET) occurs when the ligation product has formed.

[0080] In accordance with still another embodiment, the discrimination probes are internally labeled rather than end labeled. The discrimination probes can be labeled with the same or different labels as well.

[0081] It will be understood by those of skill in the art that in certain embodiments, the probes are detectably labeled by any known means, including, for example, radio-labeled, fluorescently labeled, colorimetrically labeled, etc.

[0082] In accordance with the inventive methods, the detection of the hybridized product length can be performed by gel electrophoresis, capillary electrophoresis, free solution hydrodynamic separation, chromatography methods, etc.

[0083] Alternative variations of Ligo-miR TM assay methods of the present invention. In accordance with an embodiment, a single adapter probe, a single common probe, and a set of discrimination probes are designed to perform multiplex detection where the position of the second ligation site, on the opposite strand, is exactly opposite the first ligation site. In accordance with another embodiment, the second ligation site is shifted in either the 3' or 5' direction such that it is offset from the first ligation site to increase specificity.

[0084] In accordance with yet another embodiment, the second ligation site is shifted and multiple common probes, multiple adapter probes, and a set of discrimination probes are designed to enable multiplexing by color and length.

[0085] In accordance with one or more embodiments, the first ligation is performed in the absence of ATP.

[0086] In accordance with one or more embodiments, the second ligation is performed in the absence of ATP.

[0087] In accordance with a further embodiment, polymerase chain reaction (PCR) primer sites are designed into the probe sequences to enable incorporation of a downstream PCR amplification step. The 3' end of all the discrimination probes are designed to contain a common PCR primer site. A forward primer, anti-sense to this site, is synthesized. A reverse primer, anti-sense to the adapter probe sequence, is also synthesized. In this manner, all the miRNA specific ligation products can be PCR amplified using a single PCR primer set. PCR amplification is performed after the second ligation step by adding the aforementioned PCR primers, polymerase, and PCR reagents to the ligation products and thermocycling. The

discrimination probes for miR-26a, miR-34a, and miR-29b listed in Table 1 include such a common PCR primer site.

[0088] In accordance with still another embodiment, poly(A) tailing is used in the first step instead of ligation and a poly(T) common probe is used is the second step. A sufficient amount of poly(A) polymerase, poly(A) polymerase reaction buffer, and ATP are added to the sample containing the RNA of interest and incubated at 37 ºC for 10 minutes. This adds a poly(A) tail to the end of each RNA molecule in the sample. The polymerase is then heat inactivated at 65 ºC for 20 minutes. A poly(T) common probe is then used and the assay proceeds accordingly. In yet another embodiment 5 units of poly(A) polymerase, 1X poly(A) polymerase reaction buffer, and 1 mM ATP are added to 500 ng of total RNA and incubated at 37 ºC for 10 minutes.

[0089] Adapter Probe Design. The adapter probes used in the methods of the present invention are commercially synthesized, HPLC or PAGE purified, DNA oligonucleotides. In accordance with an embodiment, a typical adapter probe is about 19 nt in length with a 5'-App (adenylation) and 3'-ddC (dideoxycytosine).

[0090] In accordance with another embodiment, the 3' terminus of the adapter probe is blocked by -ddC to prevent unwanted ligation products. Other blocking moieties known in the art may also be used, for example, an NH_2 moiety.

[0091] In accordance with one or more alternative embodiments, the 3' terminus may also be labeled with a detectable moiety known in the art, for example, an indicator such as Cy3, Cy5, Alexa488, Alexa647, FITC, etc. can be used.

[0092] In accordance with one or more embodiments, the adapter probe comprises a detectable internal label.

[0093] In accordance with one or more embodiments, the 5' terminus of the adapter probe is pre-adenylated so that ligation can be performed in the absence of ATP, reducing unwanted ligation products. Pre-adenylation can be performed using known methods either enzymatically or chemically.

[0094] In accordance with another embodiment, the 5' residue is RNA rather than DNA to reduce ligation bias and to increase ligation efficiency.

[0095] In accordance with still another embodiment, alternative nucleic acid residues such as locked nucleic acids (LNAs) or peptide nucleic acids (PNAs) may be substituted at various positions to increase binding specificity. It will be understood by those of ordinary skill in the art that LNAs can significantly enhance the binding characteristics of nucleic acid

probes. LNA bases can be substituted into the recognition sequence of the discrimination probes as well as the adapter and/or common probes. The substitution of LNA bases raises the overall melting temperatures of the individual probes. Such a strategy can be used to: 1) increase probe specificity for distinguishing closely related miRNA, 2) matching probe melting temperatures to minimize variability in reaction efficiency, and 3) increasing overall reaction efficiency by increasing probe binding strength. Peptide nucleic acids can be used similarly.

[0096] The adapter probe sequences of the present invention can be designed to either eliminate or promote formation of secondary structures, either alone, or in conjunction with the miRNA targets of interest. Such secondary structures can be useful in reducing or increasing ligation efficiency as appropriate. Typical adapter probe sequences are provided in Table 1.

[0097] In accordance with one or more embodiments, the adapter probe comprises a biotin or streptavidin at the 3' end.

[0098] In accordance with one or more embodiments, the adapter probe is bound to a magnetic bead at the 3' end.

[0099] In accordance with one or more embodiments, wherein the adapter probe is comprised of RNA nucleotides.

[0100] Common Probe Design. In accordance with one or more embodiments of the present invention, the common probes are commercially synthesized, HPLC or PAGE purified, DNA oligonucleotides. In accordance with an embodiment, typically, a single common probe is utilized in each probe set. A typical common probe is about 19 nt in length with a sequence that is complimentary to the adapter probe, and has a detectable moiety on the 5' end, for example, 5'-Cy5, and also comprises a 3'-OH end. In an alternative embodiment, the 5' end of the common probe of the present invention can be labeled with other detectable moieties, such as Cy3, Alexa488, Alexa 647, FITC, fluorescein, etc.

[0101] In accordance with one or more embodiments, the common probe comprises 10-30 DNA nucleotides in length and is complementary to the adapter probe sequence and further comprises a 3'-OH.

[0102] In accordance with one or more embodiments, the common probe comprises a detectable internal label.

[0103] In accordance with another embodiment, alternative nucleic acid residues such as locked nucleic acids or peptide nucleic acids may be substituted at various positions to increase binding specificity.

[0104] It will be understood by those of ordinary skill in the art that the common probe sequence can be designed to either eliminate or promote formation of secondary structures, either alone or in conjunction with the miRNA targets. Such secondary structures can be useful in reducing or increasing ligation efficiency as appropriate. Typical common probe sequences are provided in Table 1.

[0105] Discrimination Probe Design. The discrimination probes used with the methods of the present invention are commercially synthesized, HPLC or PAGE purified, DNA oligonucleotides. At least one discrimination probe is required for each miRNA being detected. In accordance with an embodiment, about 5-100 discrimination probes are used in each probe set. Each discrimination probe comprises a recognition sequence at its 5' end that is complimentary to the target miRNA of interest being detected, and also comprises a variable length coding sequence (VLCS) at the 3' end that can be used to identify the hybridized probe by its unique length. In one embodiment, the recognition sequence is exactly the same length as the microRNA target and entirely complementary to the microRNA sequence.

[0106] In some other embodiments, the recognition sequence may vary from full complementary, for example, in another embodiment, the length of the recognition sequence is shorter than the microRNA target while maintaining perfect complementarity. In another embodiment, the length of the recognition sequence is longer than the microRNA target while maintaining perfect complementarity and extending into the adapter sequence.

[0107] In accordance with an embodiment, the discrimination probes comprise a 5' recognition sequence having a DNA oligonucleotide sequence which has antisense complementarity to the specific miRNA of interest and having a 3' variable length coding sequence (VLCS) having at least 5 to about 200 nucleotides in length.

[0108] A VLCS on a discrimination probe, in one or more embodiments, can vary in length from 5 – 150 nt, however, in accordance with other embodiments; the lengths can be made longer or shorter to enable distinct identification. Longer discrimination probe sequences can be generated using known cloning techniques to create probes that are between 100-100,000 bp. The minimum length difference between VLCS tags is determined by the desired level of multiplexing and the sizing resolution and sizing dynamic range of the

downstream size discrimination technique used to identify them. In embodiments using with detection platforms with single nucleotide resolution such as polyacrylamide gel electrophoresis, capillary electrophoresis or, single molecule free solution hydrodynamic separation, VLCS tags are typically separated by 3-10 nucleotides. In embodiments using agarose gel electrophoresis, VLCS tags are typically separated by 10-100 nucleotides.

[0109] In accordance with an embodiment, the 3' terminus of the discrimination probe is blocked by -ddC to prevent unwanted ligation products. In alternative embodiments, other blocking moieties may also be used, for example, such as NH_2. In accordance with a further embodiment, the 3' terminus of the discrimination probe can be labeled with a detectable moiety, for example, such as Cy3, Cy5, Alexa488, Alexa647, FITC, etc.

[0110] In accordance with an embodiment, the 5' terminus of the discrimination probe is pre-adenylated so that ligation can be performed in the absence of ATP, reducing unwanted ligation products. Pre-adenylation can be performed using any known methods in the art, including enzymatically or chemically.

[0111] In accordance with another embodiment, the discrimination probe comprises alternative nucleic acid residues such as locked nucleic acids or peptide nucleic acids may be substituted at various positions to increase binding specificity.

[0112] It will be understood by those of ordinary skill in the art that each discrimination probe comprises 2 regions, a recognition sequence and a VLCS tag. In accordance with an embodiment, the recognition sequence is typically fully complementary to the target miRNA sequence of interest. In accordance with other embodiments, the recognition sequence may vary from having a fully complementary sequence to the target miRNA sequence of interest.

[0113] In accordance with some embodiments, the discrimination probe comprises a detectable internal label.

[0114] In accordance with an embodiment, the VLCS tag typically is comprised of sequences not found within the host genome to prevent unwanted hybridization to background RNA/DNA molecules. In one or more alternative embodiments, the discrimination probe sequences can be designed to either eliminate or promote formation of secondary structures, either alone or in conjunction with the miRNA targets. Such secondary structures can be useful in reducing or increasing ligation efficiency as appropriate. Examples of some discrimination probe sequences are provided in Table 1.

[0115] As used in the methods described herein, the target RNA of interest is analyzed and identified using size discrimination methods. Examples of size discrimination methods include, but are not limited to, gel electrophoresis, capillary electrophoresis and SML-FSHS.

[0116] Ligo-miR EZ methods. Figure 1 shows a schematic of the overall Ligo-miR TM analysis methods of the present invention. In accordance with an embodiment, in the first ligation step, the method begins with obtaining about 0.1 - 5 μL of total RNA sample, and adding a sufficient quantity of adenylated adapter probe (50 - 1000 nM), ligase (50 - 500 units), ligase buffer (0.1 - 5 μL), and water. It will be understood by those of ordinary skill in the art that ligases that can be used in the methods of the present invention include, for example, T4 RNA Ligase 1, T4 RNA Ligase 2 truncated, T4 RNA Ligase 2 truncated KQ, T4 RNA Ligase 2 truncated K227Q, Thermostable 5′ App DNA/RNA Ligase, etc. In certain embodiments, additives such as PEG (polyethylene glycol) (0 - 40%) and DMSO (Dimethyl sulfoxide) (0 - 30%) can also be added to increase ligation efficiency. The final reaction volume is about 10 μL. The mixture is then incubated at 25 °C for about 0.25 to 48 hours and is followed by a 65°C heat denaturing step for about 1-60 minutes, preferably about 1, 2, 5, 10, 15, 20, 30 up to 60 minutes. In some embodiments, the incubation step can also be performed at lower temperatures such as 4 °C or higher temperatures such as 65 °C to reduce ligation bias and increase ligation efficiency.

[0117] In accordance with an embodiment, in the second ligation step, about 5 μL of the previous reaction mixture is obtained and a sufficient quantity of the common probe (10 - 1000 nM), discrimination probes (10 - 1000 nM), ligase (50 - 500 units), ligase buffer (0.1 - 5 μL), and water is added. It will be understood by those of ordinary skill, that in for the second step, ligases that can be used include, for example, T4 DNA Ligase, 9 °N DNA Ligase, Ampligase, T4 RNA Ligase 2, *E. Coli* DNA Ligase, Taq Ligase, etc. In some embodiments, additives such as PEG (polyethylene glycol) (0 - 40%) and DMSO (Dimethyl sulfoxide) (0 - 30%) can also be added to increase ligation efficiency. The final reaction volume is about 10 μL. The reaction is then subjected to thermal cycling between 95 °C (0.5 - 5 minutes hold) and 45 °C (1 - 100 minutes hold) for 1-100 cycles depending on how much amplification is required.

[0118] After thermocycling, in accordance with an embodiment, the sample is then subjected to PAGE (poly-acrylamide gel electrophoresis) to separate the hybridized products. Typically, 5 μL of the sample is mixed with gel loading buffer and analyzed on a denaturing

15% TBE-urea polyacrylamide gel for about 50 minutes at 300V, enabling separation of the microRNA reaction products. In an alternative embodiment, agarose gel electrophoresis can also be used to separate the hybridized products. The separated gel products are imaged on a Typhoon Variable Mode Imager (GE). A typical gel image is provided in Figure 2. Figure 3 shows a typical gel image of 6-plex analysis of titrated miRNA targets. In alternative embodiments, CCD-based or film based imagers can also be used to obtain a gel image. Image analysis is used to identify the microRNA expression levels based on gel band position and integrated band intensity. Figure 4 shows a typical titration curve obtained from the image analysis of miR-21 using Figure 3.

[0119] In accordance with another embodiment, the thermocycled samples can be analyzed using capillary electrophoresis instead of slab gel electrophoresis. The resultant electropherogram can then be analyzed for to obtain microRNA expression levels based on peak position and integrated peak intensity.

[0120] Table 1: Sequences of Adapter, Common, and Discrimination Probes

Probe Name	Length	5' Mod	3' Mod	Sequence 5'-> 3'
rA-1 adapter probe	18	phos	ddC	5'- /5Phos/rAC TGT AGG CAC CAT CAA T/3ddC/ -3' (SEQ ID NO: 1)
rA-2 adapter probe	18	phos	Biotin-TEG	5'- /5Phos/rAC TGT AGG CAC CAT CAA TC/3BiotinTEG/ -3' (SEQ ID NO: 2)
rA-3 adapter probe	19	phos	Cy5	5'- /5Phos/rAC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 3)
rG adapter probe	19	phos	Cy5	5'- /5Phos/rGC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 4)
rU adapter probe	19	phos	Cy5	5'- /5Phos/rUC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 5)
rC adapter probe	19	phos	Cy5	5'- /5Phos/rCC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 6)
dA-1 adapter probe	19	phos	ddC	5'- /5Phos/AC TGT AGG CAC CAT CAA TC/3ddC/ -3' (SEQ ID NO: 7)
dA-2 adapter probe	19	phos	Cy5	5'- /5Phos/AC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 8)
dG adapter probe	19	phos	Cy5	5'- /5Phos/GC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 9)
dT adapter probe	19	phos	Cy5	5'- /5Phos/TC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 10)
dC adapter probe	19	phos	Cy5	5'- /5Phos/CC TGT AGG CAC CAT CAA TC/3Cy5Sp/ -3' (SEQ ID NO: 11)
Common probe	19	Cy5	-	5'- /5Cy5/GA TTG ATG GTG CCT ACA GT -3' (SEQ ID NO: 12)
miR-34a discrimination probe	50	phos	-	5'- /5Phos/AC AAC CAG CTA AGA CAC TGC CAA TCC TGT TAC CAT CGT AGG CAC CTG AAA -3' (SEQ ID

				NO: 13)
let-7a discrimination probe	57	phos	FAM	5'- /5Phos/AA CTA TAC AAC CTA CTA CCT CAA TCC TGT TAC CAG TGG CTG CTG CCA GTG GCG ATA A/36-FAM/ -3' (SEQ ID NO: 14)
miR-29b discrimination probe	64	phos	-	5'- /5Phos/AA CAC TGA TTT CAA ATG GTG CTA ATC CTG TTA CCA GTG GCT GCT GCC ATC GTA GGC ACC TGA AA -3' (SEQ ID NO: 15)
miR-16 discrimination probe	72	phos	-	5'- /5Phos/CG CCA ATA TTT ACG TGC TGC TAA TCC TGT TAC CAG TGG CTG CTG CCA GTG GCG ATA AGT CGT GTC TTA CCG G -3' (SEQ ID NO: 16)
miR-26a discrimination probe	82	phos	-	5'- /5Phos/AG CCT ATC CTG GAT TAC TTG AAA TCC TGT TAC CAG TGG CTG CTG CCA GTG GCG ATA AGT CGT GTC ATC GTA GGC ACC TGA AA -3' (SEQ ID NO: 17)
miR-21 discrimination probe	92	phos	-	5'- /5Phos/TC AAC ATC AGT CTG ATA AGC TAA TCC TGT TAC CAG TGG CTG CTG CCA GTG GCG ATA AGT CGT GTC TTA CCG GGT TGG ACT CAA GAC GAT AGT -3' (SEQ ID NO: 18)

[0121] Ligo-miR HD Methods. For the HD methods of the present invention, in an embodiment, the first and second ligation reactions proceed as detailed above for Ligo-miR TM EZ embodiments.

[0122] Separation and analysis of the hybridized products is performed using single molecule free solution hydrodynamic separation (SML-FSHS). In an embodiment, SML-FSHS is performed using a PicoSep instrument (Circulomics) fitted with a 2 μm inner diameter, 75 cm long fused silica microcapillary at 100 psi with TE buffer as the loading and elution buffer. In an alternative embodiment, a 1 μm inner diameter microcapillary can be used for higher sizing resolution. In certain other embodiments, denaturing buffer conditions can also be used.

[0123] In accordance with yet another embodiment, a microfluidic device can be used instead of a microcapillary to perform SML-FSHS analysis. A schematic illustration of SML-FSHS microfluidic device is shown in Figure 5. The device contains a serpentine shaped microfluidic channel with width and depth of 1 μm and length of 1 m.

[0124] The resultant SML-FSHS chromatogram can then be analyzed for to obtain microRNA expression levels based on peak position, which is correlated to size, and single molecule counts, which is correlated to amount. A representative chromatogram is provided in Figure 6.

[0125] Ligo-miR Bead Methods. Figure 7 shows a schematic of an embodiment of the overall method of the present invention. Adapter probes are coupled to magnetic beads by conjugating 3'-biotin labeled adapter probes with streptavidin coated magnetic beads using standard streptavidin-biotin binding chemistry known in the art. For the first ligation step, about 1.5 μL of raw sample are obtained and adapter probe coupled magnetic beads are added, along with a sufficient quantity of ligase, ligase buffer, and water. Raw samples can include blood, serum, cell lysate, etc. It will be understood by those of ordinary skill that ligases that can be used in the inventive methods include, for example, T4 RNA Ligase 1, T4 RNA Ligase 2 truncated, T4 RNA Ligase 2 truncated KQ, T4 RNA Ligase 2 truncated K227Q, Thermostable 5′ App DNA/RNA Ligase, etc. In some embodiments, additives such as PEG (polyethylene glycol) and DMSO (Dimethyl sulfoxide) can also be added to increase ligation efficiency. The final reaction volume is about 10 μL. Incubate at 25 °C for about 1 to 4 hours, followed by a 65 °C heat denaturing step for about 20 minutes. In alternative embodiments, incubation of the solution can also be done at lower temperatures such as 4 °C or higher temperatures such as 65 °C to reduce ligation bias and increase ligation efficiency. In accordance with one or more embodiments, the present method ligates the 3'-OH present on every miRNA to the 5'-App on the adapter probe-coupled beads to create miRNA templates tethered to the bead surface.

[0126] The beads are then magnetically isolated, washed and resuspended in buffer. For the second ligation step, about 5 μL of the resuspended bead mixture are obtained and add the common probe, discrimination probes, ligase, ligase buffer, and water. As in other embodiments discussed above, ligases that can be used in the inventive methods include T4 DNA Ligase, 9 °N DNA Ligase, Ampligase, T4 RNA Ligase 2, E Coli DNA Ligase, Taq Ligase, etc. In alternative embodiments, additives such as PEG (polyethylene glycol) and DMSO (Dimethyl sulfoxide) can also be added to increase ligation efficiency. The final reaction volume is about 10 μL.

[0127] The above products are then thermal cycled between 95 °C (0.5 - 5 minutes hold) and 45 °C (1 - 100 minutes hold) for 1-100 cycles depending on how much amplification is required.

[0128] The isolated samples are then separated using PAGE. In accordance with one or more embodiments, about, 5 μL of the sample is mixed with gel loading buffer and analyzed on a denaturing 15% TBE-urea polyacrylamide gel for 50 minutes at 300V, enabling separation of the microRNA reaction products. Alternatively, in some embodiments, agarose

gel electrophoresis can also be used. The separated gel is imaged on a Typhoon Variable Mode Imager (GE). Alternatively, depending on the detectable moiety used, CCD-based or film based imagers can also be used to obtain a gel image. Image analysis is used to identify the microRNA expression levels based on gel band position and integrated band intensity.

[0129] In accordance with another embodiment, the sample can also be analyzed using capillary electrophoresis, or in some embodiments the sample can be analyzed using SML-FSHS.

[0130] All references, including publications, patent applications, and patents, cited herein are hereby incorporated by reference to the same extent as if each reference were individually and specifically indicated to be incorporated by reference and were set forth in its entirety herein.

[0131] The use of the terms "a" and "an" and "the" and similar referents in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms "comprising," "having," "including," and "containing" are to be construed as open-ended terms (i.e., meaning "including, but not limited to,") unless otherwise noted. Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., "such as") provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

[0132] Preferred embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Variations of those preferred embodiments may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by

applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

Claims:

1. A method for detection of one or more target RNA of interest in a sample comprising:

a) obtaining a sample containing one or more target RNA of interest;

b) adding to the sample of a) a sufficient amount of adapter probes and a sufficient amount of a first ligase

c) performing a first ligation step by incubation and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more target RNA of interest in the sample;

d) adding to the sample of b) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase;

e) performing a second ligation step by incubating the sample of c) for a sufficient amount of time to effect hybridization of probes of d) with ligated target RNA of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; and

f) analysis of the products of by one or more size discrimination methods.

2. The method of claim 1, wherein the adapter probe comprises 5-50 DNA nucleotides in length and is phosphorylated at the 5' end.

3. The method of claim 2, wherein the adapter probe is pre-adenylated at the 5' end.

4. The method of either of claims 2 or 3, wherein the adapter probe is blocked at the 3' end with ddC or NH_2.

5. The method of any of claims 2 to 4, wherein the adapter probe comprises a detectable label on the 3' end.

6. The method of any of claims 2 to 5, wherein the adapter probe comprises a detectable internal label.

7. The method of any of claims 2 to 6, wherein the adapter probe comprises a biotin or streptavidin at the 3' end.

8. The method of any of claims 2 to 7, wherein the adapter probe is bound to a magnetic bead at the 3' end.

9. The method of claim 8, wherein a magnetic separation, wash, and resuspension are performed after c) first ligation step.

10. The method of any of claims 2 to 9, wherein the adapter probe is comprised of RNA nucleotides.

11. The method of any of claims 1 to 10, wherein the common probe comprises 5-50 DNA nucleotides in length and is complementary to the adapter probe sequence and further comprises a 3'-OH.

12. The method of claim 11, wherein the common probe comprises a detectable label on the 5' end.

13. The method of either of claims 11 or 12, wherein the common probe comprises a detectable internal label.

14. The method of any of claims 1 to 13, wherein the discrimination probes comprise a 5' recognition sequence having a DNA oligonucleotide sequence which has antisense complementarity to the specific miRNA of interest and having a 3' variable length coding sequence (VLCS) having at least 5 to about 200 nucleotides in length.

15. The method of claim 14, wherein the VLCS tag length is specific to one or more miRNAs of interest

16. The method of either of claims 14 or 15, wherein the discrimination probe is blocked at the 3' end with ddC or NH_2.

17. The method of any of claims 14 to 16, wherein the discrimination probe comprises a detectable label on the 3' end.

18. The method of any of claims 14 to 17, wherein the discrimination probe comprises a detectable internal label.

19. The method of either of claims 14 to 18, wherein the discrimination probe is phosphorylated at the 5' end.

20. The method of claim 19, wherein the discrimination probe is pre-adenylated at the 5' end.

21. The method of any of claims 1 to 20, wherein the detectable labels are selected from the group consisting of fluorescent probes, near infra-red probes, FRET probes, luminescent probes and radionuclides.

22. The method of claim 21, wherein the detectable labels are Cy dyes, Alexa Fluors, fluorescein dyes, ATTO dyes, IRDyes, WellRED dyes, and rhodamine dyes.

23. The method of any of claims 1 to 22, wherein the probes comprise modified nucleic acid bases such as LNA or PNA.

24. The method of claim 1, wherein in f) size discrimination is performed using electrophoresis.

25. The method of claim 1, wherein in f) size discrimination is performed using single molecule free solution hydrodynamic separation (SML-FSHS).

26. The method of claim 1, wherein in f) size discrimination is performed using chromatography.

27. The method of claim 1, wherein in c) first ligation is performed in the absence of ATP.

28. The method of claim 1, wherein in e) second ligation is performed in the absence of ATP.

29. The method of claim 1, wherein the incubation comprises thermocycling between 1-100 cycles of between about 0.25 to about 5 minutes @ 45 ºC to about 95 ºC denaturation followed by between about 0.25 to about 100 minutes @ 4 ºC to about 80 ºC ligation.

30. The method of any of claims 1-29, wherein additives such as PEG or DMSO are added to increase ligation efficiency and reduce ligation bias.

31. The method of any of claims 1-30, wherein internal controls are added to account for variations in ligation efficiency.

32. The method of claim 1, wherein the first ligase is T4 RNA ligase 1, T4 RNA ligase 2 truncated, T4 RNA ligase 2 K227Q, T4 RNA ligase 2 truncated KQ, and *Mth.* RNA ligase.

33. The method of claim 1, wherein the second ligase is *E. Coli* DNA ligase, Taq DNA ligase, 9 °N DNA ligase, T4 DNA ligase, T4 RNA ligase 2, and Ampligase.

34. The method of claim 1, wherein heat denaturing steps are performed post c) the first ligation step and e) the second ligation step.

34. The method of claim 1, wherein PCR amplification of the reaction products is performed after e) the second ligation step but before f) size discrimination.

35. The method of claim 34, wherein one PCR primer is anti-sense to the VLCS sequence and the second primer is anti-sense to the adapter probe sequence.

36. The method of any of claims 1 to 35, wherein the target RNA of interest is miRNA, siRNA or other RNA sequences comprising between 8 to 150 nucleotides in length.

37. A method for diagnosis or prognosis of a disease or condition in a subject comprising:

a) obtaining a biological sample from the subject;

b) analyzing the sample of a) for the expression levels of one or more target RNA of interest using the methods of any of claims 1 to 36;

c) comparing the expression levels of one or more target RNA of interest in the sample to the expression levels of one or more target RNA of interest in a control sample;

d) diagnosing the subject as having the disease or condition or prognosing the outcome or response of the disease or condition when the expression levels one or more target RNA of interest are associated with a specific pattern when compared to the control sample.

38. The method of claim 37, wherein the target RNA is miRNA, siRNA or other RNA sequences comprising between 8 to 150 nucleotides in length.

39. A method for high efficiency and low bias ligation of an adapter probe to a short RNA sequence comprising:

a) obtaining a sample containing one or more target RNA of interest;

b) adding to the sample a saturating amount of adapter probe;

c) adding to the sample a saturating amount of ligase;

d) adding to the sample a saturating amount of PEG; and

e) performing a ligation step by incubation and allowing the adapter probes to be ligated to the 3'-OH ends of the one or more target RNA of interest in the sample.

40. The method of claim 39, wherein the reaction volume is 5 - 100 μL.

41. The method of claim 40, wherein the reaction volume is 20 μL.

42. The method of claim 39, wherein the sample comprises 1 - 10,000 ng of total RNA.

43. The method of claim 42, wherein the sample comprises 500 ng of total RNA.

44. The method of claim 39, wherein the amount of ligase is between 100 - 1000 units.

45. The method of claim 44, wherein the amount of ligase is 300 units.

46. The method of claim 39, wherein the PEG is between 10-40%.

47. The method of claim 46, wherein the PEG is 25%.

48. The method of claim 39, wherein the incubation temperature is between 4 - 95 °C.

49. The method of claim 48, wherein the incubation temperature is between 25 °C.

50. The method of claim 39, wherein the incubation time is between 0.25 - 24 hours.

51. The method of claim 50, wherein the incubation time is 4 hours.

52. The method of claim 39, wherein the ligase comprises T4 RNA ligase 1, T4 RNA ligase 2 truncated, T4 RNA ligase 2 K227Q, T4 RNA ligase 2 truncated KQ, and *Mth.* RNA ligase.

53. The method of claim 52, wherein the ligase comprises T4 RNA ligase 2 K227Q.

54. The method of claim 39, wherein the adapter probe comprises DNA bases and is 5 - 50 nt in length.

55. The method of claim 39, wherein the adapter probe comprises a mixture of RNA and DNA bases and is 5 - 50 nt in length.

56. The method of claim 55, wherein the adapter probe comprises RNA bases adjacent to the ligation site.

57. A method for detection of one or more target RNA of interest in a sample comprising:

a) obtaining a sample containing one or more target RNA of interest;

b) adding to the sample of a) a sufficient amount of poly (A) polymerase;

c) performing poly (A) extension by incubating and adding a poly (A) tail to the end of one or more target RNA of interest in the sample ;

d) adding to the sample of b) a sufficient amount of common probes and discrimination probes, and a sufficient amount of a second ligase;

e) performing a second ligation step by incubating the sample of c) for a sufficient amount of time to effect hybridization of probes of d) with ligated target RNA of interest and adapter probes of b) and to allow ligation between the discrimination probes and common probes; and

f) analysis of the products of by one or more size discrimination methods.

FIGURE 1

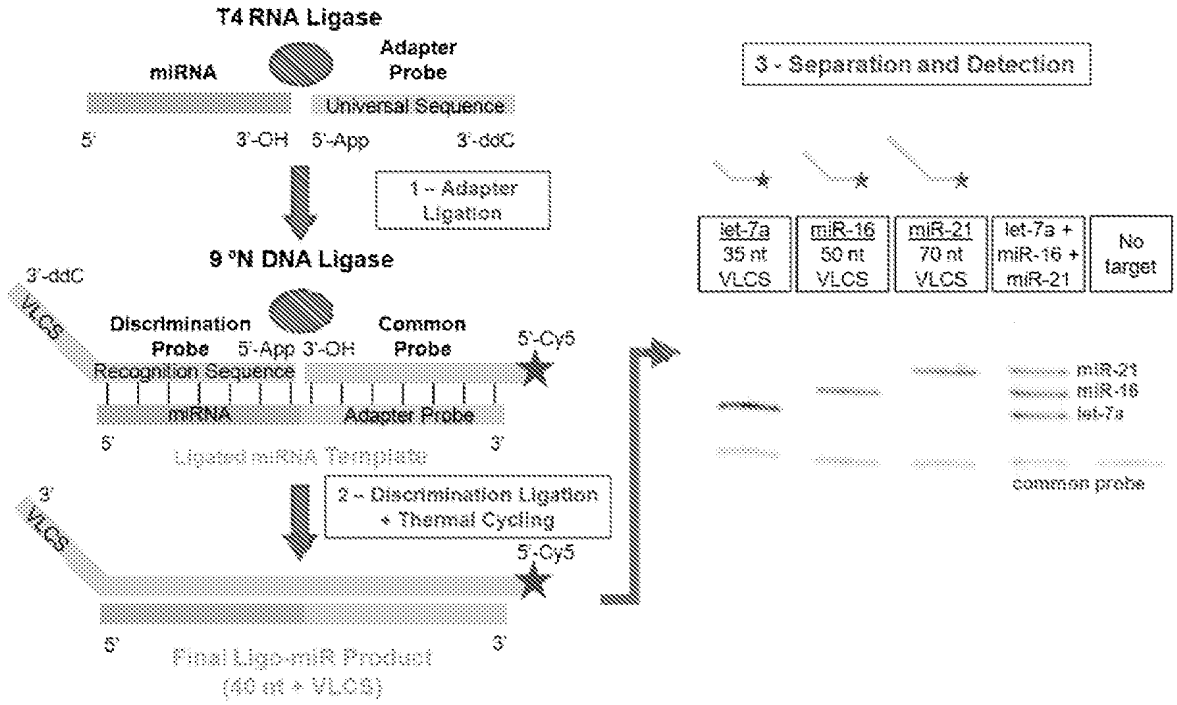

T4 RNA Ligase
miRNA
Adapter Probe
Universal Sequence
5'
3'-OH
5'-App
3'-ddC
1 – Adapter Ligation
9 °N DNA Ligase
3'-ddC
VLCS
Discrimination Probe
5'-App
3'-OH
Common Probe
5'-Cy5
Recognition Sequence
miRNA
Adapter Probe
5'
3'
Ligated miRNA Template
2 – Discrimination Ligation + Thermal Cycling
3'
VLCS
5'-Cy5
5'
3'
Final Ligo-miR Product (40 nt + VLCS)
3 - Separation and Detection
let-7a 35 nt VLCS
miR-16 50 nt VLCS
miR-21 70 nt VLCS
let-7a + miR-16 + miR-21
No target
miR-21
miR-16
let-7a
common probe

FIGURE 2


miR-34a
28 nt
VLCS
let-7a
35 nt
VLCS
miR-29b
41 nt
VLCS
miR-16
50 nt
VLCS
miR-26a
60 nt
VLCS
miR-21
70 nt
VLCS
all
miRNA

FIGURE 3

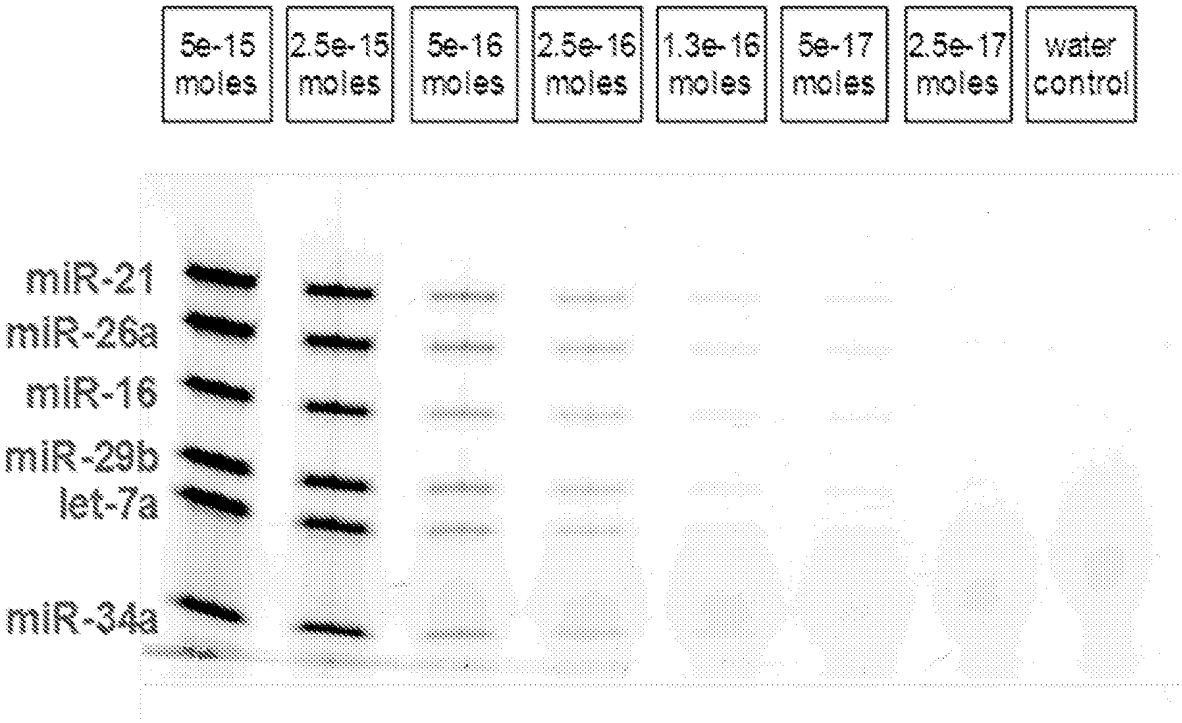

5e-15 moles
2.5e-15 moles
5e-16 moles
2.5e-16 moles
1.3e-16 moles
5e-17 moles
2.5e-17 moles
water control
miR-21
miR-26a
miR-16
miR-29b
let-7a
miR-34a

FIGURE 4

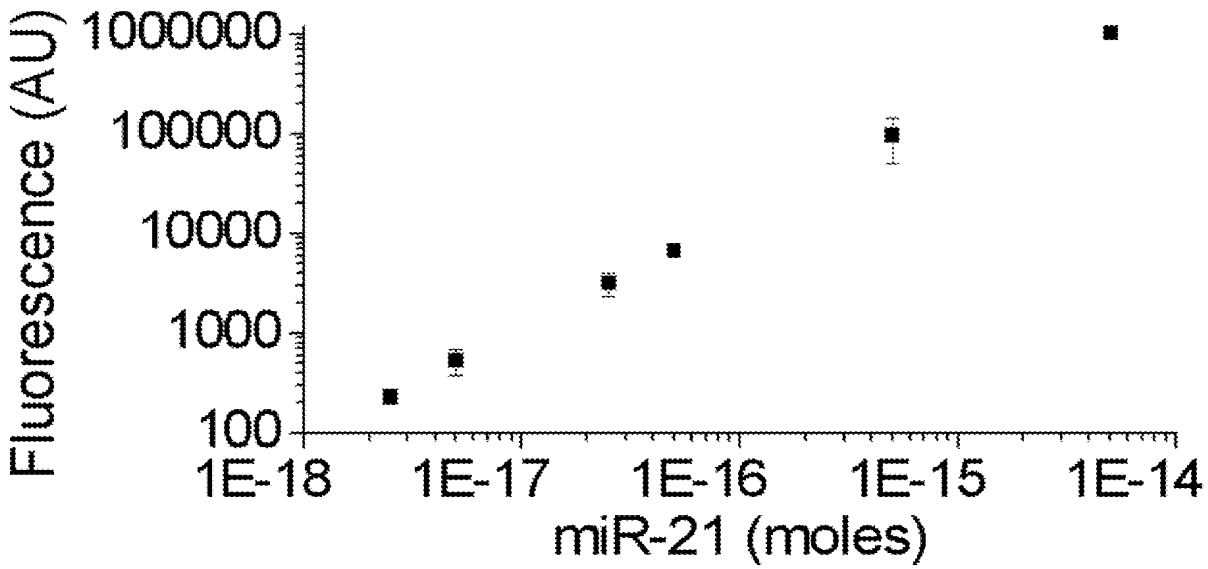

Fluorescence (AU)
1000000
100000
10000
1000
100
1E-18
1E-17
1E-16
1E-15
1E-14
miR-21 (moles)

FIGURE 5

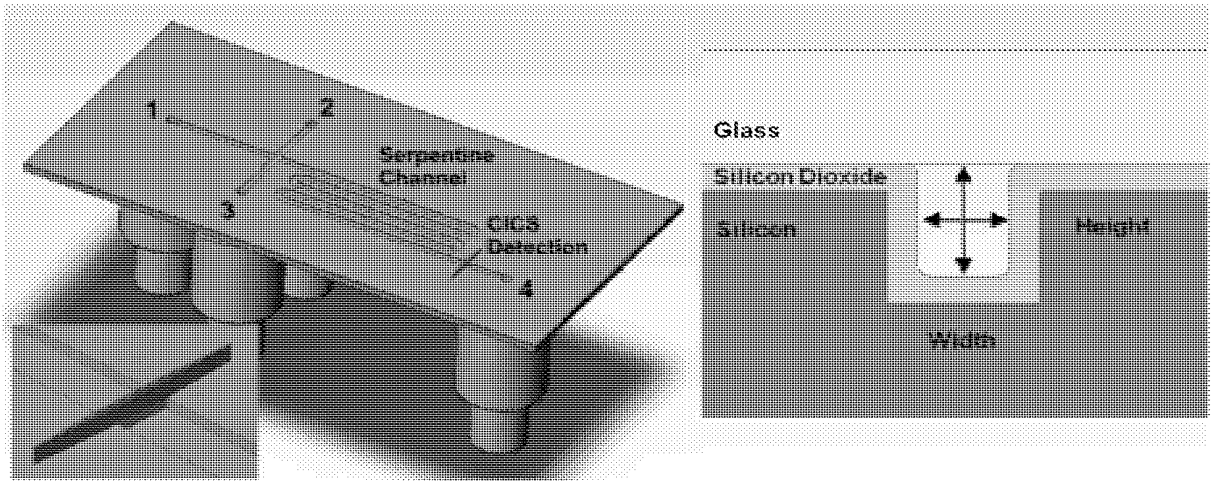

1
2
3
Serpentine Channel
CICS Detection
4
Glass
Silicon Dioxide
Silicon
Height
Width

FIGURE 6

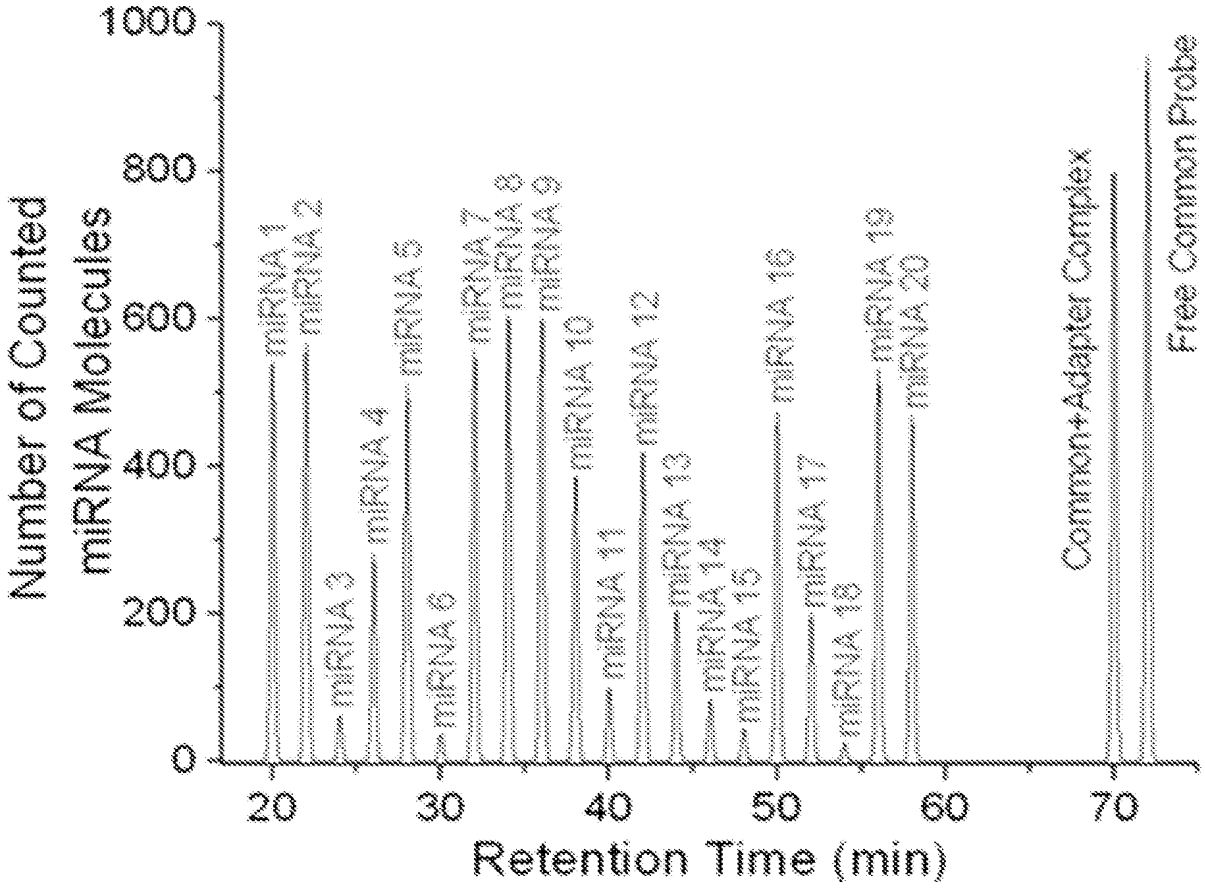

Number of Counted miRNA Molecules
1000
800
600
400
200
0
20
30
40
50
60
70
Retention Time (min)
miRNA 1
miRNA 2
miRNA 3
miRNA 4
miRNA 5
miRNA 6
miRNA 7
miRNA 8
miRNA 9
miRNA 10
miRNA 11
miRNA 12
miRNA 13
miRNA 14
miRNA 15
miRNA 16
miRNA 17
miRNA 18
miRNA 19
miRNA 20
Common+Adapter Complex
Free Common Probe

FIGURE 7

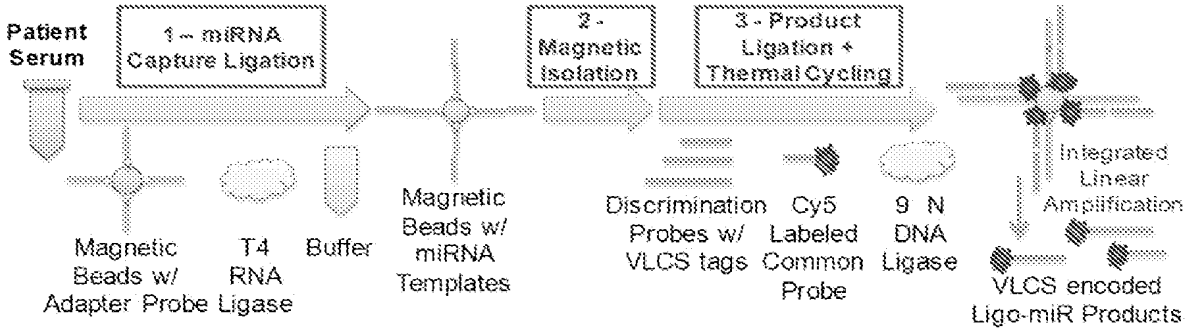
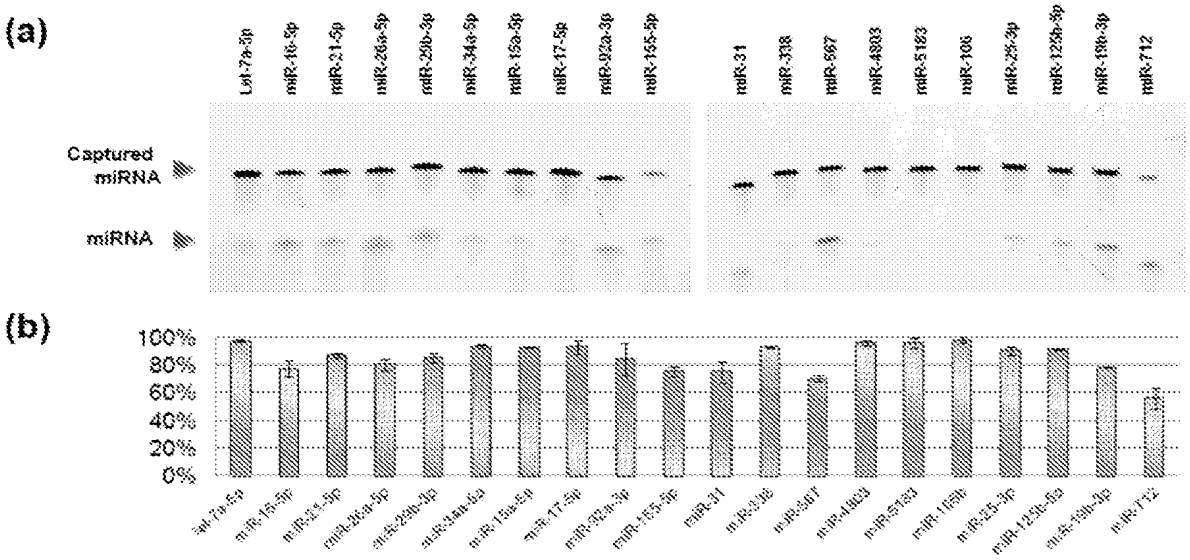

Patient Serum
1 – miRNA Capture Ligation
2 - Magnetic Isolation
3 - Product Ligation + Thermal Cycling
Magnetic Beads w/ Adapter Probe
T4 RNA Ligase
Buffer
Magnetic Beads w/ miRNA Templates
Discrimination Probes w/ VLCS tags
Cy5 Labeled Common Probe
9 N DNA Ligase
Integrated Linear Amplification
VLCS encoded Ligo-miR Products

FIGURE 8



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/025860

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/68(2006.01)i, C12N 15/11(2006.01)i, G01N 33/58(2006.01)i, G01N 33/50(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q 1/68; C12P 19/34; C07H 21/04; C12N 15/11; G01N 33/58; G01N 33/50

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS(KIPO internal) & Keywords:miRNA, assay, adapter, probe, poly-a, discriminant, bead, bias, PCR-free hybridization, ligation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZHANG et al. 'Sensitive detection of microrna with isothermal amplification and a single-quantum-dot-based nanosensor` Analytical Chemistry, Vol. 86, No.1, pp.224-231 (21 November 2011) See abstract; table 1; and scheme 1.	1-4,24-29,32-35 ,39-57
A	US 2009-0075276 A1 (LEE et al.) 19 March 2009 See claim 1 and figures 1-5.	1-4,24-29,32-35 ,39-57
A	US 2006-0078925 A1 (MOURELATOS et al.) 13 April 2006 See the whole documents.	1-4,24-29,32-35 ,39-57
A	US 2011-0104693 A1 (SELIGMANN et al.) 05 May 2011 See the whole documents.	1-4,24-29,32-35 ,39-57
A	EP 1735459 B1 (JACOBSEN et al.) 25 January 2012 See the whole documents.	1-4,24-29,32-35 ,39-57

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 June 2013 (07.06.2013)	**10 June 2013 (10.06.2013)**

Name and mailing address of the ISA/KR
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea
Facsimile No. 82-42-472-7140

Authorized officer
KIM, Seung Beom
Telephone No. 82-42-481-3371

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/025860

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 37-38
because they relate to subject matter not required to be searched by this Authority, namely:
Claims 37-38 pertain to methods for diagnosis of the human body, and thus relate to a subject matter which this International Searching Authority is not required, under Article 17(2)(a)(i) of the PCT and Rule 39.1(iv) of the Regulations under the PCT, to search.

2. ☒ Claims Nos.: 9,12,15,20,22
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
The claims 9,12,15,20,22 do not comply with PCT Article 6 because they are referring to multiple dependent claims which do not comply with PCT Rule 6.4(a).

3. ☒ Claims Nos.: 5-8,10-11,13-14,16-19,21,23,30-31,36
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required addtional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2013/025860

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2009-0075276 A1	19.03.2009	JP 2009-543582 A	10.12.2009
		JP 4909413 B2	20.01.2012
		US 2012-0003638 A1	05.01.2012
		US 2013-0089860 A1	11.04.2013
		US 7947446 B2	24.05.2011
		US 8304527 B2	06.11.2012
		WO 2008-150377 A1	11.12.2008
US 2006-0078925 A1	13.04.2006	None	
US 2011-0104693 A1	05.05.2011	AU 2010315303 A1	24.05.2012
		CA 2778249 A1	12.05.2011
		CN 102712955 A	03.10.2012
		EP 2496715 A1	12.09.2012
		JP 2013-509863 A	21.03.2013
		WO 2011-056863 A1	12.05.2011
EP 1735459 B1	25.01.2012	AT 542918 T	15.02.2012
		CA 2562390 A1	20.10.2005
		CN 1961081 A	09.05.2007
		CN 1961081 B	05.09.2012
		DK 1735459 T3	29.05.2012
		EP 1735459 A2	27.12.2006
		ES 2380923 T3	21.05.2012
		JP 2007-532100 A	15.11.2007
		JP 5192229 B2	08.02.2013
		WO 2005-098029 A2	20.10.2005
		WO 2005-098029 A3	15.12.2005

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 December 2014 (18.12.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/200767 A1

(51) International Patent Classification:
C12P 19/34 (2006.01)

(21) International Application Number:
PCT/US2014/040731

(22) International Filing Date:
3 June 2014 (03.06.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/834,111	12 June 2013 (12.06.2013)	US
61/912,054	5 December 2013 (05.12.2013)	US
61/972,940	31 March 2014 (31.03.2014)	US

(71) Applicant: **THE GENERAL HOSPITAL CORPORATION** [US/US]; 55 Fruit Street, Boston, Massachusetts 02114 (US).

(72) Inventors: **WEISSLEDER, Ralph**; 11 Nichols Lane, West Peabody, Massachusetts 01960 (US). **AGASTI, Sarit S.**; 18 Sutherland Road, Apt. 6, Brighton, Massachusetts 02135 (US). **PETERSON, Vanessa M.**; 179 Morrison Avenue, Somerville, Massachusetts 02144 (US). **ULLAL, Adeeti**; 19384 Chablis Ct., Saratoga, California 95070 (US).

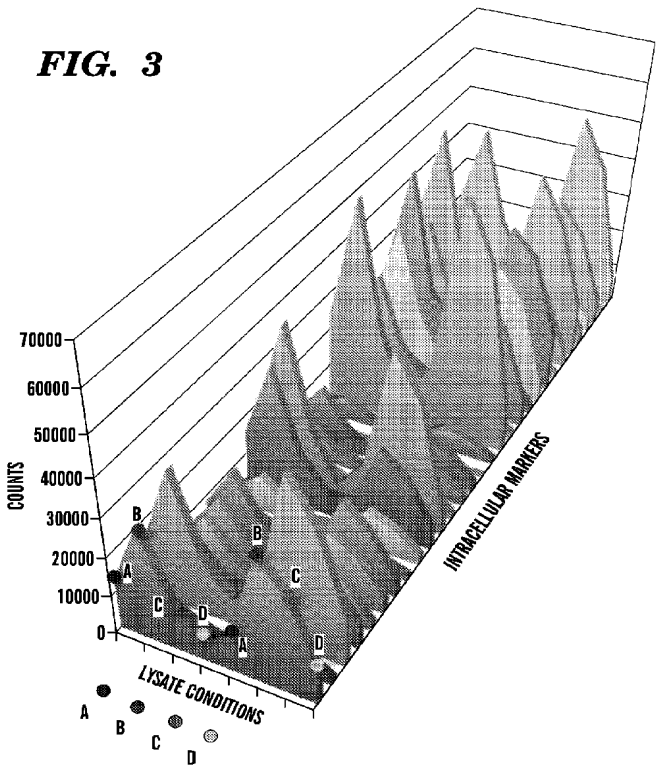
(74) Agents: **RESNICK, David S.** et al.; 100 Summer Street, Boston, Massachusetts 02110 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: METHODS, KITS, AND SYSTEMS FOR MULTIPLEXED DETECTION OF TARGET MOLECULES AND USES THEREOF

FIG. 3



(57) Abstract: Described herein are methods, compositions, kits and systems for multiplexed detection of target molecules from a sample. In some embodiments, the methods, compositions, kits and systems can be used to perform multiplexed protein analysis of a sample (e.g., a sample comprising a small number of cells or a single-cell sample). In some embodiments, the same sample subjected to a multiplexed protein analysis using the methods, compositions, kits and systems described herein can also be subjected to a nucleic acid (e.g., RNAs, microRNAs, and/or DNA) analysis, thereby creating an integrated expression profiling from a limited amount of sample.

WO 2014/200767 A1

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

METHODS, KITS, AND SYSTEMS FOR MULTIPLEXED DETECTION OF TARGET MOLECULES AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit under 35 U.S.C. § 119(e) of the U.S. Provisional Application No. 61/834,111 filed June 12, 2013, U.S. Provisional Application No. 61/912,054 filed December 5, 2013, U.S. Provisional Application No. 61/972,940 filed March 31, 2014, the contents of each of which are incorporated herein by reference in their entireties.

TECHNICAL FIELD

[0002] The present invention relates to methods, kits, and systems for detection of a plurality of target molecules in a sample. The methods, kits, and systems described herein can be used in diagnostic, prognostic, quality control and screening applications.

BACKGROUND

[0003] An increasing number of clinical trials, e.g., cancer trials, require patient samples, e.g., tissue biopsies, to measure individual drug response markers [1]. For example, surgically harvested tissues are often used to collect data at two ends of the cellular spectrum: (i) genomic analyses that reveal driver oncogenes and specific mutations [2] and (ii) protein analyses of selected biomarkers intended to monitor cellular responses [3, 4]. Ideally, clinical samples are collected serially to monitor change in expression levels of key proteins. This raises many challenges, notably risk of morbidity with repeat core biopsies, increased cost, and logistical limitations. Alternative sample collection methods include fine-needle aspirates (FNAs), "liquid biopsies" of circulating tumor cells, or analysis of scant cells present in other easily harvested fluids. However, these samples have much lower cell numbers than biopsies, thereby limiting the number of proteins that can be analyzed.

[0004] After tissues have been sampled, selecting ubiquitous biomarkers can be difficult because of heterogeneity and dynamic network changes. Typically, small-molecule drugs influence more than one target proteins, whereas numerous proteins modulate downstream specific drug actions, trigger alternative molecular pathways, and induce tumor cell death or resistance [5]. The current tools to profile these key proteins in scant clinical samples are limited; standard practice encompasses immunocytology, which often precludes broad protein analysis because of insufficient sample within FNAs or liquid biopsies [6]. Thus, the number of markers is often limited (<10) and requires time-consuming analyses of tissue sections. Proteomic analyses by mass spectrometry remain technically challenging for single cells and phosphoproteomic detection and are costly for

routine clinical purposes [7]. In research settings, multiplexed flow cytometry and mass cytometry have been used to examine an expanded set of markers (10 to 45) using single-cell populations. However, multiplexed flow cytometry often encounters limits in the amount of markers it can measure because of spectral overlap. Mass cytometry vaporizes cells during sample preparation, resulting in sample loss [8]. Accordingly, both of these existing methods do not enable isolating a rare cell of interest or performing concurrent genetic analyses once samples are used for proteomic analyses.

[0005] Hence, there remains a need for compositions and methods for simultaneous detection of a large number of target molecules from a sample.

SUMMARY

[0006] Embodiments of various aspects described herein are, in part, based on the development of a method that not only allows analysis of hundreds of proteins from a limited amount of sample, e.g., minimally invasive fine-needle aspirates (FNAs), which contain much smaller numbers of cells than core biopsies, but also preserves genetic material from the same sample to enable simultaneous measurements of proteins and genetic materials (e.g., DNA, RNA, and microRNAs). In particular, the method relies on DNA-barcoded antibody sensing, where barcodes-single strands of DNA- can be photocleaved and detected using fluorescent complementary probes without any amplification steps, and is referred to as an antibody barcoding with photocleavable DNA (ABCD) platform herein. To demonstrate the capability of the ABCD platform, inventors isolated cancer cells within the FNAs of patients and exposed these cells to a mixture of about 90 DNA-barcoded antibodies, covering the hallmark processes in cancer (for example, apoptosis and DNA damage). The inventors discovered that the single-cell protein analysis of the patients' FNAs showed high intratumor heterogeneity, indicating the ability of the ABCD platform to perform protein profiling on rare single cells, including, but not limited to circulating tumor cells. Further, the inventors discovered that patients who showed identical histopathology yet showed patient heterogeneity in proteomic profiling, indicating the ability of the ABCD platform to identify personalized targets for treatment. By profiling and clustering protein expression in patients' samples, the inventors also showed use of the ABCD platform to monitor and predict treatment response in patients receiving chemotherapy, e.g., kinase inhibitors. The protein analysis determined by the ABCD platform is scalable and can be extended to detect other target molecules, e.g., metabolites and lipids. Accordingly, various aspects described herein provide for methods, systems and kits for detecting and/or quantifying a plurality of target molecules from a sample, as well as their uses thereof in various applications, e.g., diagnosis, prognosis, personalized treatment, and/or treatment monitoring.

[0007] In one aspect, provided herein is a method for detecting a plurality of target molecules in a sample. The method comprises (a) contacting a sample with a composition comprising a plurality of target probes, wherein each target probe in the plurality comprises: (i) a target-binding molecule that specifically binds to a target molecule in the sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable linker between the target-binding molecule and the identification nucleotide sequence; (b) releasing the identification nucleotide sequences from the bound target probes; and (c) detecting signals from the released identification nucleotide sequences, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of target molecules in the sample.

[0008] Stated another way, the method comprises: (a) forming a plurality of complexes in a sample, each complex comprising a target molecule and a target probe bound thereto, wherein the target probe comprises (i) a target-binding molecule that specifically binds to the target molecule present in the sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable linker between the target-binding molecule and the identification nucleotide sequence; (b) releasing the identification nucleotide sequences from the complex; and (c) detecting signals from the released identification nucleotide sequences, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of target molecules in the sample. In some embodiments, the cleavable linker is not pre-hybridized (e.g., by basepairing) to any portion of the identified nucleotide sequences.

[0009] In some embodiments, e.g., cell assay, each complex comprising a target molecule and a target probe bound thereto does not require two or more target probes of different kinds bound to the same target molecule, where each of the target probes binds to a different region of the same target molecule. For example, each complex does not require both a first target probe binding to a first region of a target molecule, and a second target probe binding to a second region of the same target molecule. Stated another way, in some embodiments, a single target probe as described herein binding to a target molecule is sufficient for enablement of the methods described herein. In these embodiments, the method described herein does not require another target probe binding to the same target molecule for attachment to a solid substrate (e.g., a bead).

[0010] In some embodiments, the method can further comprise separating unbound target probes from target probes that are bound to the target molecules in the sample.

[0011] The signals from the released identification nucleotide sequences can be detected by various methods known in the art, including, but not limited to sequencing, quantitative polymerase chain reaction (PCR), multiplexed (PCR), mass cytometry, fluorophore-inactivated multiplexed

immunofluorescence, hybridization-based methods, fluorescence hybridization-based methods, imaging, and any combinations thereof. In some embodiments, the signals from the released identification nucleotide sequences can be determined by electrophoresis-based methods. In some embodiments, the signals from the released identification nucleotide sequences are not determined by electrophoresis-based methods.

[0012] In some embodiments, the signals from the released identification nucleotide sequences can be detected by hybridization-based methods. For example, in some embodiments, the method can further comprise, prior to the detecting in (c), coupling the released identification nucleotide sequences from (b) to a detection composition comprising a plurality of reporter probes. Each reporter probe in the plurality can comprise (i) a first target probe-specific region that is capable of binding a first portion of the identification nucleotide sequence; and (ii) a detectable label that identifies the reporter probe. In these embodiments, signals from the respective detectable labels of the reporter probes that are coupled to the released identification nucleotide sequences can be detected accordingly. Since the signals are distinguishable for each respective reporter probes that are bound to the identification nucleotide sequences, target-binding molecules can be correspondingly identified, thereby detecting a plurality of target molecules in the sample.

[0013] In some embodiments where reporter probes are used in the methods described herein, the detectable label of the reporter probes can comprise one or more labeling molecules that create a unique signal for each reporter probe. For example, a unique signal can be an optical signal. The optical signal can be a light-emitting signal or a series or sequence of light-emitting signals. In some embodiments, labeling molecules for generation of an optical signal can comprise one or a plurality of a fluorochrome moiety, a fluorescent moiety, a dye moiety, a chemiluminescent moiety, or any combinations thereof.

[0014] In some embodiments, the detection composition used in the methods described herein can additionally or alternatively comprise a plurality of capture probes. Each capture probe can comprise (i) a second target probe-specific region that is capable of binding to a second portion of the identification nucleotide sequence; and (ii) an affinity tag. The affinity tag of the capture probe is generally used to permit immobilization of the released identification nucleotide sequences, upon coupling to the detection composition, onto a solid substrate surface. In some embodiments, immobilization of the released identification nucleotide sequences can provide distinguishable spatial signals that identify the capture probes coupled to the released identification nucleotide sequences. Examples of a solid substrate include, but are not limited to, a microfluidic device, a cartridge, a microtiter plate, a tube, and an array.

[0015] In some embodiments, the detection method in (d) does not require amplification of the released identification nucleotide sequences, first target probe-specific region, or the second

target probe-specific region. Amplification-free detection methods can minimize any bias or errors introduced during amplification, e.g., due to varying amplification efficiencies among the nucleotide sequences.

[0016] In some embodiments, identification nucleotide sequences of the target probes described herein can be selected or designed such that they do not cross-react with any nucleic acid sequence in a genome of a subject, whose sample is being evaluated. Thus, the identification nucleotide sequences used to detect target molecules from a subject's sample can be selected or designed based on nucleotide sequences of a species or genus that is different from the subject. By way of example only, in some embodiments, the identification nucleotide sequences for use in an animal's sample (e.g., a mammal such as a human) can be derived from a plant genome. In one embodiment, the identification nucleotide sequences for use in a human's sample can be derived from a potato genome. In some embodiments, the identification nucleotide sequences can have sequences selected from Table 2 (SEQ ID NO: 1 to SEQ ID NO: 110), or a fragment thereof.

[0017] Generally, identification nucleotide sequences of the target probes can have any sequence length and can vary depending on a number of factors, including, but not limited to detection methods, and/or the number of target molecules to be detected. For example, in some embodiments, the length of the identification nucleotide sequences can increase to provide sufficient identification of a large number of target molecules in a sample. In some embodiments where a hybridization-based method is used to detect identification nucleotide sequences, the identification nucleotide sequences can have a length sufficient to provide reliable binding to complementary reporter probes and to generate detectable signals. In some embodiments, the identification nucleotide sequences can have a length of about 30-100 nucleotides. In some embodiments, the identification nucleotide sequences can have a length of about 70 nucleotides.

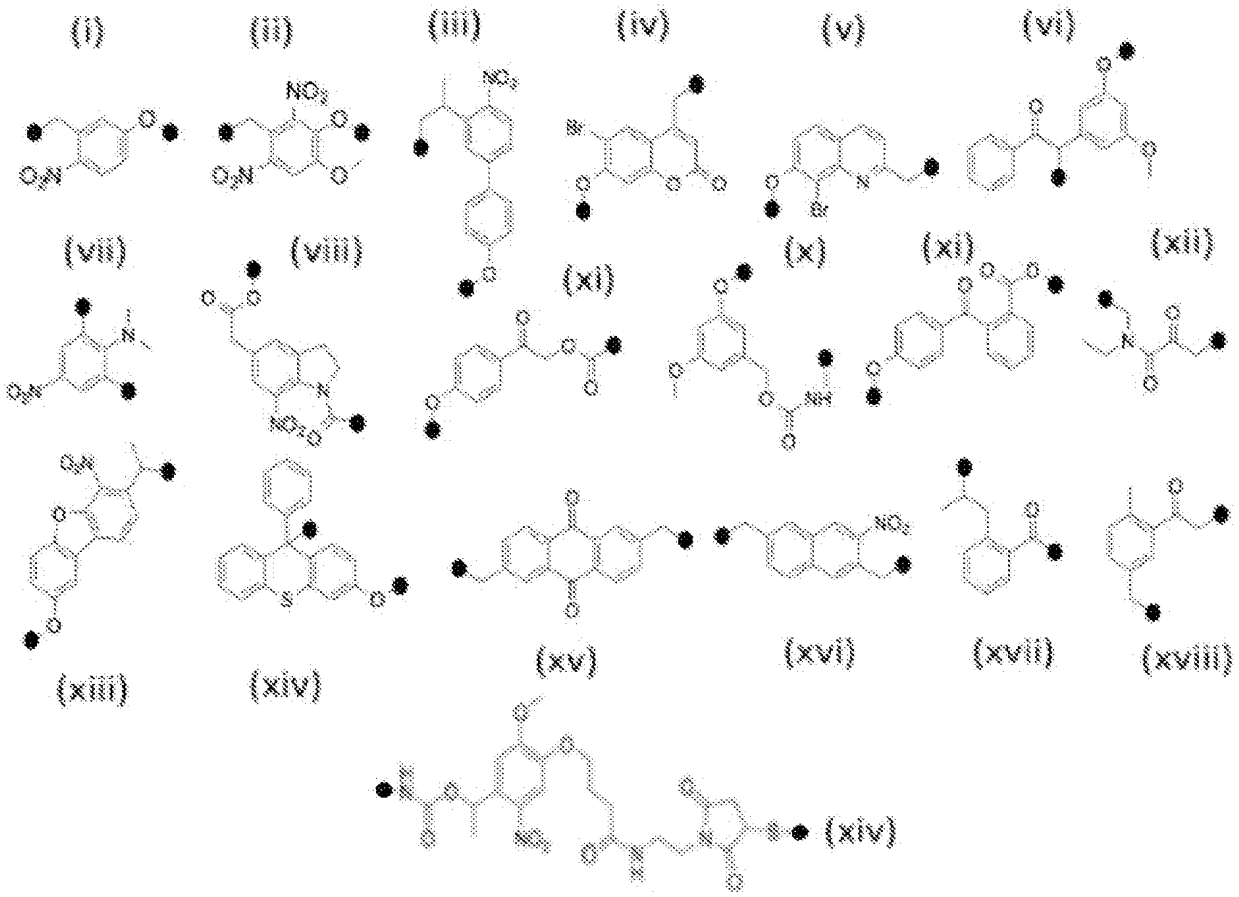
[0018] The cleavable linker coupling a target-binding molecule to an identification nucleotide sequence in a target probe can permit release of the identification nucleotide sequence from the target probe upon binding to a target molecule such that the released identification nucleotide sequence can then be detected. Cleavable linkers are known in the art, of which examples include, but are not limited to the ones that are sensitive to an enzyme, pH, temperature, light, shear stress, sonication, a chemical agent (e.g., dithiothreitol), or any combination thereof. In some embodiments, the cleavable linker can be sensitive to light and enzyme degradation.

[0019] In some embodiments, the cleavable linker does not comprise a polynucleotide sequence (e.g., a single-stranded polynucleotide sequence) that is complementary (for basepairing) to at least a portion of the identification nucleotide sequence. That is, in these embodiments, the identification nucleotide sequence is not released from the complex by detaching from the complementary polynucleotide sequence coupled to a target-binding molecule. Accordingly, in some

embodiments, a target probe comprises (i) a target-binding molecule that specifically binds to the target molecule present in the sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable, non-hybridizable linker between the target-binding molecule and the identification nucleotide sequence.

[0020] In some embodiments, the cleavable, non-hybridizable linkers can be selected from the group consisting of hydrolyzable linkers, redox cleavable linkers, phosphate-based cleavable linkers, acid cleavable linkers, ester-based cleavable linkers, peptide-based cleavable linkers, photocleavable linkers, and any combinations thereof. In some embodiments, the cleavable linker can comprise a disulfide bond, a tetrazine-trans-cyclooctene group, a sulfhydryl group, a nitrobenzyl group, a nitoindoline group, a bromo hydroxycoumarin group, a bromo hydroxyquinoline group, a hydroxyphenacyl group, a dimethozybenzoin group, or any combinations thereof.

[0021] In some embodiments, the cleavable, non-hybridizable linker can comprise a photocleavable linker. Any art-recognized photocleavable linker can be used for the target probes described herein. Exemplary photocleavable linker is selected from the group consisting of molecules (i)-(xiv) and any combinations thereof, wherein the chemical structures of the molecules (i)-(xiv) are shown as follows:



where each of the black dots in each molecule represents a connecting or coupling point that connects, directly or indirectly, to a target-binding molecule described herein or an indentification nucleotide sequence described herein. The connecting point can be a bond, or comprise an atom, a molecule, and/or a linker described herein. In some embodiments, the connecting point is a bond.

[0022] In some embodiments, the photocleavable linker can comprise the molecule (xiv).

[0023] In some embodiments where a photocleavable linker is used, the identification nucleotide sequences can be released from the bound target probes by exposing the bound target probes to a light of a specified wavelength. In some embodiments, ultraviolet light can be used to release identification nucleotide sequences from bound target probes.

[0024] In some embodiments, the method can further comprise, prior to contacting the sample with target probes, separating target cells from interfering cells in the sample. Methods to separate target cells from interfering cells are known in the sample, e.g., based on cell surface proteins that distinguish target cells from interfering cells. By way of example only, target cells or interfering cells can be labeled with ligands that target specific cells of interests (e.g., cell-specific antibodies). In some embodiments where the cell-specific ligands are fluorescently labeled, the labeled cells can then be sorted, e.g., by flow cytometry. Alternatively, if the cell-specific ligands are attached to magnetic particles, the labeled cells with bound magnetic particles can be isolated from the sample by magnetic separation.

[0025] Target cells can be prokaryotic or eukaryotic, including microbes (e.g., bacteria, fungus, virus, and/or pathogens. In some embodiments, the target cells can comprise normal cells, diseased cells, mutant cells, germ cells, somatic cells, and/or rare cells. Example of rare cells include, without limitations, circulating tumor cells, fetal cells, stem cells, immune cells, clonal cells, and any combination thereof. In some embodiments, the target cells can comprise tumor cells. In some embodiments, the tumor cells can be derived from a tissue biopsy, a fine aspirate or a liquid biopsy (e.g., peritoneal, pleural, cerebrospinal fluid, and/or blood), a mucosal swap, a skin biopsy, a stool sample, or any combinations thereof. In some embodiments, whole cells and/or cell lysates can be used in the methods and/or systems described herein to detect a plurality of target molecules in a sample. In some embodiments, the whole cells can be obtained from a fixed cell or tissue sample.

[0026] Typically, signals detected from the identification nucleotide sequences of the target probes corresponding to target molecules can be compared to a control reference to account for any non-specific binding. Accordingly, in some embodiments, the composition added to the sample can further comprise a plurality of control probes. Each control probe in the plurality can comprise: (i) a control-binding molecule that specifically binds to one control molecule in the sample; (ii) an identification control sequence that identifies the control-binding molecule; and (iii) a cleavable

linker between the control-binding molecule and the identification control sequence. The control-binding molecule can bind to a control protein present in a sample. Non-limiting examples of control proteins include housekeeping proteins, control IgG isotypes, mutant non-functional or non-binding proteins, and any combinations thereof.

[0027] Signals from the control probes can then be used to threshold the signals from the target probes. Accordingly, in some embodiments, the method can further comprise thresholding the target signals. In some embodiments, the target signals can be thresholded on the basis of nonspecific binding. In one embodiment, the threshold is generally set to be greater than that of the signals from the non-specific binding. In some embodiments, the threshold can be determined by using standard deviation and measurement error from at least one or more control proteins.

[0028] In some embodiments, the method can further comprise quantifying the signals (e.g., signals that are above the pre-determined threshold) by normalizing the signals associated with the target probes by the signals associated with the control probes. In one embodiment, the signals is quantified and expressed as number of identification nucleotide sequences detected per target-binding agent

[0029] In some embodiments, the method can further comprising extracting a nucleic acid molecule for the same sample for a nucleic acid analysis. Examples of a nucleic acid detection and analysis can include, but are not limited to sequencing, quantitative polymerase chain reaction (PCR), multiplexed PCR, DNA sequencing, RNA sequencing, de novo sequencing, next-generation sequencing such as massively parallel signature sequencing (MPSS), polony sequencing, pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing, ion semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, single molecule real time (SMRT) sequencing, nanopore DNA sequencing, sequencing by hybridization, sequencing with mass spectrometry, microfluidic Sanger sequencing, microscopy-based sequencing techniques, RNA polymerase (RNAP) sequencing, or any combinations thereof.

[0030] While the methods described herein are described in the context where the identification nucleotide sequences are released from bound target probes before detection, in some embodiments, the identification nucleotide sequences do not need to be released from the bound target probes. Accordingly, in some embodiments, the methods described herein can also apply when the identification nucleotide sequences remain bound to target probes during detection.

[0031] Various embodiments of the methods described herein can be carried out in one or more functional modules in a system or a computer system as described herein. Accordingly, another provided herein relates to a system for multiplexed detection of a plurality of target molecules in a sample. For example, the system comprises:

(a) at least one sample processing module comprising instructions for receiving said at least one test sample comprising a sample and a plurality of target probes, wherein each target probe in the plurality comprises:

i. a target-binding molecule that specifically binds to one target molecule in the sample;

ii. an identification nucleotide sequence that identifies the target-binding molecule; and

iii. a cleavable linker between the target-binding molecule and the identification nucleotide sequence; and

wherein the at least one sample processing module further comprises instructions for releasing the identification nucleotide sequences from the target probes that are bound to target molecules in the sample;

(b) a signal detection module comprising instructions for detecting signals from the released identification nucleotide sequences;

(c) at least one data storage module comprising instructions for storing the detected signals from (b) and information associated with identification nucleotide sequences of the target probes;

(d) at least one analysis module comprising instructions for determining the presence of one or more target molecules in the sample based on the detected signals; and

(e) at least one display module for displaying a content based in part on the analysis output from said analysis module, wherein the content comprises a signal indicative of the following: (i) the presence of one or more target molecules in the sample, (ii) the absence of one or more target molecules in the sample, and/or (iii) expression levels of one or more target molecules in the sample.

[0032] In some embodiments, the analysis module can further comprise instructions for (i) identifying the detectable probes of the reporter probes that correspond to the detected signals; (ii) identifying the identification nucleotide sequences of the target probes that correspond to the detectable probes based on the first target probe-specific regions of the reporter probes; and (iii) identifying the target-binding molecules that correspond to the identification nucleotide sequences, thereby determining the presence of one or more target molecules in the sample.

[0033] In some embodiments, the content can be displayed on a computer display, a screen, a monitor, an email, a text message, a website, a physical printout (e.g., paper), or provided as stored information in a data storage device.

[0034] Kits, e.g., for multiplexed detection of a plurality of different target molecules from a sample, are also provided herein. In some embodiments, the kit comprise:

(a) a plurality of target probes, wherein each target probe in the plurality comprises:

i. a target-binding molecule that specifically binds to one target molecule in the sample;

ii. an identification nucleotide sequence that identifies the target-binding molecule; and

iii. a cleavable linker between the target-binding molecule and the identification nucleotide sequence; and

(b) a plurality of reporter probes, wherein each reporter probe comprises:

i. a first target probe-specific region that is capable of binding a first portion of the identification nucleotide sequence; and

ii. a detectable label that identifies the reporter probe.

[0035] In some embodiments, the detectable label of the reporter probes can comprise one or more labeling molecules that create a unique signal for each reporter probe. An exemplary unique signal can be an optical signal. The optical signal can comprise one or a series or a sequence of light-emitting signals. In these embodiments, non-limiting examples of the labeling molecules include fluorochrome moieties, fluorescent moieties, dye moieties, chemiluminescent moieties, and any combinations thereof.

[0036] In some embodiments, the kit can further comprise a plurality of capture probes, wherein each capture probe comprises (i) a second target probe-specific region that is capable of binding a second portion of the identification nucleotide sequence; and (ii) an affinity tag.

[0037] In some embodiments, the kit can further comprise a plurality of control probes, wherein each control probe in the plurality comprises:

(i) a control-binding molecule that specifically binds to one control molecule in the sample;

(ii) an identification control sequence that identifies the control-binding molecule; and

(iii) a cleavable linker between the control-binding molecule and the identification control sequence.

[0038] In some embodiments, the kit can further comprise at least one reagent for use in one or more embodiments of the methods or systems described herein. Reagents that can be provided in the kit can include at least one or more of the following: a hybridization reagent, a purification reagent, an immobilization reagent, an imaging agent, a cell permeabilization agent, a blocking agent, a cleaving agent for the cleavable linker, primers for nucleic acid detection, nucleic acid polymerase, and any combinations thereof.

[0039] In some embodiments, the kit can further include a device for use in one or more embodiments of the methods and/or systems described herein. In some embodiments, the device can comprise a surface for immobilization of the capture probes upon coupling to the identification nucleotide sequences. In some embodiments, the device can comprise a microfluidic device for separating target cells from interfering cells as described herein.

[0040] The methods, systems and kits described herein can be used to detect any target molecules present in a sample provided that appropriate target-binding agents are used in the target probes employed in the methods described herein. Exemplary target molecules which can be

detected by the methods, systems and kits described herein include, but are not limited to proteins, peptides, metabolites, lipids, carbohydrates, toxins, growth factors, hormones, cytokines, cells (e.g., eukaryotic cells, prokaryotic cells, and microbes), and any combinations thereof. In some embodiments, the target molecules to be detected can be extracellular or secreted molecules. In some embodiments, the target molecules to be detected can be intracellular, e.g., cytoplasmic molecules or nuclear molecules.

[0041] To detect intracellular molecules (e.g., intracellular proteins), the target cells in the sample can be permeabilized or lysed such that target probes can contact the target intracellular molecules for further processing and analysis.

[0042] In some embodiments, the methods, systems and kits described herein can enable measurements of at least two target molecules of different types. For example, the methods, systems, and kits described herein can be used to measure, for example, nucleic acid molecules and proteins, or proteins and metabolites, or proteins and lipids. The measurements of at least two target molecules of different types can be performed simultaneously or sequentially. In another embodiment, by releasing identification nucleotide sequences from bound target molecules (e.g., proteins), genetic material and the identification nucleotide sequences can be concurrently extracted from a single sample, enabling analyses of protein-DNA-RNA interrelationships.

[0043] By way of example only, the methods, systems and kits described herein applied to a sample can preserve genetic materials in a sample while detecting other non-genetic target materials in the same sample. Accordingly, in some embodiments, the methods, systems and/or kits described herein for detection of non-genetic target molecules (e.g., but not limited to proteins) can be used in combination with a nuclei acid analysis for genetic materials, for example, to study the non-genetic target molecules (e.g., but not limited to proteins) that interact with genetic materials or genetic regulatory elements. In these embodiments, the methods and systems described herein for detecting a plurality of target molecules in a sample as described herein can further comprise extracting a nucleic acid molecule from the same sample in which target molecules are to be detected. In some embodiments, the methods and systems described herein can further comprise subjecting the extracted nucleic acid molecule to a nucleic acid analysis. Various methods can be used for nucleic acid analysis, including, but not limited to sequencing, next generation sequencing, quantitative polymerase chain reaction (PCR), multiplexed PCR, DNA sequencing, RNA sequencing, de novo sequencing, next-generation sequencing such as massively parallel signature sequencing (MPSS), polony sequencing, pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing, ion semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, single molecule real time (SMRT) sequencing, nanopore DNA sequencing, sequencing by hybridization, sequencing with mass spectrometry, microfluidic Sanger sequencing, microscopy-based sequencing

techniques, RNA polymerase (RNAP) sequencing, fluorescence hybridization-based technology (e.g., but not limited to nanoString nCounter® hybridization technology), any art-recognized nucleic acid detection methods, or any combinations thereof.

[0044] In some embodiments, after a sample and/or non-genetic target molecules have been labeled with a plurality of target probes described herein, the identification nucleotide sequences of the target probes can be released from the bound non-genetic target molecules simultaneously with extraction of nucleic acid molecules (cells' genetic materials) from the same labeled sample. In these embodiments, both the nucleic acid molecules of interest and the identification nucleotide sequences can be detected simultaneously in a single sample mixture. In one embodiment, both the nucleic acid molecules of interest and the identification nucleotide sequences can be detected simultaneously in a single sample mixture using nanoString nCounter® analysis system, for example, as described in U.S. Pat. No. 8,415,102, the content of which is incorporated herein by reference. Other art-recognized methods for nucleic acid analyses as described herein can also be used for simultaneous detection of both nucleic acid molecules of interest (cells' genetic materials) and released identification nucleotide sequences from bound non-genetic target molecules.

[0045] In alternative embodiments, nucleic acid molecules can be extracted from a first portion of a sample, while non-genetic target molecules can be independently derived or obtained from a second portion of the same sample. In these embodiments, the nucleic acid molecules of interest and the non-genetic target molecules can be detected separately to determine expression levels of the nucleic acid molecules (cells' genetic materials) of interest and non-genetic target molecules (e.g., but not limited to proteins) in the same sample. The nucleic acid molecules of interests can be subjected to any art-recognized nucleic acid analysis, while the non-genetic target molecules can be detected through detecting and identifying the corresponding identification nucleotide sequences released from the target probes using the methods, systems and/or kits described herein.

[0046] In some embodiments, the methods, systems and kits described herein can be adapted to measure proteins and nucleic acid molecules in the same sample. For example, the proteins can be detected by one or more embodiments of the target probes described herein, while the nucleic acid molecules can be detected by any methods known in the art, thereby creating an integrated expression profiling for the sample, which can provide information on interaction between the proteins and the nucleic acid molecules, e.g., genetic regulatory elements such as microRNAs.

[0047] The methods, systems and kits described herein can be used in any applications where detection of a plurality of target molecules in a sample is desirable. For example, a sample can be a biological sample, or a sample from an environmental source (e.g., water, soil, food products, and/or ponds).

[0048] The inventors have demonstrated that, in one embodiment, an antibody barcoding with photocleavable DNA (ABCD) platform described herein can enable analysis of hundreds of proteins from a single cell or a limited number of cells, e.g., from minimally invasive fine-needle aspirates (FNAs). Accordingly, samples amenable to the methods described herein can comprise less than 500 cells or fewer. In some embodiments, the sample can be a single-cell sample. In some embodiments, the sample can comprise cells isolated from a fine-needle aspirate.

[0049] Where the sample is a biological sample, in some embodiments, the methods, systems and kits described herein can be used in personalized treatment. For example, a biological sample can be collected from an individual subject who is in need of a treatment for a condition. Using the methods, systems and/or kits described herein, an expression profile of target molecules associated with the subject's condition can be generated to identify one or more therapeutic targets for the subject, thereby identifying a treatment regimen for the subject.

[0050] In some embodiments, the methods, systems and kits described herein can be used in monitoring response of a subject to a treatment for his/her condition. For example, biological sample(s) can be collected from the subject prior to and/or over the course of the treatment. Using the methods, systems and/or kits described herein, expression profiles of target molecules associated with the subject's condition before and/or over the course of the treatment can be generated for comparison to determine any changes in expression levels of the target molecules in the subject, thereby monitoring the treatment response in the subject.

[0051] In some embodiments, the methods, systems and kits described herein can be used in diagnosing a condition in a subject. For example, a biological sample can be collected from a subject who is at risk for a condition. Using the methods, systems and/or kits described herein, an expression profile of target molecules associated with the condition to be diagnosed can be generated for comparison with one or more reference expression profiles (e.g., corresponding to a normal healthy subject and/or a subject having the condition to be diagnosed), thereby determining whether the subject is at risk for the condition.

BRIEF DESCRIPTION OF THE DRAWINGS

[0052] **Figs. 1A-1C** shows an exemplary scheme of a multiplexed protein analysis in single cells in accordance with one or more embodiments of the methods described herein. (**Fig. 1A**) Cells were harvested from cancer patients by FNAs. In this case, a heterogeneous population of EpCAM-positive cancer cells (green) is displayed alongside mesothelial cells (red) with nuclei shown in blue (Hoechst) from an abdominal cancer FNAs. Cancer cells were enriched and isolated via magnetic separation in polydimethylsiloxane (PDMS) microfluidic devices with herringbone channels using both positive (for example, EpCAM+/CK+) and negative (for example, CD45−) selection modes.

(**Fig. 1B**) Cells of interest were incubated with a cocktail of DNA-conjugated antibodies containing a photocleavable linker (**Fig. 2A**) to allow DNA release after exposure to ultraviolet light. (**Fig. 1C**) DNA-antibody conjugates released from lysed cells (**Fig. 3**) were isolated using size separation and IgG pull-down. Released "alien" DNA barcodes were processed with a fluorescent DNA barcoding platform (NanoString). Fluorescent barcodes were hybridized and imaged using a CCD camera. The quantified barcodes were translated to protein expression levels by normalizing to DNA per antibody and housekeeping proteins and subtracting nonspecific binding from control IgGs. A representative profile of SKOV3 ovarian cancer cell lines shows high CD44 and high Her2 expression, characteristic of this cell line.

[0053] Figs. 2A-2B shows an exemplary scheme of DNA–antibody conjugation. (**Fig. 2A**) Various linker strategies were investigated to conjugate DNA to antibodies. In some embodiments, the photocleavable linker (PCL) was selected owing to its better cleavage efficiency compared with DTT, tetrazine-trans-cyclooctene (via click chemistry, linker 1), and Traut's reagent (linker 2). Linker cleavage was tested by measuring released DNA via the NanoString platform. Data are averages of two independent trials. $^{**}P < 0.01$, paired t-test. (**Fig. 2B**) Linking DNA to an antibody via the PCL. The linker was first reacted with the amine ($-NH_2$) groups on the antibody. After excess small molecule was removed, thiolated DNA was added at 10-fold excess to the antibody-linker mix. The final antibody-DNA chimera was purified via both size separation and IgG-specific pulldown. DNA could subsequently be released from the antibody by photocleavage at a specific wavelength (365 nm).

[0054] Fig. 3 shows experimental data directed to optimization of lysis and blocking methods. (Methods A to D) Four different lysis and blocking methods were used to recover DNA from labeled cells. Lysate conditions included: (Method A) Proteinase K + PKD lysis buffer; (Method B) Proteinase K + ATL lysis buffer; (Method C) ATL lysis buffer alone; and (Method D) UV cleavage alone (no cell lysis). The lysate conditions were tested in duplicate (x-axis) measuring DNA signal (y-axis) and different intracellular proteins (z-axis). The best reaction condition was method B (Proteinase K + ATL lysis buffer), with a 20% increase in signal over methods (Methods A and C).

[0055] Fig. 4 is a graph showing the readouts of DNA per antibody for each target molecule. The number of alien DNA fragments per antibody was measured by Nanostring method (shown in graph) and independently confirmed by ssDNA quantification and Qubit protein measurement. Data are displayed from triplicate measurements ± SEM.

[0056] Fig. 5 shows a multiplexed protein profiling of a human breast cancer cell line. Representative example of 88 different antibodies spanning cancer-relevant pathways (color-coded) profiled in triplicate (mean ± SEM) on the MDAMB-231 triple-negative breast cancer cell line.

DNA counts were converted to protein binding by normalizing to the amount of DNA per antibody. Nonspecific binding from expression of six control IgGs was subtracted, and expression was normalized by housekeeping proteins Cox IV, histone H3, tubulin, actin, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (far right). AU, arbitrary units; EMT, epithelial-to-mesenchymal transition.

[0057] **Figs. 6A-6B** are graphs showing effect of permeabilization schemes on antibody labeling. (**Fig. 6A**) Methanol and saponin permeabilization were similar for both intracellular and nuclear proteins. Representative examples are graphed, such as phospho-histone H3 (pH3), epithelial cell adhesion molecule (EpCAM), phosphorylated Src (pSRC), and phosphorylated glycogen synthase kinase 3β (pGSK3b). (**Fig. 6B**) Nonspecific binding was much higher with methanol permeabilization.

[0058] **Figs. 7A-7B** are experimental data showing comparison of unmodified antibodies to DNA-antibody conjugates. (**Fig. 7A**) Antibody–DNA conjugates show good correlation against unconjugated, native antibodies, as determined by flow cytometry. Experiments were performed on multiple cell lines. A representative example is shown with a head-to-head comparison of multiple antibodies on the human SKOV3 cell line ($R^2 = 0.92$). (**Fig. 7B**) Protein expression detected in different cell lysates showed similar patterns of expression whether detected by unmodified antibodies or DNA conjugates. This held true for p53 and phospho-S6RP (immunoblotting), and Ki67 (dot blot).

[0059] **Figs. 8A-8C** show validation data of DNA–antibody conjugates. (**Fig. 8A**) Concordance between two different antibody clones of EpCAM (MOC-31 and 158206) when conjugated to separate DNA barcodes. The antibodies were assayed across multiple patient samples (n=22). (**Fig. 8B**) Antibody expression was measured when cell lines were stained with a single antibody as compared to a cocktail (80+ antibodies). Data were collected from 5 antibodies (CD44, EGFR, 53BP1, p-S6RP, rabbit IgG) on 3 different human cell lines (MDA-MB-231, MDA-MB-436, and A431). The experiment was repeated in duplicate and each data point corresponds to one marker on a given cell line. Expression measurements were calculated by normalized DNA counts for the same number of cells. (**Fig. 8C**) Changes in marker expression before and after treatment were assayed and quantified using both the method in the Examples (ABCD platform) and an independent immunofluorescence screen (standard error is shown from biological triplicate).

[0060] **Fig. 9** is a set of graphs showing that protein marker expression correlates with flow cytometry. Multiple markers (CD44, Her2, EGFR, CA19-9, Keratin 7, and Muc 1) were screened across multiple cell lines (SK-OV-3, ES-2, OVCA429, UCI-107, UCI-101, TOV-112D, TOV-21G, and A2780). Each data point represents expression derived from NanoString DNA counts or flow cytometry for a particular cell line. Expression values were normalized by housekeeping proteins

GAPDH, tubulin, and actin. Cell lines with measurements below that of the negative control (IgG antibodies) either on flow cytometry or Nanostring were excluded. These measurements were compared to independently performed flow cytometry measurements, which were calculated from the mean fluorescence intensity (signal/background), where the background was the secondary antibody without the primary antibody. The inset shows the log-log plot of the data.

[0061] **Figs. 10A-10C** are experimental data on detection sensitivity using a human epidermal cancer cell line. (**Fig. 10A**) A bulk sample of 500,000cells from the epidermoid carcinoma cell line A431 was lysed and processed as shown in **Figs. 1A-1C**. Dilutions corresponding to 5, 15, and 50 cells were then compared to the bulk measurement. (**Fig. 10B**) Correlation values for single A431 cells selected by micromanipulation are compared to the bulk measurements (500,000 cells). (**Fig. 10C**) Protein expression profiles (log_2 expression values) of four single cells compared with the bulk sample. Correlations were highly significant when comparing all single cells to bulk measurements ($P < .0001$, paired t test; GraphPad Prism 6.0).

[0062] **Figs. 11A-11B** are experimental data showing single-cell variability in treated and untreated A431 cells. (**Fig. 11A**) Single cell measurements in human A431 cell lines that were either treated or untreated with the EGFR inhibitor gefitinib were clustered based on a correlation metric (MatLab). (**Fig. 11B**) Pairwise t-tests for four markers (FDR = 0.1, ***$P < 0.001$, GraphPad Prism), are shown. Markers that were most significant are shown, as well as phosphor-EGFR, which is the primary target of gefitinib inhibition. The distribution between signals from untreated cells (blue) and treated cells (yellow) are shown. Each point represents expression levels calculated from a single cell, and the mean and standard deviation are shown in the box plot.

[0063] **Figs. 12A-12B** show experimental data based on a single-cell protein analysis in a patient sample. An FNA was obtained from a patient with biopsy-proven lung adenocarcinoma. (**Fig. 12A**) Eleven harvested cells were analyzed individually, and protein expression levels in each cell (y axis) were correlated with expression levels from the bulk tumor sample (x axis). Each data point represents the expression for a given marker (n = 85markers, 3 below detection threshold). (**Fig. 12B**) Spearman R correlation coefficient values for each of the single cells in (**Fig. 12A**) relative to each other and to the bulk measurement.

[0064] **Fig. 13** shows interpatient heterogeneity in lung cancer. FNAs were obtained from six patients with biopsy-proven lung adenocarcinoma, and bulk samples (~100 cells each) were processed as shown in **Figs. 1A-1C** with 88 barcoded antibodies. Expression data were log_2-normalized by row to show differences between each patient. Expression profiles were heterogeneous despite the identical histological type: Upon genetic analysis, patients 1 and 2 had EGFR exon 19 amplification and T790M mutations, patient 3 had an exon 20 EGFR mutation,

patient 4 had an EGFR L858R mutation and an additional BRAF mutation, patient 5 had a KRAS mutation, and patient 6 had an EML4-ALK translocation.

[0065] **Figs. 14A-14B** show experimental data on effect of different therapies on protein expression profiles in MDA-MB-436 triple-negative breast cancer cell line. (**Fig. 14A**) MDA-MB-436 cells were treated with different agents, and marker proteins were measured. Unsupervised hierarchical clustering based on Euclidean distance grouped drug treatments by their mechanisms of action (molecularly targeted versus DNA damaging) and primary targets [EGFR for gefitinib/cetuximab and mammalian target of rapamycin (mTOR)/PI3K for PKI-587]. Data show the $\log_2$ fold change of marker expression in treated compared to untreated cells for n = 84 markers. All experiments were performed in triplicate. (**Fig. 14B**) Correlating drug sensitivity of four different cell lines with proteomic profile changes after treatment with cisplatin and olaparib. IC50 values (black bars) were calculated on the basis of viability curves (**Fig. 15A**). The cell profile change after treatment is represented by the number of significant markers (gray bars) that were identified by a pairwise t test of treated versus untreated samples (FDR = 0.1).

[0066] **Figs. 15A-15E** are graphs showing protein marker changes correlate with drug sensitivity. Human ovarian carcinoma (A2780, OVCAR429) and breast cancer (MDA-MB-436, MDA-MB-231) cell lines react differently to chemotherapy. Those with increased sensitivity to a drug are expected to show a greater degree of change in their cell profiles. (**Fig. 15A**) Sensitivity was determined by IC50 values calculated from MTS viability curves in biological triplicate for each cell line as shown. Exact values and the fit of the viability curves were determined by GraphPad Prism 5.0 (dose-response curve). (**Fig. 15B**) Data of a control study where cell lines were treated with cetuximab, which resulted in drug inhibition. (**Figs. 15C-15E**) Changes across a selected panel of several DNA damage markers (pH2A.X, Ku80, pChk2, pChhk1), apoptosis markers (cleaved PARP, cleaved caspase 7), and other mechanisms commonly associated with platinum treatment (pERK, Bim). Data are means ± SEM, performed in triplicate.

[0067] **Figs. 16A-16D** show that taxol treatment and dose response screens in human HT1080 cells in vitro. (**Fig. 16A**) Select marker changes from dose response taxol treatment are displayed with DNA barcoding profiles with standard error from biological duplicate. (**Fig. 16B**) EC50 values from the dose response curves are displayed along with R^2 values. (**Fig. 16C**) Markers that significantly differed from untreated (pairwise multiple t-test, FDR = 0.2) were shown to have a dose-dependent response to taxol treatment. (**Fig. 16D**) The markers that significantly different between untreated and treated conditions are shown in a venn diagram. CDCP1 was significantly different at all doses.

[0068] **Figs. 17A-17B** show expression profilings of various cancer patients for monitoring and predicting treatment response in patients receiving PI3K inhibitors. (**Fig. 17A**) Profiles of five

drug-naïve cancer patients are shown with clustering based on correlation metrics with weighted linkage. The dotted box shows the cluster including the marker that best separated responders and nonresponders (H3K79me2). Other markers in the cluster include pS6RP (a downstream target of PI3K), pH2A.X (DNA damage marker), PARP (DNA repair protein), and 4EBP1 (protein translation). (**Fig. 17B**) Four patients with biopsy-proven adenocarcinoma were treated with PI3Ki, and primary cancers were biopsied before and after treatment. The heat map is a pre-post treatment difference map showing log2 fold changes in protein expression (normalized by row to highlight differences between patients). Patient segregation is by correlation distance metric (weighted linkage). The patient in the third column received a higher dose of the PI3Ki (400 mg, twice daily) than the patient in the fourth column (150 mg, twice daily).

[0069] **Figs. 18A-18B** are block diagrams showing exemplary systems for use in the methods described herein, e.g., for multiplexed detection of target molecules in a sample.

DETAILED DESCRIPTION

[0070] Immunohistochemistry-based clinical diagnoses generally require invasive core biopsies and use a limited number of protein stains to identify and classify cancers. Fine-needle aspirates (FNAs) employ thin needles to obtain cells from tumor masses and the procedure is thus minimally invasive. While FNAs can give tremendous insight into malignancy, the number of cells in the FNAs is so small (compared to core biopsy) that current technologies for protein analysis, such as immunohistochemistry, are insufficient. Embodiments of various aspects described herein are, in part, based on the development of a scalable method that not only allows analysis of a plurality of proteins from a limited amount of sample, e.g., FNAs, but also preserves genetic material from the same sample to enable simultaneous measurements of proteins and genetic materials (e.g., DNA, RNA, epigenetic and microRNAs). In one embodiment, the method relies on DNA-barcoded antibody sensing, where barcodes-single strands of DNA- can be photocleaved and detected using fluorescent complementary probes without any amplification steps, and is referred to as an antibody barcoding with photocleavable DNA (ABCD) platform herein. Unlike the protein detection method described in U.S. Pat. App. Pub. No. US 2011/0086774, the DNA barcode and the antibody that the inventors developed is coupled together through a cleavable, non-hybridizable linker, not a hybridizable linker that is reversibly hybridized (e.g., by basepairing) to a portion of the DNA barcode. In addition, detection of a target protein is based on binding of a single DNA-barcoded antibody to the target protein, which is different from the protein detection method described in U.S. Pat. App. Pub. No. US 2011/0086774, where two antibodies (one for immobilization to a solid substrate, e.g., a bead, and another for detection purpose) are required for binding to different regions of the target protein.

[0071] To demonstrate the capability of the ABCD platform, inventors isolated cancer cells within the FNAs of patients and exposed these cells to a mixture of about 90 DNA-barcoded antibodies, covering the hallmark processes in cancer (for example, apoptosis and DNA damage). The inventors discovered that the single-cell protein analysis of the patients' FNAs showed high intratumor heterogeneity, indicating the ability of the ABCD platform to perform protein profiling on rare single cells, including, but not limited to circulating tumor cells. Further, the inventors discovered that patients who showed identical histopathology yet showed patient heterogeneity in proteomic profiling, indicating the ability of the ABCD platform to identify personalized targets for treatment. By profiling and clustering protein expression in patients' samples, the inventors also showed use of the ABCD platform to monitor and predict treatment response in patients receiving chemotherapy, e.g., kinase inhibitors. The ABCD platform for generating a protein profiling is scalable and can be extended to detect other target molecules, e.g., metabolites and lipids. Not only can the ABCD platform measure protein quantitatively, but the ABCD platform can also enable simultaneous measurements of at least 90 different proteins or more (e.g., about 100-200 different proteins) in a single sample. Further, because of the high sensitivity of the ABCD platform, the ABCD platform can enable detection of rare proteins, e.g., proteins that are not generally highly-expressed, or not easily accessible or extracted, such as intracellular proteins. Accordingly, various aspects described herein provide for methods, systems and kits for detecting and/or quantifying a plurality of target molecules from a sample, as well as their uses thereof in various applications, e.g., diagnosis, prognosis, personalized treatment, and/or treatment monitoring.

Methods for detecting or quantifying a plurality of target molecules in a sample

[0072] In one aspect, provided herein is a method for detecting a plurality of target molecules in a sample. The method comprises (a) contacting a sample with a composition comprising a plurality of target probes, wherein each target probe in the plurality comprises: (i) a target-binding molecule that specifically binds to a target molecule or a distinct target molecule in the sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable linker between the target-binding molecule and the identification nucleotide sequence; (b) releasing the identification nucleotide sequences from the bound target probes; and (c) detecting signals from the released identification nucleotide sequences, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of target molecules in the sample.

[0073] In some embodiments where each target probe in the plurality binds to a distinct target molecule, no two target probes in the plurality binds to different regions of the same target molecule.

[0074] Stated another way, the method comprises: (a) forming a plurality of complexes in a sample, each complex comprising a target molecule and a target probe bound thereto, wherein the target probe comprises (i) a target-binding molecule that specifically binds to the target molecule present in the sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable linker between the target-binding molecule and the identification nucleotide sequence; (b) releasing the identification nucleotide sequences from the complex; and (c) detecting signals from the released identification nucleotide sequences, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of target molecules in the sample. In some embodiments, the cleavable linker is not pre-hybridized (e.g., by basepairing) to any portion of the identified nucleotide sequences.

[0075] In some embodiments, e.g., cell assay, each complex comprising a target molecule and a target probe bound thereto does not require two or more target probes of different kinds bound to the same target molecule, where each of the target probes binds to a different region of the same target molecule. For example, unlike the protein detection method described in the U.S. Pat. App. No. US 2011/0086774, each complex described herein does not require both a first target probe binding to a first region of a target molecule, and a second target probe binding to a second region of the same target molecule. Stated another way, in some embodiments, a single target probe as described herein binding to a target molecule is sufficient for enablement of the methods described herein. In these embodiments, the method described herein does not require another target probe binding to the same target molecule for attachment to a solid substrate (e.g., a bead), e.g., as described in the U.S. Pat. App. No. US 2011/0086774.

[0076] In various embodiments of different aspects described herein, the cleavable linker does not comprise a polynucleotide sequence (e.g., a single-stranded polynucleotide sequence) that is complementary (for basepairing) to at least a portion of the identification nucleotide sequence. That is, in these embodiments, the identification nucleotide sequence is not released from the complex by detaching from the complementary polynucleotide sequence coupled to a target-binding molecule. Accordingly, in some embodiments, a target probe comprises (i) a target-binding molecule that specifically binds to the target molecule present in the sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable, non-hybridizable linker between the target-binding molecule and the identification nucleotide sequence.

[0077] "Target probes" is described in detail in the following "*Target Probes*" section.

[0078] In some embodiments, the method can further comprise separating unbound target probes from target probes that are bound to the target molecules in the sample.

[0079] As used herein, the term "bound target probes" refers to target probes binding to target molecules in a sample.

[0080] In some embodiments, the method can further comprise, prior to contacting the sample with target probes, separating target cells from interfering cells in the sample. Methods to separate target cells from interfering cells are known in the sample, e.g., based on cell surface proteins that distinguish target cells from interfering cells. By way of example only, target cells or interfering cells can be labeled with ligands that target specific cells of interests (e.g., cell-specific antibodies). In some embodiments where the cell-specific ligands are fluorescently labeled, the labeled cells can then be sorted, e.g., by flow cytometry. Alternatively, if the cell-specific ligands are attached to magnetic particles, the labeled cells with bound magnetic particles can be isolated from the sample by magnetic separation. In some embodiments, the cell sorting or selection can be performed in a microfluidic device. In some embodiments, methods for isolating target cells or interfering cells from a sample as described in the International Pat. App. No. WO 2013/078332, the content of which are incorporated herein by reference, can be used in combination with the methods described herein.

[0081] Target cells can be prokaryotic or eukaryotic (including microbes such as bacteria, fungi, virus and/or pathogens). In some embodiments, the target cells can comprise normal cells, diseased cells, mutant cells, germ cells, somatic cells, and/or rare cells. Example of rare cells include, without limitations, circulating tumor cells, fetal cells, stem cells, immune cells, clonal cells, and any combination thereof. In some embodiments, the target cells can comprise tumor cells. In some embodiments, the tumor cells can be derived from a tissue biopsy, a fine aspirate or a liquid biopsy (e.g., peritoneal, pleural, cerebrospinal fluid, and/or blood), a mucosal swap, a skin biopsy, a stool sample, or any combinations thereof. In some embodiments, whole cells and/or cell lysates can be analyzed by the methods described herein to detect a plurality of target molecules in a sample. In some embodiments, the whole cells can be obtained from a fixed cell or tissue sample.

[0082] Exemplary target molecules which can be detected by the methods described herein include, but are not limited to proteins, peptides, metabolites, lipids, carbohydrates, toxins, growth factors, hormones, cytokines, cells, and any combinations thereof. In some embodiments, the target molecules to be detected can be extracellular or secreted molecules. In some embodiments, the target molecules to be detected can be intracellular, e.g., cytoplasmic molecules or nuclear molecules.

[0083] To detect intracellular molecules (e.g., intracellular proteins), the target cells in the sample can be permeabilized or lysed (e.g., with a lysis buffer or a surfactant) such that target probes can contact the target intracellular molecules for further processing and analysis. In some embodiments, the lysis buffer can comprise a protease. An exemplary protease is a protease K.

[0084] The identification nucleotide sequences can be released from the bound target probes using any methods known in the art, depending on the types of the cleavable linkers. In some embodiments, the cleavable linker does not comprise a polynucleotide sequence (e.g., a single-stranded polynucleotide sequence) that is complementary (for basepairing) to at least a portion of the identification nucleotide sequence. That is, in these embodiments, the identification nucleotide sequence is not released from the complex by detaching from the complementary polynucleotide sequence (hybridizable linker) coupled to a target-binding molecule. Cleavable, non-hybridizable linkers are known in the art, of which examples include, but are not limited to the ones that are sensitive to an enzyme, pH, temperature, light, shear stress, sonication, a chemical agent (e.g., dithiothreitol), or any combination thereof. In some embodiments, the cleavable linker can be sensitive to light and enzyme degradation.

[0085] In some embodiments where a photocleavable linker is used, the identification nucleotide sequences can be released from the bound target probes by exposing the bound target probes to a light of a specified wavelength. In some embodiments, ultraviolet light can be used to release identification nucleotide sequences from bound target probes.

[0086] The signals from the released identification nucleotide sequences can be detected by various methods known in the art, including, but not limited to sequencing, quantitative polymerase chain reaction (PCR), multiplexed (PCR), mass cytometry, fluorophore-inactivated multiplexed immunofluorescence, hybridization-based methods, fluorescence hybridization-based methods, imaging, and any combinations thereof. In some embodiments, the signals from the released identification nucleotide sequences can be determined by electrophoresis-based methods. In some embodiments, the signals from the released identification nucleotide sequences are not determined by electrophoresis-based methods. Gel electrophoresis-based methods are generally not as quantitative or sensitive as other detection methods described herein such as PCR, fluorescence hybridization-based methods, and nanoString nCounter® hybridization technology, for example, as described in U.S. Pat. No. 8,415,102, and Geiss et al. Nature Biotechnology. 2008. 26(3): 317-325, the contents of each of which is incorporated herein by reference. Thus, gel electrophoresis-based methods do not necessarily have required sensitivity for detection of rare proteins, e.g., proteins that are not generally highly-expressed, or not easily accessible or extracted, such as intracellular proteins. In addition, limited size resolution on gels can limit simultaneous measurements of a large number (e.g., more than 5 or more than 10) of different target molecules, as compared to other detection methods described herein such as PCR, fluorescence hybridization-based methods, and nanoString nCounter® hybridization technology, for example, as described in U.S. Pat. No. 8,415,102, and Geiss et al. Nature Biotechnology. 2008. 26(3): 317-325, the contents of each of which is incorporated herein by reference.

[0087] The nature of the signals from the released identification nucleotide sequences can vary with choice of detection methods and/or detectable labels. In some embodiments, the signals from the released identification nucleotide sequences can be detected by hybridization-based methods. For example, in some embodiments, the method can further comprise, prior to detecting the signals from the released identification nucleotide sequences, coupling the released identification nucleotide sequences to a detection composition comprising a plurality of reporter probes. Each reporter probe in the plurality can comprise (i) a first target probe-specific region that is capable of binding a first portion of the identification nucleotide sequence; and (ii) a detectable label that identifies the reporter probe. In these embodiments, signals from the respective detectable labels of the reporter probes that are coupled to the released identification nucleotide sequences can be detected accordingly. Since the signals are distinguishable for each respective reporter probes that are bound to the identification nucleotide sequences, target-binding molecules can be correspondingly identified, thereby detecting a plurality of target molecules in the sample. Additional information of "reporter probes" will be found in the following "*Reporter Probes*" section.

[0088] In some embodiments, the detection composition used in the methods described herein can additionally comprise a plurality of capture probes as described herein. Additional information of capture probes will be found in the "*Capture Probes*" section below.

[0089] In some embodiments, the method selected to detect signals from the released identification nucleotide sequences does not require amplification of the released identification nucleotide sequences, first target probe-specific region, or the second target probe-specific region. Amplification-free detection methods can minimize any bias or errors introduced during amplification, e.g., due to varying amplification efficiencies among the nucleotide sequences.

[0090] In some embodiments, the identification nucleotide sequences can be detected by nanoString nCounter® hybridization technology, for example, as described in U.S. Pat. No. 8,415,102, and Geiss et al. Nature Biotechnology. 2008. 26(3): 317-325, the contents of each of which is incorporated herein by reference.

[0091] Typically, signals detected from the identification nucleotide sequences of the target probes corresponding to target molecules can be compared to a control reference to account for any non-specific binding. Accordingly, in some embodiments, the composition added to the sample can further comprise a plurality of control probes. Each control probe in the plurality can comprise: (i) a control-binding molecule that specifically binds to one control molecule in the sample; (ii) an identification control sequence that identifies the control-binding molecule; and (iii) a cleavable linker between the control-binding molecule and the identification control sequence. The control-binding molecule can bind to a control protein present in a sample. Non-limiting examples of control

proteins include housekeeping proteins, control IgG isotypes, mutant non-functional or non-binding proteins, and any combinations thereof.

[0092] Signals from the control probes can then be used to threshold the signals from the target probes. Accordingly, in some embodiments, the method can further comprise thresholding the target signals. In some embodiments, the target signals can be thresholded on the basis of nonspecific binding. For example, in some embodiments, the threshold can be determined by using standard deviation and measurement error from at least one or more control proteins. The threshold is generally set to be greater than that of the signals from the non-specific binding. In some embodiments, the threshold can be at least 50% or more (including, e.g., at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or higher) greater than that of the signals from the non-specific binding. In some embodiments, the threshold can be at least 1.1-fold or more (including, e.g., at least 1.2-fold, at least 1.3-fold, at least 1.4-fold, at least 1.5-fold, at least 2-fold, or higher) greater than that of the signals from the non-specific binding.

[0093] In some embodiments, the method can further comprise quantifying the signals (e.g., signals that are above the pre-determined threshold) by normalizing the signals associated with the target probes by the signals associated with the control probes. In some embodiments, the signals can be analyzed and expressed as number of identification nucleotide sequences per target-binding molecule (or target molecule).

[0094] In some embodiments, the methods described herein can complement other art-recognized single-cell proteomic techniques. Exemplary single-cell proteomic techniques include, e.g., mass cytometry and fluorophore-inactivated multiplexed immunofluorescence. See, e.g., Bendall et al. Science 332, 687-696 (2011) and Gerdes et al. Proc. Natl. Acad. Sci. U.S.A. 110, 11982-11987 (2013) for additional information about single-cell mass cytometry and fluorophore-inactivated multiplexed immunofluorescence.

[0095] In some embodiments, the methods, systems and kits described herein can enable measurements of at least two target molecules of different types. For example, the methods, systems, and kits described herein can be used to measure, for example, nucleic acid molecules and proteins, or proteins and metabolites, or proteins and lipids. The measurements of at least two target molecules of different types can be performed simultaneously or sequentially.

[0096] By way of example only, the methods, systems and kits described herein applied to a sample can preserve genetic materials in a sample while detecting other non-genetic target materials in the same sample. This is one of the advantages over existing non-genetic (e.g., proteomic) analysis methods such as flow cytometry and mass cytometry, which generally require an entire cell to measure non-genetic target molecules (e.g., but not limited to proteins). Accordingly, following the non-genetic (e.g., proteomic) measurements, the entire cell including its genetic material is lost.

In flow cytometry, the cell is lost as it goes through the flow chamber to detect fluorescence; in mass cytometry, the cellular sample is vaporized, destroying any genetic material that may be available. Cell vaporization in mass cytometry results in destruction of ~60% of the sample even for proteomic detection let alone recovery genetic material.

[0097] In contrast, the methods and/or systems presented herein employ an identification nucleotide sequence (which comprises nucleotides) as a tag or barcode to label and/or measure non-genetic target molecules (e.g., but not limited to proteins). Thus, the methods and/or systems described herein ensure that any nucleotide-containing materials (e.g., identification nucleotide sequences and even genetic material extracted from cells) will not be destroyed. As such, in one embodiment, the methods to perform simultaneous measurements on the identification nucleotide sequences (barcodes for identification of non-genetic target molecules, e.g., but not limited to proteins) as well as cells' genetic material of interest (including, but not limited to DNA, RNA, microRNAs, long non-coding RNAs, etc.) are essentially the same, except that the complementary probe set (comprising reporter probes and optionally capture probes) has to be expanded to detect not only the identification nucleotide sequences to measure the non-genetic target molecules (e.g., but not limited to proteins), but also the genetic materials (e.g., but not limited to DNA/RNA) from cells.

[0098] Accordingly, in some embodiments, the methods, systems and/or kits described herein for detection of non-genetic target molecules (e.g., but not limited to proteins) can be used in combination with a nuclei acid analysis for genetic materials, for example, to study the non-genetic target molecules (e.g., but not limited to proteins) that interact with genetic materials or genetic regulatory elements. In these embodiments, the methods and systems described herein for detecting a plurality of target molecules in a sample as described herein can further comprise extracting a nucleic acid molecule from the same sample in which target molecules are to be detected. In some embodiments, the methods and systems described herein can further comprise subjecting the extracted nucleic acid molecule to a nucleic acid analysis. Various methods can be used for nucleic acid analysis, including, but not limited to sequencing, next generation sequencing, quantitative polymerase chain reaction (PCR), multiplexed (PCR), DNA sequencing, RNA sequencing, de novo sequencing, next-generation sequencing such as massively parallel signature sequencing (MPSS), polony sequencing, pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing, ion semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, single molecule real time (SMRT) sequencing, nanopore DNA sequencing, sequencing by hybridization, sequencing with mass spectrometry, microfluidic Sanger sequencing, microscopy-based sequencing techniques, RNA polymerase (RNAP) sequencing, fluorescence hybridization-based technology

(e.g., but not limited to nanoString nCounter® technology), any art-recognized nucleic acid detection methods, or any combinations thereof.

[0099] In some embodiments, after a sample and/or non-genetic target molecules have been labeled with a plurality of target probes described herein, the identification nucleotide sequences of the target probes can be released from the bound non-genetic target molecules simultaneously with extraction of nucleic acid molecules (cells' genetic materials) from the same labeled sample. In these embodiments, both the nucleic acid molecules (cells' genetic materials) of interest and the identification nucleotide sequences can be detected simultaneously in a single sample mixture. In one embodiment, both the nucleic acid molecules (cells' genetic materials) of interest and the identification nucleotide sequences can be detected simultaneously in a single sample mixture using nanoString nCounter® analysis system, for example, as described in U.S. Pat. No. 8,415,102, the content of which is incorporated herein by reference. In this embodiment, once the nucleic acid molecules (genetic materials) from cells and the released identification nucleotide sequences are in solution, the solution mixture can be contacted with probe sets comprising both reporter probes and capture probes as described herein for the identification nucleotide sequences as well as for the cell's nucleic acid molecules (cells' genetic materials) of interest. One of the advantages of using nanoString nCounter® hybridization technology is that the analysis can be done on cell lysates, as well as on fixed samples, without the need for amplification that can introduce bias, and with minimal hands-on time preparation. However, other art-recognized methods for nucleic acid analyses or genetic analysis as described herein (e.g., but not limited to sequencing) can also be used for simultaneous detection of both nucleic acid molecules (cells' genetic materials) of interest and released identification nucleotide sequences from bound non-genetic target molecules. For example, in the case of sequencing, both the cells' genetic materials (e.g., DNA and/or mRNA) and the identification nucleotide sequences corresponding non-genetic target molecules can be sequenced together.

[00100] In alternative embodiments, nucleic acid molecules can be extracted from a first portion of a sample, while non-genetic target molecules can be independently derived or obtained from a second portion of the same sample. In these embodiments, the nucleic acid molecules of interest and the non-genetic target molecules can be detected separately to determine expression levels of the nucleic acid molecules of interest and non-genetic target molecules in the same sample. The nucleic acid molecules of interests can be subjected to any art-recognized nucleic acid analysis, while the non-genetic target molecules can be detected through detecting and identifying the corresponding identification nucleotide sequences released from the target probes using the methods, systems and/or kits described herein.

[00101] In some embodiments, the methods, systems and/or kits described herein can be adapted to measure proteins and nucleic acid molecules (cells' genetic materials) present in the same sample. For example, the proteins can be labeled by one or more embodiments of the target probes described herein and detected using the methods, systems and/or kits described herein, while the nucleic acid molecules (cells' genetic materials) can be detected separately or simultaneously by any methods known in the art (e.g., using, in one embodiment, nanoString nCounter® gene expression kit), e.g., for a multi-analyte assay on the same sample. In one embodiment, the sample can comprise cancer cells. The multi-analyte assay can enable generation of an integrated expression profiling for the sample, which can provide information on interaction between the proteins and the nucleic acid molecules, e.g., genetic regulatory elements such as microRNAs. This would be valuable or desirable in cases where rare samples with only limited sample size are available. For example, after labeling a sample or cells or non-genetic target molecules (e.g., proteins) with the target probes each comprising an unique identification nucleotide sequence (where in one embodiment, the identification nucleotide sequences are alien or foreign DNA barcodes), the identification nucleotide sequences (e.g., alien or foreign DNA barcodes) can then be released from the bound cells or target molecules (e.g., proteins) simultaneously with nucleic acid molecules (e.g., RNA and/or DNA) from the same sample or cells, e.g., using lysis buffer with or without additional cleaving agents (e.g., but not limited to, UV and/or chemical agents). Once the nucleic acid molecules (e.g., RNA and/or DNA) from the cells and unique identification nucleotide sequences (e.g., alien DNA barcodes) are in solution, a hybridization assay can be performed. In one embodiment, the hybridization assay can be nanoString nCounter ® analysis assay. In the nCounter® analysis assay, the probe sets can have both reporter probes and capture probes as described herein for the identification nucleotide sequences (e.g., alien DNA barcodes) as well as for the genes of interest. If sample size is not a concern, a sample can be aliquoted or split such that the protein assay and gene expression assay can be run separately to get a readout of both mRNA and protein on the same sample (e.g., a certain population of cells).

[00102] In some embodiments, to optimize the nanoString nCounter® hybridization technology for detection of identification nucleotide sequences and/or cells' genetic materials, one can , for example, make sure that all the probes fall into a linear range when counting them in bulk and expression of one does not saturate the system. This can readily be done with optimization by one of skill in the art depending on the kit of interest.

[00103] In another embodiment, by releasing identification nucleotide sequences from bound target molecules (e.g., proteins), genetic material and the identification nucleotide sequences can be concurrently extracted from a single sample, enabling analyses of protein-DNA-RNA interrelationships.

[00104] While the methods described herein are described in the context where the identification nucleotide sequences are released from bound target probes before detection, in some embodiments, the identification nucleotide sequences do not need to be released from the bound target probes. Accordingly, in some embodiments, the methods described herein can also apply when the identification nucleotide sequences remain bound to target probes during detection.

[00105] In certain embodiments, the methods of detection are performed in multiplex assays, whereby a plurality of target molecules are detected in the same assay (a single reaction mixture). In a one embodiment, the assay is a hybridization assay in which the plurality of target molecules are detected simultaneously. In certain embodiments, the plurality of target molecules detected in the same assay is, at least 2 different target molecules, at least 5 different target molecules, at least 10 different target molecules, at least 20 different target molecules, at least 50 different target molecules, at least 75 different target molecules, at least 100 different target molecules, at least 200 different target molecules, at least 500 different target molecules, or at least 750 different target molecules, or at least 1000 different target molecules. In other embodiments, the plurality of target molecules detected in the same assay is up to 50 different target molecules, up to 100 different target molecules, up to 150 different target molecules, up to 200 different target molecules, up to 300 different target molecules, up to 500 different target molecules, up to 750 different target molecules, up to 1000 different target molecules, up to 2000 different target molecules, or up to 5000 different target molecules. In yet other embodiments, the plurality of target molecules detected is any range in between the foregoing numbers of different target molecules, such as, but not limited to, from 20 to 50 different target molecules, from 50 to 200 different target molecules, from 100 to 1000 different target molecules, or from 500 to 5000 different target molecules.

Target probes

[00106] As used herein, the term "target probe" generally refers to a synthetic molecule that specifically binds to a target molecule for identification and detection. In accordance with various aspects described herein, each target probe comprises: (i) a target-binding molecule that specifically binds to a target molecule in a sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable linker between the target-binding molecule and the identification nucleotide sequence.

[00107] In some embodiments, the cleavable linker does not comprise a polynucleotide sequence (e.g., a single-stranded polynucleotide sequence) that is complementary (for basepairing) to at least a portion of the identification nucleotide sequence. That is, in these embodiments, the identification nucleotide sequence is not released from a target-binding molecule by detaching from the complementary polynucleotide sequence coupled to the target-binding molecule. Accordingly, in

some embodiments, a target probe comprises (i) a target-binding molecule that specifically binds to the target molecule present in the sample; (ii) an identification nucleotide sequence that identifies the target-binding molecule; and (iii) a cleavable, non-hybridizable linker between the target-binding molecule and the identification nucleotide sequence.

[00108] ***Target-binding molecules***: A target-binding molecule is a molecule that specifically binds to target molecule in a sample. As used herein, the term "specifically bind(s)" or "specific binding" refers to a target binding molecule that binds to a target molecule with a greater affinity than when it binds to other non-target molecule under the same conditions. Specific binding is generally indicated by a dissociation constant of 1 μM or lower, e.g., 500 nM or lower, 400 nM or lower, 300 nM or lower, 250 nM or lower, 200 nM or lower, 150 nM or lower, 100 nM or lower, 50 nM or lower, 40 nM or lower, 30 nM or lower, 20 nM or lower, 10 nM or lower, or 1 nM or lower. Typically the nature of the interaction or binding is noncovalent, e.g., by hydrogen, electrostatic, or van der Waals interactions, however, binding can also be covalent. Target-binding molecules can be naturally-occurring, recombinant or synthetic. Examples of the target-binding molecule can include, but are not limited to a nucleic acid, an antibody or a portion thereof, an antibody-like molecule, an enzyme, an antigen, a small molecule, a protein, a peptide, a peptidomimetic, a carbohydrate, an aptamer, and any combinations thereof. In some embodiments, the target-binding molecule does not include a nucleic acid molecule.

[00109] In some embodiments, the target-binding molecules can be modified by any means known to one of ordinary skill in the art. Methods to modify each type of target-binding molecules are well recognized in the art. Depending on the types of target-binding molecules, an exemplary modification includes, but is not limited to genetic modification, biotinylation, labeling (for detection purposes), chemical modification (e.g., to produce derivatives or fragments of the target-binding molecule), and any combinations thereof. In some embodiments, the target-binding molecule can be genetically modified. In some embodiments, the target-binding molecule can be biotinylated.

[00110] In some embodiments, the target-binding molecule can comprise an antibody or a portion thereof, or an antibody-like molecule. An antibody or a portion thereof or antibody-like molecule can detect expression level of a cellular protein (including cell surface proteins, secreted proteins, cytoplasmic proteins, and nuclear proteins), or phosphorylation or other post-translation modification state thereof. In some embodiments, the antibody or a portion thereof or antibody-like molecule can specifically bind to a protein marker present in a rare cell. Examples of a rare cell include, but are not limited to a circulating tumor cell, a fetal cell, and/or a stem cell. In some embodiments, the antibody or a portion thereof or antibody-like molecule can specifically bind to a target marker or protein associated with a condition (e.g., a normal healthy state, or a disease or

disorder). In some embodiments, the antibody or a portion thereof or antibody-like molecule can specifically bind to a target marker or protein associated with cancer. For example, target markers or proteins associated with cancer can be involved in apoptosis, epigenetic, DNA damage, kinases/oncogenes, cancer diagnostic markers, epithelial-mesenchymal transition, autophagy, proliferation, and/or immune response.

[00111] In some embodiments, the target-binding molecule can comprise a cell surface receptor ligand. As used herein, a "cell surface receptor ligand" refers to a molecule that can bind to the outer surface of a cell. Exemplary cell surface receptor ligand includes, for example, a cell surface receptor binding peptide, a cell surface receptor binding glycopeptide, a cell surface receptor binding protein, a cell surface receptor binding glycoprotein, a cell surface receptor binding organic compound, and a cell surface receptor binding drug. Additional cell surface receptor ligands include, but are not limited to, cytokines, growth factors, hormones, antibodies, and angiogenic factors.

[00112] In some embodiments, the target-binding molecule comprises an antibody selected from **Table 1** in the Example, or a fragment thereof.

[00113] In some embodiments, the target-binding molecules of the target probes used in the methods described herein can comprise at least a portion or all of the antibodies listed in **Table 1** in the Example, or fragments thereof.

[00114] ***Identification nucleotide sequences***: As used herein, the term "identification nucleotide sequence" refers to a nucleotide sequence that identifies a specific target-binding molecule. Thus, each identification nucleotide sequence acts as a unique identification code for each target-binding molecule to which it was coupled.

[00115] In some embodiments, the identification nucleotide sequences have minimal or no secondary structures such as any stable intra-molecular base-pairing interaction (e.g., hairpins). Without wishing to be bound by theory, in some embodiments, the minimal secondary structure in the identification nucleotide sequences can provide for better hybridization between a first portion of the identification nucleotide sequence and the reporter probe, and/or between a second portion of the identification nucleotide sequence and the capture probe. In addition, the minimal secondary structure in the identification nucleotide sequence can provide for better binding of the target-binding molecule to the corresponding target molecule. In some embodiments, the identification nucleotide sequences described herein have no significant intra-molecular pairing at a pre-determined annealing temperature. The pre-determined annealing temperature can range from about 65 °C- 80 °C or from about 70 °C- 80°C, or at about 70°C- 75°C.

[00116] In some embodiments, identification nucleotide sequences of the target probes described herein can be selected or designed such that they do not cross-react with or bind to any

nucleic acid sequence in a genome of a subject whose sample is being evaluated. Thus, the identification nucleotide sequences of the target probes used to detect target molecules in a subject's sample can be selected or designed based on nucleotide sequences of a species or genus that share a homology (sequence identity) with that of the subject by no more than 50% or less, including, e.g., no more than 40%, no more than 30%, no more than 20%, no more than 10% or less. In some embodiments, the identification nucleotide sequences can be pre-screened for no homology against major organisms (e.g., but not limited to human and/or other mammals) in the NCBI Reference Sequence (RefSeq) database. By way of example only, in some embodiments, the identification nucleotide sequences of the target probes used in an animal's sample (e.g., a mammal such as a human) can be derived from a plant genome. In one embodiment, the identification nucleotide sequences of the target probes used in a human's sample can be derived from a potato genome. In some embodiments, the identification nucleotide sequence can comprise a sequence selected from Table 2 (SEQ ID NO: 1 to SEQ ID NO: 110), or a fragment thereof.

[00117] Generally, identification nucleotide sequences of the target probes can have any sequence length and can vary depending on a number of factors, including, but not limited to detection methods, and/or the number of target molecules to be detected. For example, in some embodiments, the length of the identification nucleotide sequences can increase to provide sufficient identification of a large number of target molecules in a sample. In some embodiments where a hybridization-based method is used to detect identification nucleotide sequences, the identification nucleotide sequences can have a length sufficient to provide reliable binding to complementary reporter probes and/or capture probes and to generate detectable signals. In some embodiments, the identification nucleotide sequences can have a length of about 30-150 nucleotides, or about 30-100 nucleotides, or about 50-100 nucleotides. In some embodiments, the identification nucleotide sequences can have a length of at least about 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100 or more nucleotides. In some embodiments, the identification nucleotide sequences can have a length of about 70 nucleotides.

[00118] In some embodiments, the identification nucleotide sequences described herein can have a fairly consistent melting temperature (Tm). Without wishing to be bound by theory, the Tm of the identification nucleotide sequences described herein refers to the temperature at which 50% of the oligonucleotide and its complement are in duplex. The consistent Tm among a population of the identification nucleotide sequences allows for the synthesis and hybridization procedures to be tightly optimized, as the optimal conditions are substantially the same for all spots and positions. In some embodiments, the Tm of an identification nucleotide sequence when hybridized to its complementary reporter probes and/or capture probes can be selected to minimize any potential formation of secondary structures (e.g., hairpins) that could interfere with probe hybridization. In

some embodiments, the Tm of an identification nucleotide sequence when hybridized to its complementary reporter probes and/or capture probes can range from about 70-90°C, from about 75-85°C, or from about 79-82°C. In some embodiments, the Tm of an identification nucleotide sequence when hybridized to its complementary reporter probes and/or capture probes can be at least 70°C, at least 75°C, at least 80°C, at least 85°C or higher. In some embodiments, the Tm of an identification nucleotide sequence when hybridized to its complementary reporter probes and/or capture probes can be about 80°C.

[00119] The GC content of the identification nucleotide sequences can vary depending on a number of factors including, e.g., desired lengths of reporter probes and/or capture probes described and/or desired Tm temperatures. For example, when the reporter and/or capture probes are shorter, the GC content of the identification nucleotide sequences can be increased to maintain the desired Tm consistent between the reporter and/or capture probes and identification nucleotide sequences to minimize potential formation of secondary structures (e.g., hairpins) that could interfere with probe hybridization. In one embodiment, the GC content of the identification nucleotide sequence is optimized to maintain the Tm of an identification nucleotide sequence when hybridized to its complementary reporter probes and/or capture probes to be about 80 °C.

[00120] In some embodiments, the identification nucleotide sequences have a balanced GC content. For example, in some embodiments, no single nucleotide in the identification nucleotide sequence can run longer than 3nt. For example, no G nucleotide or C nucleotide runs longer than 3 nt in the identification nucleotide sequence. In one embodiment where the reporter and/or capture probes have a length of about 35 nucleotides, the GC content can be optimized to maintain the Tm of the corresponding identification nucleotide sequences to be about 80 °C; where the identification nucleotide sequence should have a balanced GC content as much as possible to avoid local regions of very high GC or poly C / poly G runs.

[00121] In some embodiments, the identification nucleotide sequences are DNA sequences.

[00122] ***Cleavable linkers***: As used herein, the term "cleavable linker" refers to a linker which is sufficiently stable under one set of conditions, but which is cleaved under a different set of conditions to release the two parts the linker is holding together. In some embodiments, the cleavable linker can be cleaved at least 1.5 times or more (including, e.g., at least 2 times, at least 3 times, at least 4 times, at least 5 times, at least 6 times, at least 7 times, at least 8 times, at least 9 times, at least 10 times, at least 20 times, at least 30 times, at least 40 times, at least 50 times or more) faster under a first reference condition (e.g., with a cleaving agent) than under a second reference condition (e.g., without a cleaving agent).

[00123] For example, a cleavable linker couples an identification nucleotide sequence and a target-binding agent together under one set of conditions and can be cleaved, digested or degraded

under a different set of conditions to release the identification nucleotide sequence. The cleavable linker coupling a target-binding molecule to an identification nucleotide sequence in a target probe can control release of the identification nucleotide sequence from the target probe when needed, e.g., upon binding to a target molecule, such that the identification nucleotide can be released for detection. Cleavable linkers are known in the art, of which examples include, but are not limited to the ones that are sensitive to an enzyme, pH, temperature, light, shear stress, sonication, a chemical agent (e.g., dithiothreitol), or any combination thereof. In some embodiments, the cleavable linker can be sensitive to light and protein degradation, e.g., by an enzyme.

[00124] Cleavable linkers are susceptible to cleavage agents, e.g., hydrolysis, pH, redox potential, and light (e.g., infra-red, and/or UV) or the presence of degradative molecules. Examples of such degradative agents include: redox agents which are selected for particular substrates or which have no substrate specificity, including, e.g., oxidative or reductive enzymes or reductive agents such as mercaptans, present in cells, that can degrade a redox cleavable linker by reduction; esterases; amidases; endosomes or agents that can create an acidic environment, e.g., those that result in a pH of five or lower; enzymes that can hydrolyze or degrade an acid cleavable linker by acting as a general acid, peptidases (which can be substrate specific) and proteases, and phosphatases. In some embodiments, the cleavable linker can be cleavable by a particular enzyme.

[00125] In some embodiments, the cleavable linker is a cleavable, non-hybridizable linker. As used herein, the term "cleavable, non-hybridizable linker" refers to a cleavable linker as defined earlier that does not comprise a polynucleotide sequence (e.g., a single-stranded polynucleotide sequence) complementary (for basepairing) to at least a portion of the identification nucleotide sequence described herein. That is, in these embodiments, the identification nucleotide sequence is not released from the target-binding molecule by detaching from the complementary polynucleotide sequence coupled to the target-binding molecule.

[00126] Exemplary cleavable, non-hybridizable linkers include, but are not limited to, hydrolyzable linkers, redox cleavable linkers (e.g., -S-S- and $-C(R)_2-S-S-$, wherein R is H or C_1-C_6 alkyl and at least one R is C_1-C_6 alkyl such as CH_3 or CH_2CH_3); phosphate-based cleavable linkers (e.g., -O-P(O)(OR)-O-, -O-P(S)(OR)-O-, -O-P(S)(SR)-O-, -S-P(O)(OR)-O-, -O-P(O)(OR)-S-, -S-P(O)(OR)-S-, -O-P(S)(OR)-S-, -S-P(S)(OR)-O-, -O-P(O)(R)-O-, -O-P(S)(R)-O-, -S-P(O)(R)-O-, -S-P(S)(R)-O-, -S-P(O)(R)-S-, -O-P(S)(R)-S-, . -O-P(O)(OH)-O-, -O-P(S)(OH)-O-, -O-P(S)(SH)-O-, -S-P(O)(OH)-O-, -O-P(O)(OH)-S-, -S-P(O)(OH)-S-, -O-P(S)(OH)-S-, -S-P(S)(OH)-O-, -O-P(O)(H)-O-, -O-P(S)(H)-O-, -S-P(O)(H)-O-, -S-P(S)(H)-O-, -S-P(O)(H)-S-, and -O-P(S)(H)-S-, wherein R is optionally substituted linear or branched C1-C10 alkyl); acid cleavable linkers (e.g., hydrazones, esters, and esters of amino acids, -C=NN- and -OC(O)-); ester-based cleavable linkers (e.g., -C(O)O-); peptide-based cleavable linkers, (e.g., linkers that are cleaved by enzymes such as

peptidases and proteases in cells, e.g., -NHCHR^{A}C(O)NHCHR^{B}C(O)-, where R^{A} and R^{B} are the R groups of the two adjacent amino acids), photocleavable linkers and any combinations thereof. A peptide based cleavable linker comprises two or more amino acids. In some embodiments, the peptide-based cleavage linkage comprises the amino acid sequence that is the substrate for a peptidase or a protease. In some embodiments, an acid cleavable linker is cleavable in an acidic environment with a pH of about 6.5 or lower (e.g., about 6.5, 6.0, 5.5, 5.0, or lower), or by agents such as enzymes that can act as a general acid.

[00127] In some embodiments, the cleavable, non-hybridizable linker can comprise a disulfide bond, a tetrazine-trans-cyclooctene group, a sulfhydryl group, a nitrobenzyl group, a nitoindoline group, a bromo hydroxycoumarin group, a bromo hydroxyquinoline group, a hydroxyphenacyl group, a dimethozybenzoin group, or any combinations thereof.

[00128] In some embodiments, the cleavable, non-hybridizable linker can comprise at least one of the linkers shown in **Fig. 2A**.

[00129] In some embodiments, the cleavable, non-hybridizable linker can comprise a photocleavable linker. A photocleavable linker is a linker that can be cleaved by exposure to electromagnetic radiation (e.g., visible light, UV light, infrared, etc.). The wavelength of light necessary to photocleave the linker is dependent upon the structure of the photocleavable linker used. Any art-recognized photocleavable linker can be used for the target probes described herein. Exemplary photocleavable linker include, but are not limited to, chemical molecules containing an o-nitrobenzyl moiety, a p-nitrobenzyl moiety, a m-nitrobenzyl moiety, a nitoindoline moiety, a bromo hydroxycoumarin moiety, a bromo hydroxyquinoline moiety, a hydroxyphenacyl moiety, a dimethozybenzoin moiety, or any combinations thereof.

[00130] Additional exemplary photocleavable groups are generally described and reviewed in Pelliccioli et al., Photoremovable protecting groups: reaction mechanisms and applications, Photochem. Photobiol. Sci. 1 441 -458 (2002); Goeldner and Givens, Dynamic Studies in Biology, Wiley-VCH, Weinheim (2005); Marriott, Methods in Enzymology, Vol. 291, Academic Press, San Diego (1998); Morrison, Bioorganic Photochemistry, Vol. 2, Wiley, New York (1993); Adams and Tsien, Annu. Rev. Physiol. 55 755-784 (1993); Mayer et al., Biologically Active Molecules with a "Light Switch," Angew. Chem. Int. Ed. 45 4900-4921 (2006); Pettit et al., Neuron 19 465-471 (1997); Furuta et al., Brominated 7- hydroxycoumarin-4-ylmethyls: Photolabile protecting groups with biologically useful cross- sections for two photon photolysis, Proc. Natl. Acad. Sci. USA 96 1 193-1200 (1999); and U.S. Patent Nos. 5,430,175; 5,635,608; 5,872,243; 5,888,829; 6,043,065; and Zebala, U.S. Patent Application No. 2010/0105120, the disclosures of which are incorporated by reference herein.

[00131] In some embodiments, the photocleavable linker can generally be described as a chromophore. Examples of chromophores which are photoresponsive to such wavelengths include, but are not limited to, acridines, nitroaromatics, and arylsulfonamides. The efficiency and wavelength at which the chromophore becomes photoactivated and thus releases the identification nucleotide sequences described herein will vary depending on the particular functional group(s) attached to the chromophore. For example, when using nitroaromatics, such as derivatives of o-nitrobenzylic compounds, the absorption wavelength can be significantly lengthened by addition of methoxy groups.

[00132] In some embodiments, the photocleavable linker can comprise a nitro-aromatic compound. Exemplary photocleavable linkers having an ortho-nitro aromatic core scaffold include, but are not limited to, ortho-nitro benzyl ("ONB"), 1 -(2-nitrophenyl)ethyl ("NPE"), alpha- carboxy-2-nitrobenzyl ("CNB"), 4,5-dimethoxy-2-nitrobenzyl ("DMNB"), l-(4,5-dimethoxy- 2-nitrophenyl)ethyl ("DMNPE"), 5-carboxymethoxy-2-nitrobenzyl ("CMNB") and ((5-carboxymethoxy-2-nitrobenzyl)oxy)carbonyl ("CMNCBZ") photolabile cores. It will be appreciated that the substituents on the aromatic core are selected to tailor the wavelength of absorption, with electron donating groups (e.g., methoxy) generally leading to longer wavelength absorption. For example, nitrobenzyl ("NB") and nitrophenylethyl ("NPE") are modified by addition of two methoxy residues into 4,5-dimethoxy-2-nitrobenzyl and 1 -(4,5- dimethoxy-2-nitrophenyl)ethyl, respectively, thereby increasing the absorption wavelength range to 340-360 nm.

[00133] Further, other ortho-nitro aromatic core scaffolds include those that trap nitroso byproducts in a hetero Diels Alder reaction as generally discussed in Zebala, U.S. Patent Application No. 2010/0105120 and Pirrung et al., J. Org. Chem. 68: 1 138 (2003). The nitrodibenzofurane ("NDBF") chromophore offers an extinction coefficient significantly higher in the near UV region but it also has a very high quantum yield for the deprotection reaction and it is suitable for two-photon activation (Momotake et al, The nitrodibenzofuran chromophore: a new caging group for ultra-efficient photolysis in living cells, Nat. Methods 3 35-40 (2006)). The NPP group is an alternative introduced by Pfleiderer et al. that yields a less harmful nitrostyryl species (Walbert et al., Photolabile Protecting Groups for Nucleosides: Mechanistic Studies of the 2-(2-Nitrophenyl)ethyl Group, Helv. Chim. Acta 84 1601 -161 1 (2001)).

[00134] In exemplary embodiments involving UV light, the photocleavable linkers can be selected from the group consisting of alpha-carboxy-2-nitrobenzyl (CNB, 260 nm), l-(2-nitrophenyl)ethyl (NPE, 260 nm), 4,5-dimethoxy-2-nitrobenzyl (DMNB, 355 nm), l-(4,5-dimethoxy-2-nitrophenyl)ethyl (DMNPE, 355 nm), (4,5-dimethoxy-2-nitrobenzoxy)carbonyl (NVOC, 355 nm), 5-carboxymethoxy-2-nitrobenzyl (CMNB, 320 nm), ((5-carboxymethoxy- 2-

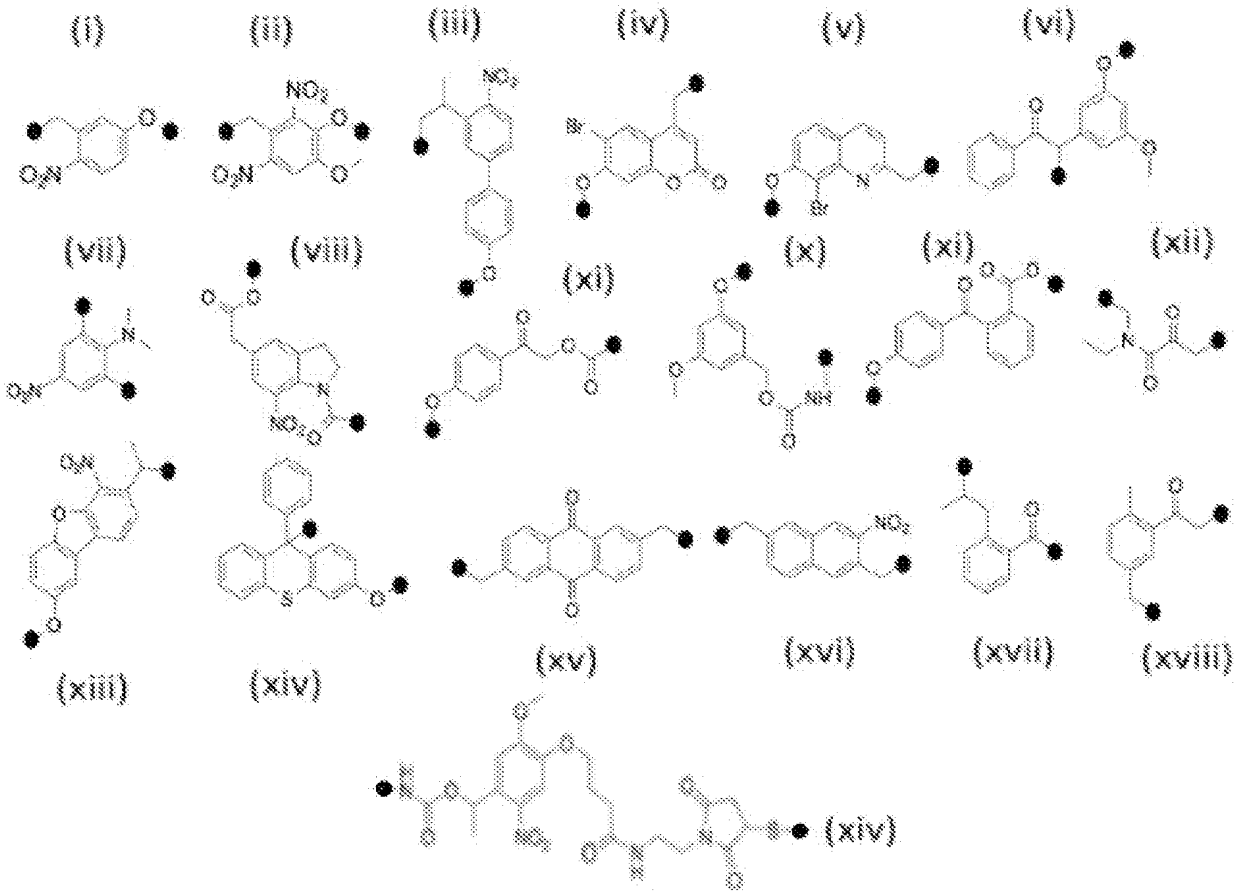
nitrobenzyl)oxy)carbonyl (CMNCBZ, 320 nm), desoxybenzoinyl (desyl, 360 nm), and anthraquino-2-ylmethoxycarbonyl (AQMOC, 350 nm).

[00135] Other suitable photocleavable linkers are based on the coumarin system, such as BHC (Furuta and Iwamura, Methods Enzymol. 291 50-63 (1998); Furuta et al., Proc. Natl. Acad. Sci. USA 96 1 193-1200 (1999); Suzuki et al., Org. Lett. 5:4867 (2003); U.S. Patent No. 6,472,541 , the disclosure of which is incorporated by reference herein). The DMACM linkage photocleaves in nanoseconds (Hagen et al., [7-(Dialkylamino)coumarin-4-yl]methyl- Caged Compounds as Ultrafast and Effective Long-Wavelength Phototriggers of 8-Bromo- Substituted Cyclic Nucleotides, Chem Bio Chem 4 434-442 (2003)) and is cleaved by visible light (U.S. Patent Application Serial No. 11/402,715 the disclosure of which is incorporated by reference herein). Coumarin-based photolabile linkages are also available for linking to aldehydes and ketones (Lu et al., Bhc-diol as a photolabile protecting group for aldehydes and ketones, Org. Lett. 5 2119-2122 (2003)). Closely related analogues, such as BHQ, are also suitable (Fedoryak et al., Brominated hydroxyquinoline as a photolabile protecting group with sensitivity to multiphoton excitation, Org. Lett. 4 3419-3422 (2002)). Another suitable photocleavable linker comprises the pHP group (Park and Givens, J. Am. Chem. Soc. 119:2453 (1997), Givens et al., New Phototriggers 9: p -Hydroxyphenacyl as a C-Terminal Photoremovable Protecting Group for Oligopeptides, J. Am. Chem. Soc. 122 2687-2697 (2000); Zhang et al., J. Am. Chem. Soc. 121 5625-5632, (1999); Conrad et al., J. Am. Chem. Soc. 122 9346-9347 (2000); Conrad et al., Org. Lett. 2 1545-1547 (2000)). A ketoprofen derived photolabile linkage is also suitable (Lukeman et al., Carbanion-Mediated Photocages: Rapid and Efficient Photorelease with Aqueous Compatibility, J. Am. Chem. Soc. 127 7698-7699 (2005)).

[00136] In some embodiments, a photocleavable linker is one whose covalent attachment to an identification nucleotide sequence and/or target-binding agent is reversed (cleaved) by exposure to light of an appropriate wavelength. In some embodiments, release of the identification nucleotide sequences occurs when the conjugate is subjected to ultraviolet light. For example, photorelease of the identification nucleotide sequences can occur at a wavelength ranging from about 200 to 380 nm (the exact wavelength or wavelength range will depend on the specific photocleavable linker used, and can be, for example, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, or 380 or some range therebetween). In some embodiments, release of the identification nucleotide sequences occurs when the conjugate is subjected to visible light. For example, photorelease of the identification nucleotide sequences can occur at a wavelength ranging from about 380 to 780 nm (the exact wavelength or wavelength range will depend on the specific photocleavable linker used, and could be, for example, 380, 400, 450, 500, 550, 600, 650, 700, 750, or 780, or some range therebetween). In some embodiments, release of the identification nucleotide sequences occurs when the conjugate is subjected to infrared light. For example, photorelease of the

identification nucleotide sequences can occur at a wavelength ranging from about 780 to 1200 nm (the exact wavelength or wavelength range will depend on the specific photocleavable linker used, and could be for example, 780, 800, 850, 900, 950, 1000, 1050, 1100, 1150, or 1200, or some range therebetween).

[00137] In some embodiments, a photocleavable linker can be selected from the group consisting of molecules (i)-(xiv) and any combinations thereof, wherein the chemical structures of the molecules (i)-(xiv) are shown as follows:



where each of the black dots in each molecule represents a connecting or coupling point that connects, directly or indirectly, to a target-binding molecule described herein or an indentification nucleotide sequence described herein. The connecting point can be a bond, or comprise an atom, a molecule, and/or a linker described herein. In some embodiments, the connecting point is a bond.

[00138] In some embodiments, the photocleavable linker can comprise the molecule (xiv).

[00139] In some embodiments, the photocleavable linker is a photocleavable bifunctional linker. In some embodiments, the photocleavable linker is a photocleavable multi-functional linker.

[00140] In some embodiments where a photocleavable linker is used, the identification nucleotide sequences can be released from the bound target probes by exposing the bound target probes to a light of a specified wavelength. In some embodiments, ultraviolet (UV) light or near UV light can be used to release identification nucleotide sequences from bound target probes. In some embodiments, release of the identification nucleotide sequences can occur at a wavelength ranging from about 200 nm to about 450 nm.

[00141] Activation agents can be used to activate the components to be conjugated together (e.g., identification nucleotide sequences and/or target-binding molecules). Without limitations, any process and/or reagent known in the art for conjugation activation can be used. Exemplary surface activation method or reagents include, but are not limited to, 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC or EDAC), hydroxybenzotriazole (HOBT), N-Hydroxysuccinimide (NHS), 2-(1H-7-Azabenzotriazol-1-yl)--1,1,3,3-tetramethyl uronium hexafluorophosphate methanaminium (HATU), silanization, sulfosuccinimidyl 6-[3'(2-[pyridyldithio)-propionamido] hexanoate (sulfo-LC-SPDP), 2-iminothiolane (Traut's agent), trans-cyclooctene N-hydroxy-succinimidyl ester (TCO-NHS), surface activation through plasma treatment, and the like.

[00142] Again, without limitations, any art known reactive group can be used for coupling a photocleavable linker between an identification nucleotide sequence and a target-binding molecule. For example, various surface reactive groups can be used for surface coupling including, but not limited to, alkyl halide, aldehyde, amino, bromo or iodoacetyl, carboxyl, hydroxyl, epoxy, ester, silane, thiol, and the like.

Control probes

[00143] As used herein, the term "control probe" generally refers to a synthetic molecule that specifically binds to a control molecule for identification and detection. In accordance with various aspects described herein, each control probe comprises: (i) a control-binding molecule that specifically binds to a control molecule in a sample; (ii) an identification control sequence that identifies the control-binding molecule; and (iii) a cleavable linker between the control-binding molecule and the identification control sequence.

[00144] ***Control-binding molecules:*** A control-binding molecule is a molecule that specifically binds to a control molecule in a sample. Examples of a control protein include, but are not limited to, housekeeping proteins (e.g., GAPDH, actin and/or tubulin), control IgG isotypes, mutant non-functional or non-binding proteins (e.g., nonfunctional or non-binding antibodies, or mutated proteins such as a protein G that has been mutated at the binding site), and any combinations thereof. Typically the nature of the interaction or binding is noncovalent, e.g., by

hydrogen, electrostatic, or van der Waals interactions, however, binding can also be covalent. Control-binding molecules can be naturally-occurring, recombinant or synthetic. Examples of the control-binding molecule can include, but are not limited to a nucleic acid, an antibody or a portion thereof, an antibody-like molecule, an enzyme, an antigen, a small molecule, a protein, a peptide, a peptidomimetic, a carbohydrate, an aptamer, and any combinations thereof. In some embodiments, the control-binding molecule does not include a nucleic acid molecule.

[00145] In some embodiments, the control-binding molecules can be modified by any means known to one of ordinary skill in the art. Methods to modify each type of control-binding molecules are well recognized in the art. Depending on the types of control-binding molecules, an exemplary modification includes, but is not limited to genetic modification, biotinylation, labeling (for detection purposes), chemical modification (e.g., to produce derivatives or fragments of the control-binding molecule), and any combinations thereof. In some embodiments, the control-binding molecule can be genetically modified. In some embodiments, the control-binding molecule can be biotinylated.

[00146] In some embodiments, the control-binding molecule can comprise an antibody or a portion thereof, or an antibody-like molecule. An antibody or a portion thereof or antibody-like molecule can detect expression level of a housekeeping protein, e.g., but not limited to GAPDH, actin, and/or tubulin. In some embodiments, the antibody or a portion thereof or antibody-like molecule can specifically bind to a control IgG isotype. In some embodiments, the antibody or a portion thereof or antibody-like molecule can specifically bind to a mutant non-function or non-binding protein, e.g., a protein G that has been mutated at the binding site.

[00147] ***Identification control sequences:*** As used herein, the term "identification control sequence" refers to a nucleotide sequence that identifies a specific control-binding molecule. Thus, each identification control sequence acts as a unique identification code for each control-binding molecule to which it was coupled.

[00148] In some embodiments, the identification control sequences have minimal or no secondary structures such as any stable intra-molecular base-pairing interaction (e.g., hairpins). Without wishing to be bound by theory, in some embodiments, the minimal secondary structure in the identification control sequences can provide for better hybridization between a first portion of the identification control sequence and the reporter probe, and/or between a second portion of the identification control sequence and the capture probe. In addition, the minimal secondary structure in the identification control sequence can provide for better binding of the control-binding molecule to the corresponding control molecule. In some embodiments, the identification control sequences described herein have no significant intra-molecular pairing at a pre-determined annealing

temperature. The pre-determined annealing temperature can range from about 65 °C- 80 °C or from about 70 °C- 80°C, or at about 70°C- 75°C.

[00149] In some embodiments, identification control sequences of the control probes described herein can be selected or designed such that they do not cross-react with or bind to any nucleic acid sequence in a genome of a subject whose sample is being evaluated. Thus, the identification control sequences of the control probes used to detect control molecules in a subject's sample can be selected or designed based on nucleotide sequences of a species or genus that share a homology (sequence identity) with that of the subject by no more than 50% or less, including, e.g., no more than 40%, no more than 30%, no more than 20%, no more than 10% or less. By way of example only, in some embodiments, the identification control sequences of the control probes used in an animal's sample (e.g., a mammal such as a human) can be derived from a plant genome. In one embodiment, the identification control sequences of the control probes used in a human's sample can be derived from a potato genome. In some embodiments, the identification control sequence can comprise a sequence selected from Table 2 (SEQ ID NO: 1 to SEQ ID NO: 110), or a fragment thereof.

[00150] Generally, identification control sequences of the control probes can have any sequence length and can vary depending on a number of factors, including, but not limited to detection methods, and/or the number of control molecules to be detected. For example, in some embodiments, the length of the identification control sequences can increase to provide sufficient identification of a large number of control molecules in a sample. In some embodiments where a hybridization-based method is used to detect identification control sequences, the identification control sequences can have a length sufficient to provide reliable binding to complementary reporter probes and/or capture probes and to generate detectable signals. In some embodiments, the identification control sequences can have a length of about 30-150 nucleotides, or about 30-100 nucleotides, or about 50-100 nucleotides. In some embodiments, the identification control sequences can have a length of at least about 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100 or more nucleotides. In some embodiments, the identification control sequences can have a length of about 70 nucleotides.

[00151] In some embodiments, the identification control sequences described herein can have a fairly consistent melting temperature (Tm). Without wishing to be bound by theory, the Tm of the identification control sequences described herein refers to the temperature at which 50% of the oligonucleotide and its complement are in duplex. The consistent Tm among a population of the identification control sequences allows for the synthesis and hybridization procedures to be tightly optimized, as the optimal conditions are substantially the same for all spots and positions. In some embodiments, the Tm of an identification control sequence when hybridized to its complementary

reporter probes and/or capture probes can range from about 70-90 °C, from about 75-85 °C, or from about 79-82 °C. In some embodiments, the Tm of an identification control sequence when hybridized to its complementary reporter probes and/or capture probes can be at least 70 °C, at least 75 °C, at least 80°C, at least 85 °C or higher.

[00152] ***Cleavable linkers:*** Any cleavable linkers used in the target probes can be used in the control probes. In some embodiments, the cleavable linker comprises a photocleavable linker. In some embodiments, the photocleavable linker can be selected from the group consisting of molecules (i)-(xiv) shown herein and any combinations thereof. In some embodiments, the photocleavable linker can comprise the molecule (xiv).

Reporter probes

[00153] As used herein, the term "reporter probe" generally refers to a synthetic molecule that binds a first portion of the identification nucleotide sequence of a target probe and generates a detectable signal that is distinguishable for the reporter probe and the bound identification nucleotide sequence.

[00154] In some embodiments, the reporter probes have minimal or no secondary structures such as any stable intra-molecular base-pairing interaction (e.g., hairpins). Without wishing to be bound by theory, the minimal secondary structure in the reporter probes can provide for better hybridization between the reporter probe's backbone and a portion of the identification nucleotide sequences. In addition, the minimal secondary structure in the reporter probes can provide for better detection of the detectable label in the reporter probes. In some embodiments, the reporter probes described herein have no significant intra-molecular pairing at a pre-determined annealing temperature. The pre-determined annealing temperature can range from about 65° C- 80° C or from about 70° C- 80° C, or at about 70°C- 75° C. Secondary structures can be predicted by programs known in the art such as MFOLD.

[00155] In various aspects described herein, a reporter probe generally comprises a detectable label that identifies the reporter probe. As used herein, the term "detectable label" refers to a composition capable of producing a detectable signal indicative of the presence of a target, e.g., a reporter probe bound to an identification nucleotide sequence of a target probe. Detectable labels include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Suitable detectable labels can include fluorescent molecules, radioisotopes, nucleotide chromophores, enzymes, substrates, chemiluminescent moieties, bioluminescent moieties, and the like. As such, a detectable label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means needed for the methods and devices described herein.

[00156] In some embodiments, the detectable label of the reporter probes can comprise one or more labeling molecules that create a unique signal for each reporter probe. In some embodiments, the detectable label of the reporter probes can comprise one labeling molecule. In some embodiments, the detectable label of the reporter probes can comprise at least two or more (e.g., at least 2, at least 3, at least 4, at least 5, at least 6 or more) labeling molecules. As used herein, the term "labeling molecule" is a molecule that is capable of producing a detectable signal, which constitutes at least part of the detectable signal produced by the detectable label. Accordingly, a labeling molecule can be a fluorescent molecule, a radioisotope, a nucleotide chromophore, an enzyme, a substrate, a chemiluminescent moiety, a bioluminescent moiety, or any combinations thereof.

[00157] In some embodiments, the detectable label and/or labeling molecule(s) can generate an optical signal. The optical signal can be a light-emitting signal or a series or sequence of light-emitting signals. In some embodiments, labeling molecules for generation of an optical signal can comprise one or a plurality of (e.g., at least 2 or more, including, e.g., at least 3, at least 4, at least 5 or more) a fluorochrome moiety, a fluorescent moiety, a dye moiety, a chemiluminescent moiety, or any combinations thereof.

[00158] A wide variety of fluorescent reporter dyes are known in the art. Typically, the fluorophore is an aromatic or heteroaromatic compound and can be a pyrene, anthracene, naphthalene, acridine, stilbene, indole, benzindole, oxazole, thiazole, benzothiazole, cyanine, carbocyanine, salicylate, anthranilate, coumarin, fluorescein, rhodamine or other like compound.

[00159] Exemplary fluorophores include, but are not limited to, 1,5 IAEDANS; 1,8-ANS ; 4-Methylumbelliferone; 5-carboxy-2,7-dichlorofluorescein; 5-Carboxyfluorescein (5-FAM); 5-Carboxynapthofluorescein (pH 10); 5-Carboxytetramethylrhodamine (5-TAMRA); 5-FAM (5-Carboxyfluorescein); 5-Hydroxy Tryptamine (HAT); 5-ROX (carboxy-X-rhodamine); 5-TAMRA (5-Carboxytetramethylrhodamine); 6-Carboxyrhodamine 6G; 6-CR 6G; 6-JOE; 7-Amino-4-methylcoumarin; 7-Aminoactinomycin D (7-AAD); 7-Hydroxy-4-methylcoumarin; 9-Amino-6-chloro-2-methoxyacridine; ABQ; Acid Fuchsin; ACMA (9-Amino-6-chloro-2-methoxyacridine); Acridine Orange; Acridine Red; Acridine Yellow; Acriflavin; Acriflavin Feulgen SITSA; Aequorin (Photoprotein); Alexa Fluor 350™; Alexa Fluor 430™; Alexa Fluor 488™; Alexa Fluor 532™; Alexa Fluor 546™; Alexa Fluor 568™; Alexa Fluor 594™; Alexa Fluor 633™; Alexa Fluor 647™; Alexa Fluor 660™; Alexa Fluor 680™; Alizarin Complexon; Alizarin Red; Allophycocyanin (APC); AMC, AMCA-S; AMCA (Aminomethylcoumarin); AMCA-X; Aminoactinomycin D; Aminocoumarin; Anilin Blue; Anthrocyl stearate; APC-Cy7; APTS; Astrazon Brilliant Red 4G; Astrazon Orange R; Astrazon Red 6B; Astrazon Yellow 7 GLL; Atabrine; ATTO-TAG™ CBQCA; ATTO-TAG™ FQ; Auramine; Aurophosphine G; Aurophosphine; BAO 9

(Bisaminophenyloxadiazole); BCECF (high pH); BCECF (low pH); Berberine Sulphate; Beta Lactamase; BFP blue shifted GFP (Y66H); BG-647; Bimane; Bisbenzamide; Blancophor FFG; Blancophor SV; BOBO™ -1; BOBO™ -3; Bodipy 492/515; Bodipy 493/503; Bodipy 500/510; Bodipy 505/515; Bodipy 530/550; Bodipy 542/563; Bodipy 558/568; Bodipy 564/570; Bodipy 576/589; Bodipy 581/591; Bodipy 630/650-X; Bodipy 650/665-X; Bodipy 665/676; Bodipy Fl; Bodipy FL ATP; Bodipy Fl-Ceramide; Bodipy R6G SE; Bodipy TMR; Bodipy TMR-X conjugate; Bodipy TMR-X, SE; Bodipy TR; Bodipy TR ATP; Bodipy TR-X SE; BO-PRO™ -1; BO-PRO™ -3; Brilliant Sulphoflavin FF; Calcein; Calcein Blue; Calcium Crimson™; Calcium Green; Calcium Green-1 Ca2+ Dye; Calcium Green-2 Ca2+; Calcium Green-5N Ca2+; Calcium Green-C18 Ca2+; Calcium Orange; Calcofluor White; Carboxy-X-rhodamine (5-ROX); Cascade Blue™; Cascade Yellow; Catecholamine; CFDA; CFP - Cyan Fluorescent Protein; Chlorophyll; Chromomycin A; Chromomycin A; CMFDA; Coelenterazine ; Coelenterazine cp; Coelenterazine f; Coelenterazine fcp; Coelenterazine h; Coelenterazine hcp; Coelenterazine ip; Coelenterazine O; Coumarin Phalloidin; CPM Methylcoumarin; CTC; Cy2™; Cy3.1 8; Cy3.5™; Cy3™; Cy5.1 8; Cy5.5™; Cy5™; Cy7™; Cyan GFP; cyclic AMP Fluorosensor (FiCRhR); d2; Dabcyl; Dansyl; Dansyl Amine; Dansyl Cadaverine; Dansyl Chloride; Dansyl DHPE; Dansyl fluoride; DAPI; Dapoxyl; Dapoxyl 2; Dapoxyl 3; DCFDA; DCFH (Dichlorodihydrofluorescein Diacetate); DDAO; DHR (Dihydorhodamine 123); Di-4-ANEPPS; Di-8-ANEPPS (non-ratio); DiA (4-Di-16-ASP); DIDS; Dihydorhodamine 123 (DHR); DiO (DiOC18(3)); DiR; DiR (DiIC18(7)); Dopamine; DsRed; DTAF; DY-630-NHS; DY-635-NHS; EBFP; ECFP; EGFP; ELF 97; Eosin; Erythrosin; Erythrosin ITC; Ethidium homodimer-1 (EthD-1); Euchrysin; Europium (III) chloride; Europium; EYFP; Fast Blue; FDA; Feulgen (Pararosaniline); FITC; FL-645; Flazo Orange; Fluo-3; Fluo-4; Fluorescein Diacetate; Fluoro-Emerald; Fluoro-Gold (Hydroxystilbamidine); Fluor-Ruby; FluorX; FM 1-43™; FM 4-46; Fura Red™ (high pH); Fura-2, high calcium; Fura-2, low calcium; Genacryl Brilliant Red B; Genacryl Brilliant Yellow 10GF; Genacryl Pink 3G; Genacryl Yellow 5GF; GFP (S65T); GFP red shifted (rsGFP); GFP wild type, non-UV excitation (wtGFP); GFP wild type, UV excitation (wtGFP); GFPuv; Gloxalic Acid; Granular Blue; Haematoporphyrin; Hoechst 33258; Hoechst 33342; Hoechst 34580; HPTS; Hydroxycoumarin; Hydroxystilbamidine (FluoroGold); Hydroxytryptamine; Indodicarbocyanine (DiD); Indotricarbocyanine (DiR); Intrawhite Cf; JC-1; JO-JO-1; JO-PRO-1; LaserPro; Laurodan; LDS 751; Leucophor PAF; Leucophor SF; Leucophor WS; Lissamine Rhodamine; Lissamine Rhodamine B; LOLO-1; LO-PRO-1; Lucifer Yellow; Mag Green; Magdala Red (Phloxin B); Magnesium Green; Magnesium Orange; Malachite Green; Marina Blue; Maxilon Brilliant Flavin 10 GFF; Maxilon Brilliant Flavin 8 GFF; Merocyanin; Methoxycoumarin; Mitotracker Green FM; Mitotracker Orange; Mitotracker Red; Mitramycin; Monobromobimane; Monobromobimane (mBBr-GSH); Monochlorobimane; MPS (Methyl Green Pyronine Stilbene);

NBD; NBD Amine; Nile Red; Nitrobenzoxadidole; Noradrenaline; Nuclear Fast Red; Nuclear Yellow; Nylosan Brilliant Iavin E8G; Oregon Green™; Oregon Green 488-X; Oregon Green™ 488; Oregon Green™ 500; Oregon Green™ 514; Pacific Blue; Pararosaniline (Feulgen); PE-Cy5; PE-Cy7; PerCP; PerCP-Cy5.5; PE-TexasRed (Red 613); Phloxin B (Magdala Red); Phorwite AR; Phorwite BKL; Phorwite Rev; Phorwite RPA; Phosphine 3R; PhotoResist; Phycoerythrin B [PE]; Phycoerythrin R [PE]; PKH26 ; PKH67; PMIA; Pontochrome Blue Black; POPO-1; POPO-3; PO-PRO-1; PO-PRO-3; Primuline; Procion Yellow; Propidium Iodid (PI); PyMPO; Pyrene; Pyronine; Pyronine B; Pyrozal Brilliant Flavin 7GF; QSY 7; Quinacrine Mustard; Resorufin; RH 414; Rhod-2; Rhodamine; Rhodamine 110; Rhodamine 123; Rhodamine 5 GLD; Rhodamine 6G; Rhodamine B 540; Rhodamine B 200 ; Rhodamine B extra; Rhodamine BB; Rhodamine BG; Rhodamine Green; Rhodamine Phallicidine; Rhodamine Phalloidine; Rhodamine Red; Rhodamine WT; Rose Bengal; R-phycoerythrin (PE); red shifted GFP (rsGFP, S65T); S65A; S65C; S65L; S65T; Sapphire GFP; Serotonin; Sevron Brilliant Red 2B; Sevron Brilliant Red 4G; Sevron Brilliant Red B; Sevron Orange; Sevron Yellow L; sgBFP™; sgBFP™ (super glow BFP); sgGFP™; sgGFP™ (super glow GFP); SITS; SITS (Primuline); SITS (Stilbene Isothiosulphonic Acid); SPQ (6-methoxy-N-(3-sulfopropyl)-quinolinium); Stilbene; Sulphorhodamine B can C; Sulphorhodamine G Extra; Tetracycline; Tetramethylrhodamine ; Texas Red™; Texas Red-X™ conjugate; Thiadicarbocyanine (DiSC3); Thiazine Red R; Thiazole Orange; Thioflavin 5; Thioflavin S; Thioflavin TCN; Thiolyte; Thiozole Orange; Tinopol CBS (Calcofluor White); TMR; TO-PRO-1; TO-PRO-3; TO-PRO-5; TOTO-1; TOTO-3; TriColor (PE-Cy5); TRITC (TetramethylRodamineIsoThioCyanate); True Blue; TruRed; Ultralite; Uranine B; Uvitex SFC; wt GFP; WW 781; XL665; X-Rhodamine; XRITC; Xylene Orange; Y66F; Y66H; Y66W; Yellow GFP; YFP; YO-PRO-1; YO-PRO-3; YOYO-1; and YOYO-3. Many suitable forms of these fluorescent compounds are available and can be used.

[00160] Other exemplary detectable labels and/or labeling molecules include luminescent and bioluminescent markers (e.g., biotin, luciferase (e.g., bacterial, firefly, click beetle and the like), luciferin, and aequorin), radiolabels (e.g., 3H, 125I, 35S, 14C, or 32P), enzymes (e.g., galactosidases, glucorinidases, phosphatases (e.g., alkaline phosphatase), peroxidases (e.g., horseradish peroxidase), and cholinesterases), and calorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, and latex) beads. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837, 3,850,752, 3,939,350, 3,996,345, 4,277,437, 4,275,149, and 4,366,241, each of which is incorporated herein by reference.

[00161] Means of detecting such detectable labels and/or labeling molecules are well known to those of skill in the art. Thus, for example, radiolabels can be detected using photographic film or scintillation counters, fluorescent markers can be detected using a photo-detector to detect emitted light. Enzymatic labels are typically detected by providing the enzyme with an enzyme substrate and

detecting the reaction product produced by the action of the enzyme on the enzyme substrate, and calorimetric labels can be detected by visualizing the colored label.

[00162] In some embodiments, the detectable label and/or labeling molecules can comprise at least one or more (e.g., at least two, at least three, at least four, at least five, at least six, at least or seven or more) fluorophores or quantum dots. Without wishing to be bound by a theory, using a fluorescent reagent can reduce signal-to-noise in the imaging/readout, thus maintaining sensitivity. The color sequence of the labeling molecules in the detectable label can provide an identity to the corresponding reporter probe. For example, a reporter probe I comprises a detectable label with three fluorophores in the following order: fluorophore A; fluorophore B and fluorophore C; whereas a reporter probe II comprises a detectable label with the same three fluorophores but in a different order: fluorophore A; fluorophore C and fluorophore B. While the reporter probes I and II have the same fluorophores, the color sequences of the reporter probe I and reporter probe II are distinct, which identifies the individual reporter probes.

[00163] In some embodiments, the labeling molecule can comprise an enzyme that produces a change in color of an enzyme substrate. A variety of enzymes such as horseradish peroxidase (HRP) and alkaline peroxide (AP) can be used, with either colorimetric or fluorogenic substrates. In some embodiments, the reporter-enzyme produces a calorimetric change which can be measured as light absorption at a particular wavelength. Exemplary enzymes include, but are not limited to, beta-galactosidases, peroxidases, catalases, alkaline phosphatases, and the like.

[00164] In some embodiments, the reporter probe can further comprise a first target probe-specific region that binds to a first portion of the identification nucleotide sequence of a target probe. Accordingly, in some embodiments, a reporter probe can comprise: (a) a first target probe-specific region that binds to a first portion of the identification nucleotide sequence; and (b) a detectable label that identifies the reporter probe.

[00165] As used herein, the term "first target probe-specific region" refers to a region of a reporter probe that binds to a first portion of the identification nucleotide sequence of a target probe. The first target probe-specific region can comprise a fairly regularly-spaced pattern of a nucleotide residue and/or a group (e.g., at least 2 or more) of nucleotide residues in the backbone. In some embodiments, a nucleotide residue and/or a group (e.g., at least 2 or more) of nucleotide residues can be spaced at least an average of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 4, 15, 20, 25, 30, 35, 40, 45, or 50 bases apart within the first target probe-specific region. This allows for a first target probe-specific region having a regularly spaced pattern of a nucleotide or a group of nucleotides binds to a nucleotide sequence that has a complementary nucleotide or a group of complementary nucleotides regularly spaced apart by about the same number of bases. For example, in some embodiments, when the first target probe-specific region contain a fairly regularly-spaced pattern of adenine

residues in the backbone, it can bind a nucleotide sequence that has a thymine residue fairly regularly spaced apart by about the same of number of bases.

[00166] In some embodiments, at least 30% or more (including, e.g., at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 97%, at least 99%, or 100%) of the first target probe-specific region is complementary to a first portion of the identification nucleotide sequence. As used herein and throughout the specification, the term "complementary" refers to a first nucleic acid strand able to form hydrogen bond(s) with a second nucleic acid strand by either traditional Watson-Crick or other non- traditional types. A percent complementarity indicates the percentage of residues in a nucleic acid molecule that can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., 5, 6, 7, 8, 9,10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary). "Perfectly complementary" or 100% complementarity means that all the contiguous residues of a nucleic acid sequence will form hydrogen bonds with the same number of contiguous residues in a second nucleic acid sequence. Less than perfect complementarity refers to the situation in which some, but not all, nucleotides of two strands can form hydrogen bonds with each other. "Substantial complementarity" refers to polynucleotide strands exhibiting 90% or greater complementarity, excluding regions of the polynucleotide strands, such as overhangs, that are selected so as to be non-complementary. Specific binding requires a sufficient degree of complementarity to avoid non-specific binding of the oligomeric sequence to non-target sequences under conditions in which specific binding is desired, i.e., in vitro assays, under conditions in which the assays are performed. The non-target sequences typically differ by at least 1, 2, 3, 4, or 5 nucleotides.

[00167] In some embodiments, the first target probe-specific region can be identified for use in the reporter probe using the methods and computer systems described in U.S. Patent No. 8,415,102 to NanoString Technologies, Inc.

[00168] In some embodiments, the first target probe-specific region and a detectable label can be coupled to each other by at least one or more linkers as described herein. In some embodiments, the linker between the first target probe-specific region and the detectable label can comprise an amide bond. In some embodiments, the linker between the first target probe-specific region and the detectable label can comprise a chemical linker as described herein.

[00169] In some embodiments, the detectable label and/or labeling molecules can be detected using an epifluorescent microscope. In some embodiments, the detectable label and/or labeling molecules can be detected using a fluorescent microscope.

[00170] In some embodiments, the detectable label and/or labeling molecules can be detected through use of spectroscopy. Numerous types of spectroscopic methods can be used. Examples of such methods include, but are not limited to, ultraviolet spectroscopy, visible light spectroscopy,

infrared spectroscopy, x-ray spectroscopy, fluorescence spectroscopy, mass spectroscopy, plasmon resonance (e.g., Cherif et al., Clinical Chemistry, 52:255-262 (2006) and U.S. Pat. No. 7,030,989; herein incorporated by reference), nuclear magnetic resonance spectroscopy, Raman spectroscopy, fluorescence quenching, fluorescence resonance energy transfer, intrinsic fluorescence, ligand fluorescence, and the like.

[00171] In some embodiments, the detectable label and/or labeling molecules can be detected through use of fluorescence anisotropy. Fluorescence anisotropy is based on measuring the steady state polarization of sample fluorescence imaged in a confocal arrangement. A linearly polarized laser excitation source preferentially excites fluorescent target molecules with transition moments aligned parallel to the incident polarization vector. The resultant fluorescence is collected and directed into two channels that measure the intensity of the fluorescence polarized both parallel and perpendicular to that of the excitation beam. With these two measurements, the fluorescence anisotropy, r, can be determined from the equation: r = (Intensity parallel-Intensity perpendicular)/ (Intensity parallel+2(Intensity perpendicular)) where the I terms indicate intensity measurements parallel and perpendicular to the incident polarization. Fluorescence anisotropy detection of fluorescent molecules has been described. Accordingly, fluorescence anisotropy can be coupled to numerous fluorescent labels as have been described herein and as have been described in the art.

[00172] In some embodiments, the detectable label and/or labeling molecules can be detected through use of fluorescence resonance energy transfer (FRET). Fluorescence resonance energy transfer refers to an energy transfer mechanism between two fluorescent molecules. A fluorescent donor is excited at its fluorescence excitation wavelength. This excited state is then nonradiatively transferred to a second molecule, the fluorescent acceptor. Fluorescence resonance energy transfer may be used within numerous configurations to detect the detectable label and/or labeling molecules. For example, in some embodiments, a first labeling molecule can be labeled with a fluorescent donor and second labeling molecule can be labeled with a fluorescent acceptor. Accordingly, such labeled first and second labeling molecules can be used within competition assays to detect the detectable label and/or labeling molecules. Numerous combinations of fluorescent donors and fluorescent acceptors can be used for detection.

[00173] In some embodiments, the detectable and/or labeling molecules can be detected through use of polynucleotide analysis. Examples of such methods include, but are not limited to, those based on polynucleotide hybridization, polynucleotide ligation, polynucleotide amplification, polynucleotide degradation, and the like. Methods that utilize intercalation dyes, fluorescence resonance energy transfer, capacitive deoxyribonucleic acid detection, and nucleic acid amplification have been described, for example, in U.S. Pat. No. 7,118, 910 and No. 6,960,437; herein incorporated by reference). Such methods can be adapted to provide for detection of the detectable

label and/or labeling molecules. In some embodiments, fluorescence quenching, molecular beacons, electron transfer, electrical conductivity, and the like can be used to analyze polynucleotide interaction. Such methods are known and have been described, for example, in Jarvius, DNA Tools and Microfluidic Systems for Molecular Analysis, Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 161, ACTA UNIVERSITATIS UPSALIENSIS UPPSALA 2006, ISBN: 91-554-6616-8; Singh-Zocchi et al, Proc. Natl. Acad. Sci, 100:7605-7610 (2003); Wang et al. Anal. Chem, 75:3941-3945 (2003); and Fan et al, Proc. Natl. Acad. Sci, 100:9134-9137 (2003) and in U.S. Pat. No. 6,958,216; No. 5,093,268; and 6,090,545, the content of all of which is incorporated herein by reference. In some embodiments, the polynucleotide analysis is by polymerase chain reaction (PCR). The fundamentals of PCR are well-known to the skilled artisan, see, e.g. McPherson, et al., PCR, A Practical Approach, IRL Press, Oxford, Eng. (1991), hereby incorporated by reference.

[00174] In some embodiments, the reporter probes can further comprise an affinity tag, which is described in detail in the "*Capture probes*" section below.

[00175] In some embodiments, an affinity tag is placed near or at one end of the reporter probe such that attachment of the reporter probe to a solid substrate surface does not significantly interfere with detection of the detectable label.

[00176] In some embodiments, the reporter probe(s) described herein refers to a "reporter probe" or "labeled nanoreporter probe" or "nanoreporter probe(s)" as described in the U.S. Patent No. 8519115; and US Patent App. Pub. Nos. US 2014/0017688; US 2014/0037620; US2013/0017971; US 2013/0230851; US 2011/0201515; US 2011/0086774; US 2011/0229888; and US 2010/0261026, all of which are assigned to Nanostring Technologies, Inc. and are incorporated herein by reference.

Capture probes

[00177] As used herein, the term "capture probe" generally refers to a synthetic molecule that binds a second portion of the identification nucleotide sequence of a target probe and optionally comprise an affinity tag. As used herein, the term "affinity tag" refers to a molecule that permits reversible or reversible immobilization of the capture probe and bound identification nucleotide sequence to a solid substrate surface, e.g., to allow visualization and/or imaging of the bound complex. In some embodiments, immobilization of the released identification nucleotide sequences can provide distinguishable spatial signals that identify the capture probes coupled to the released identification nucleotide sequences. Examples of a solid substrate include, but are not limited to, a microfluidic device, a cartridge, a microtiter plate, a tube, a magnetic particle, a scaffold, and an array.

[00178] The affinity tag of the capture probe can attach to a solid substrate surface through a covalent or non-covalent interaction. The immobilization or attachment of the affinity tag to a solid substrate surface can occur covalently or non-covalently using any of the methods known to those of skill in the art. For example, covalent immobilization can be accomplished through, for example, silane coupling. See, e.g., Weetall, 15 Adv. Mol. Cell Bio. 161 (2008); Weetall, 44 Meths. Enzymol. 134 (1976). The covalent interaction between the affinity tag and the solid substrate surface can also be mediated by other art-recognized chemical reactions, such as NHS reaction or a conjugation agent. The non-covalent interaction between the affinity tag and the solid substrate surface can be formed based on ionic interactions, van der Waals interactions, dipole-dipole interactions, hydrogen bonds, electrostatic interactions, and/or shape recognition interactions.

[00179] In some embodiments, the affinity tag can comprise a linker as described herein. For example, in some embodiments, the affinity tag can comprise a member of a coupling molecule pair as described in the "Linkers" section below. In some embodiments, the affinity tag can comprise a member of the biotin-avidin or biotin-streptavidin coupling pair. For example, in some embodiments, the affinity tag can comprise a biotin molecule, while the solid surface can be coupled with avidin and/or streptavidin.

[00180] In some embodiments, the affinity tag can comprise a physical linker. For example, the affinity tag can comprise a magnetic particle, where the affinity tag is immobilized to a solid substrate surface under a magnetic force.

[00181] In some embodiments, the capture probes have minimal or no secondary structures such as any stable intra-molecular base-pairing interaction (e.g., hairpins). Without wishing to be bound by theory, the minimal secondary structure in the capture probes can provide for better hybridization between the capture probe's backbone and a portion of the identification nucleotide sequences. In addition, the minimal secondary structure in the capture probes can provide for better attachment of the bound complex (i.e., a complex comprising a capture probe bound to an identification nucleotide sequence) to a solid substrate surface. In some embodiments, the capture probes described herein have no significant intra-molecular pairing at a pre-determined annealing temperature. The pre-determined annealing temperature can range from about 65° C- 80° C or from about 70° C- 80° C, or at about 70°C- 75° C.

[00182] In some embodiments, the capture probe can comprise a second target probe-specific region that binds to a second portion of the identification nucleotide sequence of a target probe. Accordingly, in some embodiments, a capture probe can comprise: (a) a second target probe-specific region that binds to a second portion of the identification nucleotide sequence; and optionally (b) an affinity tag.

[00183] As used herein, the term "second target probe-specific region" refers to a region of a capture probe that binds to a second portion of the identification nucleotide sequence of a target probe. The second target probe-specific region can comprise a fairly regularly-spaced pattern of a nucleotide residue and/or a group (e.g., at least 2 or more) of nucleotide residues in the backbone. In some embodiments, a nucleotide residue and/or a group (e.g., at least 2 or more) of nucleotide residues can be spaced at least an average of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 4, 15, 20, 25, 30, 35, 40, 45, or 50 bases apart within the second target probe-specific region. This allows for a second target probe-specific region having a regularly spaced pattern of a nucleotide or a group of nucleotides binds to a nucleotide sequence that has a complementary nucleotide or a group of complementary nucleotides regularly spaced apart by about the same number of bases. For example, in some embodiments, when the second target probe-specific region contain a fairly regularly-spaced pattern of adenine residues in the backbone, it can bind a nucleotide sequence that has a thymine residue fairly regularly spaced apart by about the same of number of bases.

[00184] In some embodiments, at least 30% or more (including, e.g., at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 97%, at least 99%, or 100%) of the second target probe-specific region is complementary to a second portion of the identification nucleotide sequence.

[00185] In some embodiments, the second target probe-specific region can be identified for use in the capture probe using the methods and computer systems described in U.S. Patent No. 8,415,102 to NanoString Technologies, Inc.

[00186] In some embodiments, the second target probe-specific region and an affinity tag can be coupled to each other by at least one or more linkers as described herein. In some embodiments, the linker between the second target probe-specific region and the affinity tag can comprise an amide bond. In some embodiments, the linker between the second target probe-specific region and the affinity tag can comprise a chemical linker as described herein.

[00187] In some embodiments, the capture probe(s) described herein refers to a "capture probe" or "unlabeled nanoreporter probe" or "nanoreporter probe(s)" as described in the U.S. Patent No. 8519115; and US Patent App. Pub. Nos. US 2014/0017688; US 2014/0037620; US2013/0017971; US 2013/0230851; US 2011/0201515; US 2011/0086774; US 2011/0229888; and US 2010/0261026, all of which are assigned to Nanostring Technologies, Inc. and are incorporated herein by reference.

[00188] Where both reporter probes and capture probes are used in the methods and/or systems described herein, the first target probe-specific region of a reporter probe and the second target probe-specific region of a capture probe should hybridize to a portion of an identification nucleotide sequence at different positions. For example, the portions of the identification nucleotide

sequences to which the target-specific regions of the reporter and capture probes hybridize can be at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 30, at least 40 or more base pairs apart.

Systems, e.g., for multiplexed detection of target molecules in a sample

[00189] Various embodiments of the methods described herein can be carried out in one or more functional modules in a system or a computer system as described herein. Accordingly, another provided herein relates to a system for multiplexed detection of a plurality of target molecules in a sample.

[00190] **Fig. 18A** depicts a device or a computer system 600 comprising one or more processors 630 and a memory 650 storing one or more programs 620 for execution by the one or more processors 630.

[00191] In some embodiments, the device or computer system 600 can further comprise a non-transitory computer-readable storage medium 700 storing the one or more programs 620 for execution by the one or more processors 630 of the device or computer system 600.

[00192] In some embodiments, the device or computer system 600 can further comprise one or more input devices 640, which can be configured to send or receive information to or from any one from the group consisting of: an external device (not shown), the one or more processors 630, the memory 650, the non-transitory computer-readable storage medium 700, and one or more output devices 660.

[00193] In some embodiments, the device or computer system 600 can further comprise one or more output devices 660, which can be configured to send or receive information to or from any one from the group consisting of: an external device (not shown), the one or more processors 630, the memory 650, and the non-transitory computer-readable storage medium 700.

[00194] In some embodiments, the device or computer system 600 for multiplexed detection of target molecules in a sample comprises: one or more processors; and memory to store one or more programs, the one or more programs comprising instructions for:

(a) receiving said at least one test sample comprising a sample and a plurality of target probes described herein;

(b) releasing the identification nucleotide sequences from the target probes that are bound to target molecules in the sample;

(c) detecting signals from the released identification nucleotide sequences;

(d) determining the presence of one or more target molecules in the sample based on the detected signals by performing the following:

i. identifying the detectable probes of the reporter probes that correspond to the detected signals;

ii. identifying the identification nucleotide sequences of the target probes that correspond to the detectable probes based on the first target probe-specific regions of the reporter probes; and

iii. identifying the target-binding molecules that correspond to the identification nucleotide sequences; and

(e) displaying a content based in part on the analysis output from said analysis module, wherein the content comprises a signal indicative of the following: (i) the presence of one or more target molecules in the sample, (ii) the absence of one or more target molecules in the sample, and/or (iii) expression levels of one or more target molecules in the sample.

[00195] **Fig. 18B** depicts a device or a system 600 (e.g., a computer system) for obtaining data from at least one test sample obtained from at least one subject is provided. The system can be used for multiplexed detection of target molecules in a sample. The system comprises:

(a) at least one sample processing module 601 comprising instructions for

receiving said at least one test sample comprising a sample and a plurality of target probes described herein; and

releasing the identification nucleotide sequences from the target probes that are bound to target molecules in the sample;

(b) a signal detection module 602 comprising instructions for detecting signals from the released identification nucleotide sequences;

(c) at least one data storage module 604 comprising instructions for storing the detected signals from (b) and information associated with identification nucleotide sequences of the target probes;

(d) at least one analysis module 606 comprising instructions for determining the presence of one or more target molecules in the sample based on the detected signals; and

(e) at least one display module 610 for displaying a content based in part on the analysis output from said analysis module, wherein the content comprises a signal indicative of the following: (i) the presence of one or more target molecules in the sample, (ii) the absence of one or more target molecules in the sample, and/or (iii) expression levels of one or more target molecules in the sample.

[00196] In some embodiments, the sample processing module 601 can be adapted for isolating target cells, as single cells or as a population, from the sample. In some embodiments, the sample processing module can comprise a microfluidic device for magnetic separation of target cells or interfering cells from a sample using the methods and devices as described in as described in the

International Pat. App. No. WO 2013/078332, the content of which is incorporated herein by reference.

[00197] In some embodiments, the sample processing module 601 can comprise a multi-well plate (e.g., 96-well, 384 wells, or nano- or micro-wells) for single-cell analyses.

[00198] In some embodiments, the sample processing module 601 can be adapted for extracting nucleic acid molecules from the same sample for nucleic acid analysis. Techniques for nucleic acid analysis are known in the art and can be used to assay the test sample to determine nucleic acid or gene expression measurements, for example, but not limited to, DNA sequencing, RNA sequencing, de novo sequencing, next-generation sequencing such as massively parallel signature sequencing (MPSS), polony sequencing, pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing, ion semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, single molecule real time (SMRT) sequencing), nanopore DNA sequencing, sequencing by hybridization, sequencing with mass spectrometry, microfluidic Sanger sequencing, microscopy-based sequencing techniques, RNA polymerase (RNAP) sequencing, or any combinations thereof.

[00199] Accordingly, in some embodiments, the system described herein can be used to generate integrate profiling, e.g., expression profiles of proteins and nucleic acid molecules from the same sample.

[00200] In some embodiments, the sample processing module 601 or the signal detection module 602 can further comprise instructions for contacting the released identification nucleotide sequences with reporter probes described herein.

[00201] In some embodiments, the sample processing module 601 or the signal detection module 602 can further comprise instructions for contacting the released identification nucleotide sequences with capture probes described herein.

[00202] In some embodiments, the sample processing module 601 or the signal detection module 602 can further comprise instructions for immobilizing the released identification nucleotides to a solid substrate through the affinity tag described herein. Examples of a solid substrate include, but are not limited to a microfluidic device, a cartridge, a tube, a microtiter plate, a magnetic particle, and any combinations thereof.

[00203] In some embodiments, the analysis module 606 can further comprise instructions for (i) identifying the detectable probes of the reporter probes that correspond to the detected signals; (ii) identifying the identification nucleotide sequences of the target probes that correspond to the detectable probes based on the first target probe-specific regions of the reporter probes; and (iii) identifying the target-binding molecules that correspond to the identification nucleotide sequences,

thereby determining the presence of one or more target molecules in the sample based on the detected signals.

[00204] In some embodiments, the analysis module 606 can further comprise instructions for identifying a detectable label corresponding for a plurality of light signals emitted from each detectable label, wherein a spatial or temporal order of the plurality of the light signals is unique for each detectable label.

[00205] In some embodiments, the analysis module 606 can further comprise instructions for thresholding the detected signals. For example, the signals can be thresholded on the basis of nonspecific binding. In some embodiments, the threshold is greater than that of the signals from the non-specific binding. By way of example only, the threshold can be determined by using standard deviation and measurement error from at least one control protein. In some embodiments, the threshold can be at least 50% or more (including, e.g., at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or higher) greater than that of the signals from the non-specific binding. In some embodiments, the threshold can be at least 1.1-fold or more (including, e.g., at least 1.2-fold, at least 1.3-fold, at least 1.4-fold, at least 1.5-fold, at least 2-fold, or higher) greater than that of the signals from the non-specific binding.

[00206] In some embodiments, the analysis module 606 can further comprise instructions for quantifying the signals by normalizing the signals associated with the target probes by the signals associated with the control probes. In one embodiment, the signals is quantified and expressed as number of identification nucleotide sequences detected per target-binding agent.

[00207] Depending on the nature of test samples and/or applications of the systems as desired by users, the display module 610 can further display additional content. In some embodiments where the test sample is collected or derived from a subject for diagnostic assessment, the content displayed on the display module 610 can further comprise a signal indicative of a diagnosis of a condition (e.g., disease or disorder such as cancer)

[00208] In some embodiments wherein the test sample is collected or derived from a subject for selection and/or evaluation of a treatment regimen for a subject, the content can further comprise a signal indicative of a treatment regimen personalized to the subject. In some embodiments, the content can further comprise a signal indicative of the treatment response.

[00209] A tangible and non-transitory (e.g., no transitory forms of signal transmission) computer readable medium 700 having computer readable instructions recorded thereon to define software modules for implementing a method on a computer is also provided herein. In some embodiments, the computer readable medium 700 stores one or more programs for multiplexed detection of target molecules in a sample. The one or more programs for execution by one or more processors of a computer system comprises (a) instructions for determining the presence of one or

more target molecules in the sample based on the detected signals from the released identification nucleotide sequences by performing the following: (i) identifying the detectable probes of the reporter probes that correspond to the detected signals; (ii) identifying the identification nucleotide sequences of the target probes that correspond to the detectable probes based on the first target probe-specific regions of the reporter probes; and (iii) identifying the target-binding molecules that correspond to the identification nucleotide sequences; and (b) instructions for displaying a content based in part on the analysis output from said analysis module, wherein the content comprises a signal indicative of the following: (i) the presence of one or more target molecules in the sample, (ii) the absence of one or more target molecules in the sample, and/or (iii) expression levels of one or more target molecules in the sample.

[00210] Depending on the nature of test samples and/or applications of the systems as desired by users, the computer readable storage medium 700 can further comprise instructions for displaying additional content. In some embodiments where the test sample is collected or derived from a subject for diagnostic assessment, the content displayed on the display module can further comprise a signal indicative of a diagnosis of a condition (e.g., disease or disorder) in the subject. In some embodiments wherein the test sample is collected or derived from a subject for selection and/or evaluation of a treatment regimen for a subject, the content can further comprise a signal indicative of a treatment regimen personalized to the subject. In some embodiments, the content can further comprise a signal indicative of the treatment response.

[00211] Embodiments of the systems described herein have been described through functional modules, which are defined by computer executable instructions recorded on computer readable media and which cause a computer to perform method steps when executed. The modules have been segregated by function for the sake of clarity. However, it should be understood that the modules need not correspond to discrete blocks of code and the described functions can be carried out by the execution of various code portions stored on various media and executed at various times. Furthermore, it should be appreciated that the modules may perform other functions, thus the modules are not limited to having any particular functions or set of functions.

[00212] Computing devices typically include a variety of media, which can include computer-readable storage media and/or communications media, in which these two terms are used herein differently from one another as follows. Computer-readable storage media or computer readable media (e.g., 700) can be any available tangible media (e.g., tangible storage media) that can be accessed by the computer, is typically of a non-transitory nature, and can include both volatile and nonvolatile media, removable and non-removable media. By way of example, and not limitation, computer-readable storage media can be implemented in connection with any method or technology for storage of information such as computer-readable instructions, program modules, structured data,

or unstructured data. Computer-readable storage media can include, but are not limited to, RAM (random access memory), ROM (read only memory), EEPROM (erasable programmable read only memory), flash memory or other memory technology, CD-ROM (compact disc read only memory), DVD (digital versatile disk) or other optical disk storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or other tangible and/or non-transitory media which can be used to store desired information. Computer-readable storage media can be accessed by one or more local or remote computing devices, e.g., via access requests, queries or other data retrieval protocols, for a variety of operations with respect to the information stored by the medium.

[00213] On the other hand, communications media typically embody computer-readable instructions, data structures, program modules or other structured or unstructured data in a data signal that can be transitory such as a modulated data signal, e.g., a carrier wave or other transport mechanism, and includes any information delivery or transport media. The term "modulated data signal" or signals refers to a signal that has one or more of its characteristics set or changed in such a manner as to encode information in one or more signals. By way of example, and not limitation, communication media include wired media, such as a wired network or direct-wired connection, and wireless media such as acoustic, radio frequency (RF), infrared and other wireless media.

[00214] In some embodiments, the computer readable storage media 700 can include the "cloud" system, in which a user can store data on a remote server, and later access the data or perform further analysis of the data from the remote server.

[00215] Computer-readable data embodied on one or more computer-readable media, or computer readable medium 700, may define instructions, for example, as part of one or more programs, that, as a result of being executed by a computer, instruct the computer to perform one or more of the functions described herein (e.g., in relation to system 600, or computer readable medium 700), and/or various embodiments, variations and combinations thereof. Such instructions may be written in any of a plurality of programming languages, for example, Java, J#, Visual Basic, C, C#, C++, Fortran, Pascal, Eiffel, Basic, COBOL assembly language, and the like, or any of a variety of combinations thereof. The computer-readable media on which such instructions are embodied may reside on one or more of the components of either of system 600, or computer readable medium 700 described herein, may be distributed across one or more of such components, and may be in transition there between.

[00216] The computer-readable media can be transportable such that the instructions stored thereon can be loaded onto any computer resource to implement the assays and/or methods described herein. In addition, it should be appreciated that the instructions stored on the computer readable media, or computer-readable medium 700, described above, are not limited to instructions embodied

as part of an application program running on a host computer. Rather, the instructions may be embodied as any type of computer code (e.g., software or microcode) that can be employed to program a computer to implement the assays and/or methods described herein. The computer executable instructions may be written in a suitable computer language or combination of several languages. Basic computational biology methods are known to those of ordinary skill in the art and are described in, for example, Setubal and Meidanis et al., Introduction to Computational Biology Methods (PWS Publishing Company, Boston, 1997); Salzberg, Searles, Kasif, (Ed.), Computational Methods in Molecular Biology, (Elsevier, Amsterdam, 1998); Rashidi and Buehler, Bioinformatics Basics: Application in Biological Science and Medicine (CRC Press, London, 2000) and Ouelette and Bzevanis Bioinformatics: A Practical Guide for Analysis of Gene and Proteins (Wiley & Sons, Inc., 2nd ed., 2001).

[00217] The functional modules of certain embodiments of the system or computer system described herein can include a sample processing module, a signal detection module, a storage device, an analysis module and a display module. The functional modules can be executed on one, or multiple, computers, or by using one, or multiple, computer networks. The signal detection module 602 can have computer executable instructions to detect signals from the released identification nucleotide sequences.

[00218] In some embodiments, the signal detection module 602 can have computer executable instructions to provide sequence information in computer readable form, e.g., for RNA sequencing. In these embodiments, the system can enable simultaneous measurements of target molecules (e.g., proteins) and nucleic acid molecules from the same sample. For example, the integrated expression profiles of the protein and nucleic acid molecules can be used to study proteins that interact with genetic regulatory elements such as microRNAs. As used herein, "sequence information" refers to any nucleotide and/or amino acid sequence, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, or mutated sequences. Moreover, information "related to" the sequence information includes detection of the presence or absence of a sequence (e.g., detection of a mutation or deletion), determination of the concentration of a sequence in the sample (e.g., amino acid sequence expression levels, or nucleotide (RNA or DNA) expression levels), and the like. The term "sequence information" is intended to include the presence or absence of post-translational modifications (e.g. phosphorylation, glycosylation, summylation, farnesylation, and the like).

[00219] As an example, signal detection modules 602 for determining sequence information may include known systems for automated sequence analysis including but not limited to Hitachi FMBIO® and Hitachi FMBIO® II Fluorescent Scanners (available from Hitachi Genetic Systems, Alameda, California); Spectrumedix® SCE 9610 Fully Automated 96-Capillary Electrophoresis

Genetic Analysis Systems (available from SpectruMedix LLC, State College, Pennsylvania); ABI PRISM® 377 DNA Sequencer, ABI® 373 DNA Sequencer, ABI PRISM® 310 Genetic Analyzer, ABI PRISM® 3100 Genetic Analyzer, and ABI PRISM® 3700 DNA Analyzer (available from Applied Biosystems, Foster City, California); Molecular Dynamics FluorImager™ 575, SI Fluorescent Scanners, and Molecular Dynamics FluorImager™ 595 Fluorescent Scanners (available from Amersham Biosciences UK Limited, Little Chalfont, Buckinghamshire, England); GenomyxSC™ DNA Sequencing System (available from Genomyx Corporation (Foster City, California); and Pharmacia ALF™ DNA Sequencer and Pharmacia ALFexpress™ (available from Amersham Biosciences UK Limited, Little Chalfont, Buckinghamshire, England).

[00220] Alternative methods for determining sequence information, i.e. signal detection modules 602, include systems for protein and DNA analysis. For example, mass spectrometry systems including Matrix Assisted Laser Desorption Ionization - Time of Flight (MALDI-TOF) systems and SELDI-TOF-MS ProteinChip array profiling systems; systems for analyzing gene expression data (see, for example, published U.S. Patent Application Pub. No. U.S. 2003/0194711); systems for array based expression analysis: e.g., HT array systems and cartridge array systems such as GeneChip® AutoLoader, Complete GeneChip® Instrument System, GeneChip® Fluidics Station 450, GeneChip® Hybridization Oven 645, GeneChip® QC Toolbox Software Kit , GeneChip® Scanner 3000 7G plus Targeted Genotyping System, GeneChip® Scanner 3000 7G Whole-Genome Association System, GeneTitan™ Instrument , and GeneChip® Array Station (each available from Affymetrix, Santa Clara, California); automated ELISA systems (e.g., DSX® or DS2® (available from Dynax, Chantilly, VA) or the Triturus® (available from Grifols USA, Los Angeles, California), The Mago® Plus (available from Diamedix Corporation, Miami, Florida) ; Densitometers (e.g. X-Rite-508-Spectro Densitometer® (available from RP Imaging™, Tucson, Arizona), The HYRYS™ 2 HIT densitometer (available from Sebia Electrophoresis, Norcross, Georgia); automated Fluorescence in situ hybridization systems (see for example, United States Patent 6,136,540); 2D gel imaging systems coupled with 2-D imaging software; microplate readers; Fluorescence activated cell sorters (FACS) (e.g. Flow Cytometer FACSVantage SE, (available from Becton Dickinson, Franklin Lakes, New Jersey); and radio isotope analyzers (e.g. scintillation counters).

[00221] The signals from the released identification nucleotide sequences determined in the signal detection module can be read by the storage device 604. As used herein the "storage device" 604 is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the system described herein can include stand-alone computing apparatus, data telecommunications networks, including local area networks (LAN), wide area networks (WAN), Internet, Intranet, and

Extranet, and local and distributed computer processing systems. Storage devices 604 also include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage media, magnetic tape, optical storage media such as CD-ROM, DVD, electronic storage media such as RAM, ROM, EPROM, EEPROM and the like, general hard disks and hybrids of these categories such as magnetic/optical storage media. The storage device 604 is adapted or configured for having recorded thereon sequence information or expression level information. Such information may be provided in digital form that can be transmitted and read electronically, e.g., via the Internet, on diskette, via USB (universal serial bus) or via any other suitable mode of communication, e.g., the "cloud."

[00222] As used herein, "expression level information" refers to expression levels of any target molecules to be measured, e.g., but not limited to, proteins, peptides, lipids, metabolites, carbohydrates, toxins, growth factors, hormones, cytokines, cells, and any combinations thereof. In some embodiments, the expression level information can be determined from the detected signals from the released identification nucleotide sequences corresponding to each target molecule.

[00223] As used herein, "stored" refers to a process for encoding information on the storage device 604. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the sequence information or expression level information.

[00224] A variety of software programs and formats can be used to store the sequence information or expression level information on the storage device. Any number of data processor structuring formats (e.g., text file or database) can be employed to obtain or create a medium having recorded thereon the sequence information or expression level information.

[00225] By providing sequence information and/or expression level information in computer-readable form, one can use the sequence information and/or expression level information in readable form in the analysis module 606 to generate expression profiles for the sample being tested. The analysis made in computer-readable form provides a computer readable analysis result which can be processed by a variety of means. Content 608 based on the analysis result can be retrieved from the analysis module 606 to indicate the presence or absence of one or more target molecules present in a sample.

[00226] The "analysis module" 606 can use a variety of available software programs and formats for calculating expression profiles of various target molecules. In one embodiment, the analysis module 606 can calculate proteomic expression profiles as follows. First, raw counts of the released identification nucleotide sequences can be first normalized by using the nSolver analysis software to account for hybridization differences on a cartridge, before normalization via the mean of the internal positive controls, which account for hybridization efficiency. These counts can then

be converted to expression values using the relative counts of identification nucleotide sequences per a target-binding agent (e.g., an antibody). Next, average background signal from control IgG can be subtracted. Housekeeping genes can then be used for normalization that accounted for cell number variations. Signals can then be normalized via a housekeeping protein, e.g., GAPDH, actin, and/or β-tubulin.

[00227] In some embodiments, the analysis module 606 can comprise, e.g., MATLAB or functionally equivalent thereof to generate heat maps and clustergrams with a matrix input of marker expression values that were calculated as described above. In some embodiments, the clustergrams can be performed as a weighted linkage. In some embodiments, the clustergrams can be clustered using correlation values as a distance metric. If a target molecule was not detectable, it can be removed from the matrix or heat map and is not displayed.

[00228] In some embodiments, the analysis module 606 can comprise one or more programs for analyzing reporter probes and/or capture probes as described in U.S. Patent No. 7,941,279 to NanoString Technologies, Inc.

[00229] In some embodiments, the analysis module 606 can compare protein expression profiles. Any available comparison software can be used, including but not limited to, the Ciphergen Express (CE) and Biomarker Patterns Software (BPS) package (available from Ciphergen Biosystems, Inc., Freemont, California). Comparative analysis can be done with protein chip system software (e.g., The Protein chip Suite (available from Bio-Rad Laboratories, Hercules, California). Algorithms for identifying expression profiles can include the use of optimization algorithms such as the mean variance algorithm (e.g. JMP Genomics algorithm available from JMP Software Cary, North Carolina).

[00230] The analysis module 606, or any other module of the system described herein, may include an operating system (e.g., UNIX) on which runs a relational database management system, a World Wide Web application, and a World Wide Web server. World Wide Web application includes the executable code necessary for generation of database language statements (e.g., Structured Query Language (SQL) statements). Generally, the executables will include embedded SQL statements. In addition, the World Wide Web application may include a configuration file which contains pointers and addresses to the various software entities that comprise the server as well as the various external and internal databases which must be accessed to service user requests. The Configuration file also directs requests for server resources to the appropriate hardware--as may be necessary should the server be distributed over two or more separate computers. In one embodiment, the World Wide Web server supports a TCP/IP protocol. Local networks such as this are sometimes referred to as "Intranets." An advantage of such Intranets is that they allow easy communication with public domain databases residing on the World Wide Web (e.g., the GenBank or Swiss Pro World Wide

Web site). Thus, in a particular embodiment, users can directly access data (via Hypertext links for example) residing on Internet databases using a HTML interface provided by Web browsers and Web servers. In another embodiment, users can directly access data residing on the "cloud" provided by the cloud computing service providers.

[00231] The analysis module 606 provides computer readable analysis result that can be processed in computer readable form by predefined criteria, or criteria defined by a user, to provide a content based in part on the analysis result that may be stored and output as requested by a user using a display module 610. The display module 610 enables display of a content 608 based in part on the comparison result for the user, wherein the content 608 is a signal indicative of the presence of one or more target molecules in the sample, a signal indicative of the absence of one or more target molecules in the sample, a signal indicative of expression levels of one or more target molecules in the sample, or any combinations thereof. Such signal, can be for example, a display of content 608 on a computer monitor, a printed page of content 608 from a printer, or a light or sound indicative of the absence of a target molecule in a sample.

[00232] In various embodiments of the computer system described herein, the analysis module 606 can be integrated into the signal detection module 602.

[00233] Depending on the nature of test samples and/or applications of the systems as desired by users, the content 608 based on the analysis result can also include a signal indicative of a diagnosis of a condition (e.g., disease or disorder) in the subject. In some embodiments, the content 608 based on the analysis result can further comprise a signal indicative of a treatment regimen personalized to the subject. In some embodiments, the content 608 based on the analysis result can further comprise a signal indicative of a response of a subject to a treatment, which provides a means of monitoring the treatment response in a subject.

[00234] In some embodiments, the content 608 based on the analysis result can include a graphical representation reflecting the expression profiles of target molecules, e.g., as shown in **Fig. 5.**

[00235] In one embodiment, the content 608 based on the analysis result is displayed a on a computer monitor. In one embodiment, the content 608 based on the analysis result is displayed through printable media. The display module 610 can be any suitable device configured to receive from a computer and display computer readable information to a user. Non-limiting examples include, for example, general-purpose computers such as those based on Intel PENTIUM-type processor, Motorola PowerPC, Sun UltraSPARC, Hewlett-Packard PA-RISC processors, any of a variety of processors available from Advanced Micro Devices (AMD) of Sunnyvale, California, or any other type of processor, visual display devices such as flat panel displays, cathode ray tubes and the like, as well as computer printers of various types.

[00236] In some embodiments, the content can be displayed on a computer display, a screen, a monitor, an email, a text message, a website, a physical printout (e.g., paper), or provided as stored information in a data storage device.

[00237] In one embodiment, a World Wide Web browser is used for providing a user interface for display of the content 608 based on the analysis result. It should be understood that other modules of the system described herein can be adapted to have a web browser interface. Through the Web browser, a user may construct requests for retrieving data from the analysis module. Thus, the user will typically point and click to user interface elements such as buttons, pull down menus, scroll bars and the like conventionally employed in graphical user interfaces. The requests so formulated with the user's Web browser are transmitted to a Web application which formats them to produce a query that can be employed to extract the pertinent information related to the expression profile of target molecules in a sample, e.g., display of an indication of the presence or absence of one or more target molecules in the sample, or display of information based thereon. In one embodiment, the information of the control reference is also displayed.

[00238] In any embodiments, the analysis module can be executed by a computer implemented software as discussed earlier. In such embodiments, a result from the analysis module can be displayed on an electronic display. The result can be displayed by graphs, numbers, characters or words. In additional embodiments, the results from the analysis module can be transmitted from one location to at least one other location. For example, the comparison results can be transmitted via any electronic media, e.g., internet, fax, phone, a "cloud" system, and any combinations thereof. Using the "cloud" system, users can store and access personal files and data or perform further analysis on a remote server rather than physically carrying around a storage medium such as a DVD or thumb drive.

[00239] Each of the above identified modules or programs corresponds to a set of instructions for performing a function described above. These modules and programs (i.e., sets of instructions) need not be implemented as separate software programs, procedures or modules, and thus various subsets of these modules may be combined or otherwise re-arranged in various embodiments. In some embodiments, memory may store a subset of the modules and data structures identified above. Furthermore, memory may store additional modules and data structures not described above.

[00240] The illustrated aspects of the disclosure may also be practiced in distributed computing environments where certain tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules can be located in both local and remote memory storage devices.

[00241] Moreover, it is to be appreciated that various components described herein can include electrical circuit(s) that can include components and circuitry elements of suitable value in

order to implement the embodiments of the subject innovation(s). Furthermore, it can be appreciated that many of the various components can be implemented on one or more integrated circuit (IC) chips. For example, in one embodiment, a set of components can be implemented in a single IC chip. In other embodiments, one or more of respective components are fabricated or implemented on separate IC chips.

[00242] What has been described above includes examples of the embodiments of the present invention. It is, of course, not possible to describe every conceivable combination of components or methodologies for purposes of describing the claimed subject matter, but it is to be appreciated that many further combinations and permutations of the subject innovation are possible. Accordingly, the claimed subject matter is intended to embrace all such alterations, modifications, and variations that fall within the spirit and scope of the appended claims. Moreover, the above description of illustrated embodiments of the subject disclosure, including what is described in the Abstract, is not intended to be exhaustive or to limit the disclosed embodiments to the precise forms disclosed. While specific embodiments and examples are described herein for illustrative purposes, various modifications are possible that are considered within the scope of such embodiments and examples, as those skilled in the relevant art can recognize.

[00243] In particular and in regard to the various functions performed by the above described components, devices, circuits, systems and the like, the terms used to describe such components are intended to correspond, unless otherwise indicated, to any component which performs the specified function of the described component (e.g., a functional equivalent), even though not structurally equivalent to the disclosed structure, which performs the function in the herein illustrated exemplary aspects of the claimed subject matter. In this regard, it will also be recognized that the innovation includes a system as well as a computer-readable storage medium having computer-executable instructions for performing the acts and/or events of the various methods of the claimed subject matter.

[00244] The aforementioned systems/circuits/modules have been described with respect to interaction between several components/blocks. It can be appreciated that such systems/circuits and components/blocks can include those components or specified sub-components, some of the specified components or sub-components, and/or additional components, and according to various permutations and combinations of the foregoing. Sub-components can also be implemented as components communicatively coupled to other components rather than included within parent components (hierarchical). Additionally, it should be noted that one or more components may be combined into a single component providing aggregate functionality or divided into several separate sub-components, and any one or more middle layers, such as a management layer, may be provided to communicatively couple to such sub-components in order to provide integrated functionality. Any

components described herein may also interact with one or more other components not specifically described herein but known by those of skill in the art.

[00245] In addition, while a particular feature of the subject innovation may have been disclosed with respect to only one of several implementations, such feature may be combined with one or more other features of the other implementations as may be desired and advantageous for any given or particular application. Furthermore, to the extent that the terms "includes," "including," "has," "contains," variants thereof, and other similar words are used in either the detailed description or the claims, these terms are intended to be inclusive in a manner similar to the term "comprising" as an open transition word without precluding any additional or other elements.

[00246] As used in this application, the terms "component," "module," "system," or the like are generally intended to refer to a computer-related entity, either hardware (e.g., a circuit), a combination of hardware and software, software, or an entity related to an operational machine with one or more specific functionalities. For example, a component may be, but is not limited to being, a process running on a processor (e.g., digital signal processor), a processor, an object, an executable, a thread of execution, a program, and/or a computer. By way of illustration, both an application running on a controller and the controller can be a component. One or more components may reside within a process and/or thread of execution and a component may be localized on one computer and/or distributed between two or more computers. Further, a "device" can come in the form of specially designed hardware; generalized hardware made specialized by the execution of software thereon that enables the hardware to perform specific function; software stored on a computer-readable medium; or a combination thereof.

[00247] In view of the exemplary systems described above, methodologies that may be implemented in accordance with the described subject matter will be better appreciated with reference to the flowcharts of the various figures. For simplicity of explanation, the methodologies are depicted and described as a series of acts. However, acts in accordance with this disclosure can occur in various orders and/or concurrently, and with other acts not presented and described herein. Furthermore, not all illustrated acts may be required to implement the methodologies in accordance with the disclosed subject matter. In addition, those skilled in the art will understand and appreciate that the methodologies could alternatively be represented as a series of interrelated states via a state diagram or events. Additionally, it should be appreciated that the methodologies disclosed in this specification are capable of being stored on an article of manufacture to facilitate transporting and transferring such methodologies to computing devices. The term article of manufacture, as used herein, is intended to encompass a computer program accessible from any computer-readable device or storage media.

[00248] The system 600, and computer readable medium 700, are merely illustrative embodiments, e.g., for multiplexed detection of target molecules in a sample and/or for use in the methods of various aspects described herein and is not intended to limit the scope of the inventions described herein. Variations of system 600, and computer readable medium 700, are possible and are intended to fall within the scope of the inventions described herein.

[00249] The modules of the machine, or used in the computer readable medium, may assume numerous configurations. For example, function may be provided on a single machine or distributed over multiple machines.

Kits, e.g., for multiplexed detection of target molecules in a sample

[00250] Kits, e.g., for multiplexed detection of different target molecules from a sample, are also provided herein. In some embodiments, the kit comprises (a) a plurality of target probes in accordance with one or more embodiments described herein; and (b) a plurality of reporter probes in accordance with one or more embodiments described herein.

[00251] In some embodiments, each subset of the target probes in the plurality binds to a different target molecule, wherein the target probes in the subset comprise the same target-binding molecule. That is, no two target probes in the subset binds to different regions of the same target molecule.

[00252] In some embodiments, the kit comprises at least 3 or more (including at least 4, at least 5, at least 10, at least 15, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250 or more) different target probes described herein, wherein each target probe specifically binds to a different target molecule. Accordingly, in some embodiments, the kit further comprises at least 3 or more (including at least 4, at least 5, at least 10, at least 15, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250 or more) different reporter probes, wherein each reporter probe identifies a distinct target probe.

[00253] In some embodiments, depending on the design of the identification nucleotide sequences of the target probes, the kit can further comprise at least one or more (including at least 4, at least 5, at least 10, at least 15, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250 or more) capture probes described herein. In some embodiments, the same capture probe can be used, e.g., for immobilization of different identification nucleotide sequences to a solid substrate surface for visualization and/or imaging. In some embodiments, different capture probes can be used, e.g., for immobilization of different identification nucleotide sequences to a solid substrate surface for visualization and/or imaging.

[00254] In some embodiments, reporter probes and capture probes can be provided in the kit individually or in a mixture.

[00255] In some embodiments, the target-binding molecules of the target probes can comprise antibodies or fragments thereof. In some embodiments, the antibodies or fragments thereof can be selected from Table 1.

[00256] In some embodiments, the cleavable linker of the target probes can comprise a photocleavable linker. In some embodiments, the photocleavable linker can be selected from the molecules (i)-(xiv) as shown earlier. In some embodiments, the photocleavable linker can comprise molecule (xiv).

[00257] In some embodiments, the detectable label of the reporter probes can comprise one or more labeling molecules that create a unique signal for each reporter probe. An exemplary unique signal can be an optical signal. The optical signal can comprise one or a series or a sequence of light-emitting signals. In these embodiments, non-limiting examples of the labeling molecules include fluorochrome moieties, fluorescent moieties, dye moieties, chemiluminescent moieties, and any combinations thereof.

[00258] In some embodiments, the kit can further comprise a plurality of (e.g., at least 2 or more, including, at least 3, at least 4, at least 5, at least 6, at least 7 or more) control probes in accordance with one or more embodiments described herein.

[00259] In some embodiments, the kit can further comprise reagents for detecting a plurality of nucleic acid molecules. Example reagents for nucleic acid detection and analysis can include, but are not limited to, nucleic acid polymerase, primers, nucleotides, an agent for nucleic acid extraction, a buffered solution, control nucleic acid sequences, and any combination thereof. Such kit can be used to generate an integrated profiling that combines both target molecule (e.g., protein) and genetic material information (e.g., DNA, RNA, epigenetic, and microRNAs). Thus, the kit can be used to study target molecules that interact with genetic materials such as genetic regulatory elements.

[00260] In some embodiments, the kit can further comprise at least one reagent for use in one or more embodiments of the methods or systems described herein. Reagents that can be provided in the kit can include at least one or more of the following: a hybridization reagent, a purification reagent, an immobilization reagent, an imaging agent, a cell permeabilization agent, a blocking agent, a cleaving agent for the cleavable linker, and any combinations thereof.

[00261] In some embodiments, the kit can further include at least one or more devices (e.g., sample cartridges or microfluidic devices) or tubes for use in one or more embodiments of the methods and/or systems described herein. In some embodiments, the device can comprise a surface for immobilization of the capture probes upon coupling to the identification nucleotide sequences. In

some embodiments, the device can comprise a microfluidic device for separating target cells from interfering cells as described herein. For example, a microfluidic device for magnetic separation of target cells or interfering cells from a sample as described in the International Pat. App. No. WO 2013/078332, the content of which are incorporated herein by reference, can be included in the kit.

[00262] In some embodiments, the kit can further include a computer-readable (non-transitory) storage medium in accordance with one or more embodiments described herein. For example, in one embodiment, the computer-readable (non-transitory) storage medium included in the kit can provide instructions to determine the presence or expression levels of one or more target molecules in a sample. The computer-readable (non-transitory) storage medium can be in a CD, DVD, and/or USB drive.

[00263] In all such embodiments of the aspect, the kit includes the necessary packaging materials and informational material therein to store and use said kits. The informational material can be descriptive, instructional, marketing or other material that relates to the methods described herein and/or the use of an agent(s) described herein for the methods described herein. In one embodiment, the informational material can include instructions to perform a multiplexed detection of target molecules in a sample. In one embodiment, the information material can include instructions to analyze the signal readouts.

[00264] The informational material of the kits is not limited in its form. In many cases, the informational material, e.g., instructions, is provided in printed matter, e.g., a printed text, drawing, and/or photograph, e.g., a label or printed sheet. However, the informational material can also be provided in other formats, such as Braille, computer readable material, video recording, or audio recording. In another embodiment, the informational material of the kit is contact information, e.g., a physical address, email address, website, or telephone number, where a user of the kit can obtain substantive information about a compound described herein and/or its use in the methods described herein. Of course, the informational material can also be provided in any combination of formats.

[00265] In all embodiments of the aspects described herein, the kit will typically be provided with its various elements included in one package, e.g., a fiber-based, e.g., a cardboard, or polymeric, e.g., a styrofoam box. The enclosure can be configured so as to maintain a temperature differential between the interior and the exterior, e.g., it can provide insulating properties to keep the reagents at a preselected temperature for a preselected time. The kit can include one or more containers for the composition containing a compound(s) described herein. In some embodiments, the kit contains separate containers (e.g., two separate containers for the two agents), dividers or compartments for the composition(s) and informational material. For example, the composition can be contained in a bottle, vial, or syringe, and the informational material can be contained in a plastic sleeve or packet. In other embodiments, the separate elements of the kit are contained within a

single, undivided container. For example, the composition is contained in a bottle, vial or syringe that has attached thereto the informational material in the form of a label. In some embodiments, the kit includes a plurality (e.g., a pack) of individual containers, each containing one or more unit usage forms of target probes described herein. For example, the kit includes a plurality of syringes, ampules, foil packets, or blister packs, each containing a single unit usage of target probes described herein. The containers of the kits can be air tight, waterproof (e.g., impermeable to changes in moisture or evaporation), and/or light-tight.

Exemplary uses of the methods, systems and kits described herein

[00266] The methods, systems and kits described herein can be used in any applications where detection of a plurality of target molecules in a sample is desirable. For example, a sample can be a biological sample, or a sample from an environmental source (e.g., water, soil, food products, and/or ponds). Other samples that can be analyzed with the methods, systems, and kits described herein are discussed in the "*Sample*" section below.

[00267] The inventors have demonstrated that, in one embodiment, an antibody barcoding with photocleavable DNA (ABCD) platform described herein can enable analysis of hundreds of proteins from a single cell or a limited number of cells, e.g., from minimally invasive fine-needle aspirates (FNAs). Accordingly, samples amenable to the methods described herein can comprise less than 500 cells or fewer. In some embodiments, the sample can comprise less than 400 cells, less than 300 cells, less than 200 cells, less than 100 cells, less than 50 cells, less than 25 cells, less than 5 cells or fewer. In some embodiments, the sample can be a single-cell sample. In some embodiments, the sample can comprise cells isolated from a fine-needle aspirate.

[00268] Where the sample is a biological sample, in some aspects, the methods, systems and kits described herein can be used in personalized treatment. For example, a biological sample can be collected from an individual subject who is in need of a treatment for a condition. Using the methods, systems and/or kits described herein, an expression profile of target molecules associated with the subject's condition can be generated to identify one or more therapeutic targets for the subject, thereby identifying a treatment regimen for the subject. Accordingly, methods for identifying a treatment regimen for an individual subject are also provided herein. In this aspect, the method comprises: (i) contacting a sample derived from a subject who is in need of a treatment for a condition, with a composition comprising a plurality of target probes that bind to target molecules associated with the condition; (ii) releasing the identification nucleotide sequences from the bound target probes; (iii) detecting signals from the released identification nucleotide sequences, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and determining the presence of one or more target

molecules in the sample; and (iv) generating an expression profile of the target molecules detected by the target probes, thereby selecting a treatment regimen for the individual subject based on the expression profile. The methods can be applied to any condition described in the later section. In some embodiments, the condition is cancer. In some embodiments, the signals from the released identification nucleotide sequences are not detected by gel electrophoresis-based methods.

[00269] In some aspects, the methods, systems and kits described herein can be used to assess how drug dosing corresponds to cellular pharmacodynamics and thus used in monitoring response of a subject to a treatment for his/her condition. For example, biological sample(s) can be collected from the subject prior to and/or over the course of the treatment. Using the methods, systems and/or kits described herein, expression profiles of target molecules associated with the subject's condition before and/or over the course of the treatment can be generated for comparison to determine any changes in expression levels of the target molecules in the subject, thereby monitoring the treatment response in the subject. Accordingly, another aspect provided herein relates to a method of monitoring a treatment for a condition in a subject. The method comprises: (i) contacting a sample derived from a subject after a treatment for a condition, with a composition comprising a plurality of target probes that bind to target molecules associated with the condition; (ii) releasing the identification nucleotide sequences from the bound target probes; (iii) detecting signals from the released identification nucleotide sequences, wherein the signals are distinguishable for the released identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and determining the presence of one or more target molecules in the sample; (iv) generating an expression profile of the target molecules detected by the target probes; (v) comparing the expression profile with an expression profile generated from a sample derived from the same subject prior to the treatment or after treatment at an earlier time point; and (vi) determining changes in expression levels of the target molecules, thereby monitoring the treatment for the condition in the subject. In some embodiments, the signals from the released identification nucleotide sequences are not detected by gel electrophoresis-based methods.

[00270] In some embodiments, the method can further comprise administering an alternative treatment for the condition, when there are no substantial changes in expression levels of the target molecules or the changes in expression levels of the target molecules do not represent a reduction in symptoms associated with the condition.

[00271] In some embodiments, the method can further comprise continuing the same treatment for the condition, when the changes in expression levels of the target molecules represent a reduction in symptoms associated with the condition.

[00272] In some aspects, the methods, systems and kits described herein can be used in diagnosing a condition in a subject. For example, a biological sample can be collected from a subject

who is at risk for a condition. Using the methods, systems and/or kits described herein, an expression profile of target molecules associated with the condition to be diagnosed can be generated for comparison with one or more reference expression profiles (e.g., corresponding to a normal healthy subject and/or a subject having the condition to be diagnosed), thereby determining whether the subject is at risk for the condition. Accordingly, provided here is also a method for diagnosing a condition in a subject. The method comprises: (i) contacting a sample derived from a subject who is at risk for a condition, with a composition comprising a plurality of target probes that bind to target molecules associated with the condition; (ii) releasing the identification nucleotide sequences from the bound target probes; (iii) detecting signals from the released identification nucleotide sequences, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and determining the presence of one or more target molecules in the sample; (iv) generating an expression profile of the target molecules detected by the target probes, (v) comparing the expression profile to at least one reference expression profile, thereby determining whether the subject is at risk for the condition. In some embodiments, the signals from the released identification nucleotide sequences are not detected by gel electrophoresis-based methods.

[00273] In some embodiments, a reference expression profile is associated with the condition. In some embodiments, a reference expression profile is associated with a normal healthy subject.

[00274] In some embodiments, the methods described herein can be used to determine intratumoral heterogeneity, which can be used as a biomarker for diagnosis and/or prognosis.

Conditions (e.g., diseases or disorders) amenable to diagnosis, prognosis/monitoring, and/or treatment using methods, systems, kits, or various aspects described herein

[00275] Different embodiments of the methods, systems and/or kits described herein can be used for diagnosis and/or treatment of a disease or disorder in a subject, e.g., a condition afflicting a certain tissue in a subject. For example, the disease or disorder in a subject can be associated with breast, pancreas, blood, prostate, colon, lung, skin, brain, ovary, kidney, oral cavity, throat, cerebrospinal fluid, liver, or other tissues, and any combination thereof.

[00276] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a breast disease or disorder. Exemplary breast disease or disorder includes breast cancer.

[00277] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a pancreatic disease or disorder. Nonlimiting examples of pancreatic diseases or disorders include acute pancreatitis, chronic

pancreatitis, hereditary pancreatitis, pancreatic cancer (e.g., endocrine or exocrine tumors), etc., and any combinations thereof.

[00278] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a blood disease or disorder. Examples of blood disease or disorder include, but are not limited to, platelet disorders, von Willebrand diseases, deep vein thrombosis, pulmonary embolism, sickle cell anemia, thalassemia, anemia, aplastic anemia, fanconi anemia, hemochromatosis, hemolytic anemia, hemophilia, idiopathic thrombocytopenic purpura, iron deficiency anemia, pernicious anemia, polycythemia vera, thrombocythemia and thrombocytosis, thrombocytopenia, and any combinations thereof.

[00279] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a prostate disease or disorder. Non-limiting examples of a prostate disease or disorder can include prostatis, prostatic hyperplasia, prostate cancer, and any combinations thereof.

[00280] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a colon disease or disorder. Exemplary colon diseases or disorders can include, but are not limited to, colorectal cancer, colonic polyps, ulcerative colitis, diverticulitis, and any combinations thereof.

[00281] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a lung disease or disorder. Examples of lung diseases or disorders can include, but are not limited to, asthma, chronic obstructive pulmonary disease, infections, e.g., influenza, pneumonia and tuberculosis, and lung cancer.

[00282] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a skin disease or disorder, or a skin condition. An exemplary skin disease or disorder can include skin cancer.

[00283] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a brain or mental disease or disorder (or neural disease or disorder). Examples of brain diseases or disorders (or neural disease or disorder) can include, but are not limited to, brain infections (e.g., meningitis, encephalitis, brain abscess), brain tumor, glioblastoma, stroke, ischemic stroke, multiple sclerosis (MS), vasculitis, and neurodegenerative disorders (e.g., Parkinson's disease, Huntington's disease, Pick's disease, amyotrophic lateral sclerosis (ALS), dementia, and Alzheimer's disease), Timothy syndrome, Rett symdrome, Fragile X, autism, schizophrenia, spinal muscular atrophy, frontotemporal dementia, any combinations thereof.

[00284] In some embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include a liver disease or disorder. Examples

of liver diseases or disorders can include, but are not limited to, hepatitis, cirrhosis, liver cancer, biliary cirrhosis, fatty liver, nonalcoholic steatohepatitis (NASH), fibrosis, primary sclerosing cholangitis, Budd-Chiari syndrome, hemochromatosis, transthyretin-related hereditary amyloidosis, Gilbert's syndrome, and any combinations thereof.

[00285] In other embodiments, the condition (e.g., disease or disorder) amenable to diagnosis and/or treatment using any aspects described herein can include cancer. A "cancer" or "tumor" as used herein refers to an uncontrolled growth of cells which interferes with the normal functioning of the bodily organs and systems. A subject that has a cancer or a tumor is a subject having objectively measurable cancer cells present in the subject's body. Included in this definition are benign and malignant cancers, as well as dormant tumors or micrometastases. Cancers which migrate from their original location and seed vital organs can eventually lead to the death of the subject through the functional deterioration of the affected organs. Hemopoietic cancers, such as leukemia, are able to out-compete the normal hemopoietic compartments in a subject, thereby leading to hemopoietic failure (in the form of anemia, thrombocytopenia and neutropenia) ultimately causing death.

[00286] By "metastasis" is meant the spread of cancer from its primary site to other places in the body. Cancer cells can break away from a primary tumor, penetrate into lymphatic and blood vessels, circulate through the bloodstream, and grow in a distant focus (metastasize) in normal tissues elsewhere in the body. Metastasis can be local or distant. Metastasis is a sequential process, contingent on tumor cells breaking off from the primary tumor, traveling through the bloodstream, and stopping at a distant site. At the new site, the cells establish a blood supply and can grow to form a life-threatening mass. Both stimulatory and inhibitory molecular pathways within the tumor cell regulate this behavior, and interactions between the tumor cell and host cells in the distant site are also significant.

[00287] Metastases are most often detected through the sole or combined use of magnetic resonance imaging (MRI) scans, computed tomography (CT) scans, blood and platelet counts, liver function studies, chest X-rays and bone scans in addition to the monitoring of specific symptoms.

[00288] Examples of cancer include, but are not limited to carcinoma, lymphoma, blastoma, sarcoma, and leukemia. More particular examples of such cancers include, but are not limited to, basal cell carcinoma, biliary tract cancer; bladder cancer; bone cancer; brain and CNS cancer; breast cancer; cancer of the peritoneum; cervical cancer; choriocarcinoma; colon and rectum cancer; connective tissue cancer; cancer of the digestive system; endometrial cancer; esophageal cancer; eye cancer; cancer of the head and neck; gastric cancer (including gastrointestinal cancer); glioblastoma; hepatic carcinoma; hepatoma; intra-epithelial neoplasm; kidney or renal cancer; larynx cancer; leukemia; liver cancer; lung cancer (e.g., small-cell lung cancer, non-small cell lung cancer, adenocarcinoma of the lung, and squamous carcinoma of the lung); lymphoma including Hodgkin's

and non-Hodgkin's lymphoma; melanoma; myeloma; neuroblastoma; oral cavity cancer (e.g., lip, tongue, mouth, and pharynx); ovarian cancer; pancreatic cancer; prostate cancer; retinoblastoma; rhabdomyosarcoma; rectal cancer; cancer of the respiratory system; salivary gland carcinoma; sarcoma; skin cancer; squamous cell cancer; stomach cancer; testicular cancer; thyroid cancer; uterine or endometrial cancer; cancer of the urinary system; vulval cancer; as well as other carcinomas and sarcomas; as well as B-cell lymphoma (including low grade/follicular non-Hodgkin's lymphoma (NHL); small lymphocytic (SL) NHL; intermediate grade/follicular NHL; intermediate grade diffuse NHL; high grade immunoblastic NHL; high grade lymphoblastic NHL; high grade small non-cleaved cell NHL; bulky disease NHL; mantle cell lymphoma; AIDS-related lymphoma; and Waldenstrom's Macroglobulinemia); chronic lymphocytic leukemia (CLL); acute lymphoblastic leukemia (ALL); Hairy cell leukemia; chronic myeloblastic leukemia; and post-transplant lymphoproliferative disorder (PTLD), as well as abnormal vascular proliferation associated with phakomatoses, edema (such as that associated with brain tumors), and Meigs' syndrome.

[00289] In some embodiments, the methods and systems described herein can be used for determining in a subject a given stage of cancer, e.g., based on the expression profiling generated using the methods described herein. The stage of a cancer generally describes the extent the cancer has progressed and/or spread. The stage usually takes into account the size of a tumor, how deeply the tumor has penetrated, whether the tumor has invaded adjacent organs, how many lymph nodes the tumor has metastasized to (if any), and whether the tumor has spread to distant organs. Staging of cancer is generally used to assess prognosis of cancer as a predictor of survival, and cancer treatment is primarily determined by staging.

Sample

[00290] In accordance with various embodiments described herein, a sample, including any fluid or specimen (processed or unprocessed) or other biological sample, can be subjected to the methods of various aspects described herein.

[00291] In some embodiments, the sample can include a biological fluid obtained from a subject. Exemplary biological fluids obtained from a subject can include, but are not limited to, blood (including whole blood, plasma, cord blood and serum), lactation products (e.g., milk), amniotic fluids (e.g., a sample collected during amniocentesis), sputum, saliva, urine, peritoneal fluid, pleural fluid, semen, cerebrospinal fluid, bronchial aspirate, perspiration, mucus, liquefied feces or stool samples, synovial fluid, lymphatic fluid, tears, tracheal aspirate, and fractions thereof. In some embodiments, a biological fluid can include a homogenate of a tissue specimen (e.g., biopsy) from a subject. In one embodiment, a test sample can comprises a suspension obtained from

homogenization of a solid sample obtained from a solid organ or a fragment thereof. In some embodiments, a sample can be obtained from a mucosal swab. In some embodiments, a sample can be obtained from a tissue biopsy (e.g. but not limited to skin biopsy). In some embodiments, a sample can be a fine needle aspirate.

[00292] In some embodiments, a sample can be obtained from a subject who has or is suspected of having a disease or disorder. In some embodiments, the sample can be obtained from a subject who has or is suspected of having cancer, or who is suspected of having a risk of developing cancer.

[00293] In some embodiments, a sample can be obtained from a subject who is being treated for a disease or disorder. In other embodiments, the sample can be obtained from a subject whose previously-treated disease or disorder is in remission. In other embodiments, the test sample can be obtained from a subject who has a recurrence of a previously-treated disease or disorder. For example, in the case of cancer such as breast cancer, a test sample can be obtained from a subject who is undergoing a cancer treatment, or whose cancer was treated and is in remission, or who has cancer recurrence.

[00294] As used herein, a "subject" can mean a human or an animal. Examples of subjects include primates (e.g., humans, and monkeys). Usually the animal is a vertebrate such as a primate, rodent, domestic animal or game animal. Primates include chimpanzees, cynomologous monkeys, spider monkeys, and macaques, e.g., Rhesus. Rodents include mice, rats, woodchucks, ferrets, rabbits and hamsters. Domestic and game animals include cows, horses, pigs, deer, bison, buffalo, feline species, e.g., domestic cat, canine species, e.g., dog, fox, wolf, and avian species, e.g., chicken, emu, ostrich. A patient or a subject includes any subset of the foregoing, e.g., all of the above, or includes one or more groups or species such as humans, primates or rodents. In certain embodiments of the aspects described herein, the subject is a mammal, e.g., a primate, e.g., a human. The terms, "patient" and "subject" are used interchangeably herein. A subject can be male or female. The term "patient" and "subject" does not denote a particular age. Thus, any mammalian subjects from adult to newborn subjects, as well as fetuses, are intended to be covered.

[00295] In one embodiment, the subject or patient is a mammal. The mammal can be a human, non-human primate, mouse, rat, dog, cat, horse, or cow, but are not limited to these examples. In one embodiment, the subject is a human being. In another embodiment, the subject can be a domesticated animal and/or pet.

[00296] In some embodiments, a sample that can be analyzed by the methods, systems and kits described herein can be obtained from an environmental source. Examples of an environmental source include, but are not limited to, water, soil, food products, ponds, reservoir, and any combinations thereof.

Linkers

[00297] As used herein, the term "linker" generally refers to a molecular entity that can directly or indirectly connect at two parts of a composition. For example, in some embodiments with respect to a reporter probe, the linker directly or indirectly connects a first target probe-specific region to a detectable label described herein. In some embodiments with respect to a capture probe, the linker directly or indirectly connects a second target probe-specific region to an affinity tag described herein. In some embodiments with respect to a target probe, the linker directly or indirectly connects an identification nucleotide sequence to a target-binding molecule. In some embodiments with respect to a control probe, the linker directly or indirectly connects an identification control sequence to a control-binding molecule.

[00298] In some embodiments, a linker can comprise a peptide or nucleic acid linker. The peptide or nucleic acid linker can be configured to have a sequence comprising at least one of the amino acids selected from the group consisting of glycine (Gly), serine (Ser), asparagine (Asn), threonine (Thr), methionine (Met) or alanine (Ala), or at least one of codon sequences encoding the aforementioned amino acids (i.e., Gly, Ser, Asn, Thr, Met or Ala). Such amino acids and corresponding nucleic acid sequences are generally used to provide flexibility of a linker. However, in some embodiments, other uncharged polar amino acids (e.g., Gln, Cys or Tyr), nonpolar amino acids (e.g., Val, Leu, Ile, Pro, Phe, and Trp), or nucleic acid sequences encoding the amino acids thereof can also be included in a linker sequence. In alternative embodiments, polar amino acids or nucleic acid sequence thereof can be added to modulate the flexibility of a linker. One of skill in the art can control flexibility of a linker by varying the types and numbers of residues in the linker. See, e.g., Perham, 30 Biochem. 8501 (1991); Wriggers et al., 80 Biopolymers 736 (2005).

[00299] In alternative embodiments, a linker can comprise a chemical linker of any length. In some embodiments, chemical linkers can comprise a direct bond or an atom such as oxygen or sulfur, a unit such as NH, C(O), C(O)NH, SO, SO_2, SO_2NH, or a chain of atoms, such as substituted or unsubstituted alkyl, substituted or unsubstituted alkenyl, substituted or unsubstituted alkynyl, arylalkyl, arylalkenyl, arylalkynyl, heteroarylalkyl, heteroarylalkenyl, heteroarylalkynyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, aryl, heteroaryl, heterocyclyl, cycloalkyl, cycloalkenyl, alkylarylalkyl, alkylarylalkenyl, alkylarylalkynyl, alkenylarylalkyl, alkenylarylalkenyl, alkenylarylalkynyl, alkynylarylalkyl, alkynylarylalkenyl, alkynylarylalkynyl, alkylheteroarylalkyl, alkylheteroarylalkenyl, alkylheteroarylalkynyl, alkenylheteroarylalkyl, alkenylheteroarylalkenyl, alkenylheteroarylalkynyl, alkynylheteroarylalkyl, alkynylheteroarylalkenyl, alkynylheteroarylalkynyl, alkylheterocyclylalkyl, alkylheterocyclylalkenyl, alkylhererocyclylalkynyl, alkenylheterocyclylalkyl,

alkenylheterocyclylalkenyl, alkenylheterocyclylalkynyl, alkynylheterocyclylalkyl, alkynylheterocyclylalkenyl, alkynylheterocyclylalkynyl, alkylaryl, alkenylaryl, alkynylaryl, alkylheteroaryl, alkenylheteroaryl, alkynylhereroaryl, where one or more methylenes can be interrupted or terminated by O, S, S(O), SO_2, NH, $C(O)N(R^1)_2$, C(O), cleavable linker, substituted or unsubstituted aryl, substituted or unsubstituted heteroaryl, substituted or unsubstituted heterocyclic; where R^1 is hydrogen, acyl, aliphatic or substituted aliphatic. In some embodiments, the chemical linker can be a polymer chain (branched or linear).

[00300] In some embodiments, the chemical linker can comprise a stable or labile (e.g., cleavable) bond or conjugation agent. Exemplary conjugations include, but are not limited to, covalent bond, amide bond, additions to carbon-carbon multiple bonds, azide alkyne Huisgen cycloaddition, Diels-Alder reaction, disulfide linkage, ester bond, Michael additions, silane bond, urethane, nucleophilic ring opening reactions: epoxides, non-aldol carbonyl chemistry, cycloaddition reactions: 1,3-dipolar cycloaddition, temperature sensitive, radiation (IR, near-IR, UV) sensitive bond or conjugation agent, pH-sensitive bond or conjugation agent, non-covalent bonds (e.g., ionic charge complex formation, hydrogen bonding, pi-pi interactions, cyclodextrin/adamantly host guest interaction) and the like.

[00301] As used herein, the term "conjugation agent" means an organic moiety that connects two parts of a compound. Without limitations, any conjugation chemistry known in the art for conjugating two molecules or different parts of a composition together can be used for coupling two parts of a compound. Exemplary coupling molecules and/or functional groups for coupling two parts of a compound include, but are not limited to, a polyethylene glycol (PEG, NH2-PEGX-COOH which can have a PEG spacer arm of various lengths X, where $1 < X < 100$, e.g., PEG-2K, PEG-5K, PEG-10K, PEG-12K, PEG-15K, PEG-20K, PEG-40K, and the like), maleimide conjugation agent, PASylation, HESylation, Bis(sulfosuccinimidyl) suberate conjugation agent, DNA conjugation agent, peptide conjugation agent, silane conjugation agent, hydrolyzable conjugation agent, and any combinations thereof.

[00302] In some embodiments, the linker includes a coupling molecule pair. The terms "coupling molecule pair" and "coupling pair" as used interchangeably herein refer to the first and second molecules that specifically bind to each other. One member of the coupling pair is conjugated to a first entity while the second member is conjugated to a second entity, which is desired to be connected to the first entity. By way of example only, the first entity can be a detectable label of a reporter probe described herein, and the second entity can be a first target probe-specific region of the reporter probe. Thus, the detectable label can be coupled to the first target probe-specific region via a coupling molecule pair. As another example, a solid substrate surface can comprise a first member of the coupling pair, while an affinity tag of a capture probe described here can comprise a

second member of the coupling pair. As used herein, the phrase "first and second molecules that specifically bind to each other" refers to binding of the first member of the coupling pair to the second member of the coupling pair with greater affinity and specificity than to other molecules.

[00303] Exemplary coupling molecule pairs include, without limitations, any haptenic or antigenic compound in combination with a corresponding antibody or binding portion or fragment thereof (e.g., digoxigenin and anti-digoxigenin; mouse immunoglobulin and goat antimouse immunoglobulin) and nonimmunological binding pairs (e.g., biotin-avidin, biotin-streptavidin), hormone (e.g., thyroxine and cortisol-hormone binding protein), receptor-receptor agonist, receptor-receptor antagonist (e.g., acetylcholine receptor-acetylcholine or an analog thereof), IgG-protein A, lectin-carbohydrate, enzyme-enzyme cofactor, enzyme-enzyme inhibitor, and complementary oligonucleotide pairs capable of forming nucleic acid duplexes). The coupling molecule pair can also include a first molecule that is negatively charged and a second molecule that is positively charged.

[00304] One example of using coupling pair conjugation is the biotin-avidin or biotin-streptavidin conjugation. In this approach, a first entity is biotinylated (i.e., the first entity comprise a biotin molecule) and a second entity desired to be connected to the first entity can comprise an avidin or streptavidin. Many commercial kits are also available for biotinylating molecules, such as proteins. For example, an aminooxy-biotin (AOB) can be used to covalently attach biotin to a molecule with an aldehyde or ketone group.

[00305] Still another example of using coupling pair conjugation is double-stranded nucleic acid conjugation. In this approach, a first entity can comprise a first strand of the double-stranded nucleic acid and a second entity desired to be connected to the first entity can comprise a second strand of the double-stranded nucleic acid. Nucleic acids can include, without limitation, defined sequence segments and sequences comprising nucleotides, ribonucleotides, deoxyribonucleotides, nucleotide analogs, modified nucleotides and nucleotides comprising backbone modifications, branchpoints and nonnucleotide residues, groups or bridges.

[00306] In some embodiments, a linker can be a physical substrate, *e.g.*, microparticles or magnetic particles.

[00307] The linkers can be of any shape. In some embodiments, the linkers can be linear. In some embodiments, the linkers can be folded. In some embodiments, the linkers can be branched. In other embodiments, the linker adopts the shape of the physical substrate.

[00308] In some embodiments, the linker can comprise a cleavable linker described herein.

[00309] *Embodiments of various aspects described herein can be defined in any of the following numbered paragraphs:*

1. A method for detecting a plurality of target molecules in a sample comprising:

a. contacting a sample with a composition comprising a plurality of target probes, wherein each target probe in the plurality comprises:
   i. a target-binding molecule that specifically binds to a distinct target molecule in the sample;
   ii. an identification nucleotide sequence that identifies the target-binding molecule; and
   iii. a cleavable linker between the target-binding molecule and the identification nucleotide sequence;
b. separating unbound target probes from a plurality of complexes in the sample, each complex having a target molecule and a single target probe bound thereto, wherein the complex does not have a second target probe binding to a different region of the target molecule;
c. releasing the identification nucleotide sequences from the plurality of complexes;
d. detecting signals from the released identification nucleotide sequences based on a non-gel electrophoresis method, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of different target molecules in the sample.

2. The method of paragraph 1, wherein the non-gel electrophoresis method comprises sequencing, quantitative polymerase chain reaction (PCR), multiplexed (PCR), mass cytometry, fluorophore-inactivated multiplexed immunofluorescence, hybridization-based methods, fluorescence hybridization-based methods, or any combinations thereof.

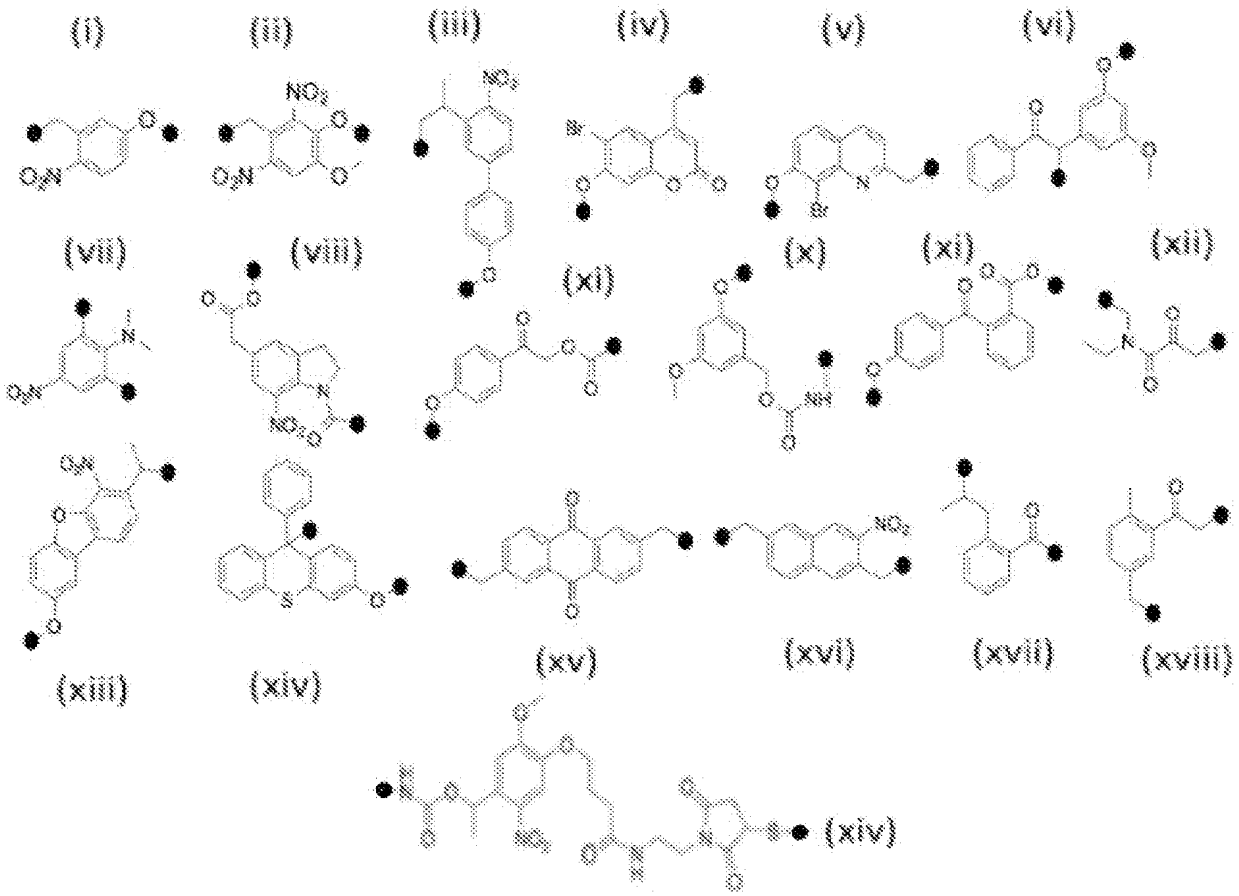
3. The method of paragraph 1 or 2, further comprising, prior to the detecting step (d), coupling the released identification nucleotide sequences from the releasing step (c) to a detection composition comprising a plurality of reporter probes, wherein each reporter probe in the plurality comprises: a first target probe-specific region that is capable of binding a first portion of the identification nucleotide sequence; and a detectable label that identifies the reporter probe.

4. The method of paragraph 3, wherein the detecting comprises detecting signals from the respective detectable labels of the reporter probes that are coupled to the released identification nucleotide sequences, wherein the signals are distinguishable for the respective reporter probes and bound the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of target molecules in the sample.

5. The method of paragraph 3 or 4, wherein the detection composition further comprises a plurality of capture probes, wherein each capture probe comprises (i) a second target probe-specific region that is capable of binding a second portion of the identification nucleotide sequence; and (ii) an affinity tag.

6. The method of paragraph 5, wherein the affinity tag of the capture probe permits immobilization of the released identification nucleotide sequences onto a solid substrate, upon coupling to the detection composition.
7. The method of any of paragraphs 3-6, wherein the detectable label of the reporter probes comprises one or more labeling molecules that create a unique signal for each reporter probe.
8. The method of paragraph 7, wherein the unique signal is an optical signal.
9. The method of paragraph 8, wherein the optical signal comprises a sequence of light-emitting signals.
10. The method of any of paragraphs 7-9, wherein the one or more labeling molecules are selected from the group consisting of a fluorochrome moiety, a fluorescent moiety, a dye moiety, a chemiluminescent moiety, and any combinations thereof.
11. The method of any of paragraphs 1-10, wherein the detecting step (d) comprises no amplification of the released identification nucleotide sequences.
12. The method of any of paragraphs 3-11, wherein the detecting step (d) comprises no amplification of the first target probe-specific region, or the second target probe-specific region.
13. The method of any of paragraphs 1-12, wherein the identification nucleotide sequences are selected such that they do not cross-react with a human genome.
14. The method of paragraph 13, wherein the identification nucleotide sequences are derived from a potato genome.
15. The method of any of paragraphs 1-14, wherein the identification nucleotide sequences have a length of about 30-100 nucleotides.
16. The method of any of paragraphs 1-15, wherein the identification nucleotide sequences have a length of about 70 nucleotides.
17. The method of paragraph 16, wherein the identification nucleotide sequences have a sequence selected from Table 2 (SEQ ID NO: 1 to SEQ ID NO: 110).
18. The method of any of paragraphs 1-17, wherein the cleavable linker is a cleavable, non-hybridizable linker.
19. The method of paragraph 18, wherein the cleavable, non-hybridizable linker is sensitive to an enzyme, pH, temperature, light, shear stress, sonication, a chemical agent (e.g., dithiothreitol), or any combination thereof.
20. The method of paragraph 18 or 19, wherein the cleavable, non-hybridizable linkers are selected from the group consisting of hydrolyzable linkers, redox cleavable linkers, phosphate-based cleavable linkers, acid cleavable linkers, ester-based cleavable linkers, peptide-based cleavable linkers, photocleavable linkers, and any combinations thereof.

21. The method of any of paragraphs 18-20, wherein the cleavable, non-hybridizable linker comprises a disulfide bond, a tetrazine-trans-cyclooctene group, a sulfhydryl group, a nitrobenzyl group, a nitoindoline group, a bromo hydroxycoumarin group, a bromo hydroxyquinoline group, a hydroxyphenacyl group, a dimethozybenzoin group, or any combinations thereof.
22. The method of any of paragraphs 18-21, wherein the cleavable, non-hybridizable linker comprises a photocleavable linker.
23. The method of paragraph 22, wherein the photocleavable linker is selected from the group consisting of molecules (i)-(xiv) and any combinations thereof, wherein the chemical structures of the molecules (i)-(xiv) are shown as follows:



wherein each of the black dots in each molecule represents a connecting or coupling point that connects, directly or indirectly, to the target-binding molecule or the indentification nucleotide sequence.

24. The method of paragraph 22, wherein the photocleavable linker comprises the molecule (xiv).

25. The method of any of paragraphs 22-24, wherein the releasing of the identification nucleotide sequences from the bound target probes comprises exposing the bound target probes to ultraviolet light.
26. The method of any of paragraphs 1-25, wherein the sample comprises less than 500 cells.
27. The method of any of paragraphs 1-26, wherein the sample is a single-cell sample.
28. The method of any of paragraphs 1-27, wherein the sample comprises cells isolated from a fine-needle aspirate.
29. The method of any of paragraphs 1-28, further comprising, prior to the contacting, separating target cells from interfering cells in the sample.
30. The method of paragraph 29, wherein the separating comprises labeling the interfering cells or target cells with magnetic particles and separating them from the sample by magnetic separation.
31. The method of paragraph 30, wherein the magnetic separation is performed in a microfluidic device.
32. The method of any of paragraphs 29-31, wherein the target cells comprise rare cells.
33. The method of paragraph 32, wherein the rare cells are selected from the group consisting of circulating tumor cells, fetal cells, stem cells, immune cells, clonal cells, and any combination thereof.
34. The method of any of paragraphs 29-33, wherein the target cells comprise tumor cells from a liquid biopsy (e.g., peritoneal, pleural, cerebrospinal fluid, and/or blood), a mucosal swap, a skin biopsy, a stool sample, or any combinations thereof.
35. The method of any of paragraphs 1-34, wherein the target molecules comprise proteins, peptides, metabolites, lipids, carbohydrates, toxins, growth factors, hormones, cytokines, cells, and any combinations thereof.
36. The method of any of paragraphs 1-35, further comprising permeabilizing the target cells in the sample.
37. The method of any of paragraphs 1-36, wherein the composition further comprises a plurality of control probes, wherein each control probe in the plurality comprises:
    i. a control-binding molecule that specifically binds to one control molecule in the sample;
    ii. an identification control sequence that identifies the control-binding molecule; and
    iii. a cleavable linker between the control-binding molecule and the identification control sequence.
38. The method of paragraph 37, wherein the control-binding molecule binds to a control protein.

39. The method of paragraph 38, wherein the control protein is selected from the group consisting of housekeeping proteins, control IgG isotypes, mutant non-functional or non-binding proteins, and any combinations thereof.
40. The method of any of paragraphs 1-39, further comprising thresholding the signals.
41. The method of paragraph 40, wherein the signals are thresholded on the basis of nonspecific binding.
42. The method of paragraph 41, wherein the threshold is greater than that of the signals from the non-specific binding.
43. The method of any of paragraphs 40-42, wherein the threshold is determined by using standard deviation and measurement error from at least one control protein.
44. The method of any of paragraphs 1-43, further comprising quantifying the signals by normalizing the signals associated with the target probes by the signals associated with the control probes.
45. The method of any of paragraphs 1-44, further comprising extracting a nucleic acid molecule from the same sample for nucleic acid analysis.
46. The method of paragraph 45, further comprising subjecting the nucleic acid molecule to a nucleic acid analysis.
47. The method of paragraph 46, wherein the nucleic acid analysis comprises sequencing, quantitative polymerase chain reaction (PCR), multiplexed PCR, DNA sequencing, RNA sequencing, de novo sequencing, next-generation sequencing such as massively parallel signature sequencing (MPSS), polony sequencing, pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing, ion semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, single molecule real time (SMRT) sequencing, nanopore DNA sequencing, sequencing by hybridization, sequencing with mass spectrometry, microfluidic Sanger sequencing, microscopy-based sequencing techniques, RNA polymerase (RNAP) sequencing, or any combinations thereof.
48. The method of any of paragraphs 45-47, wherein the target molecules to be detected in the sample comprise proteins, thereby detecting proteins and nucleic acid molecules from the same sample.
49. A kit for multiplexed detection of a plurality of different target molecules from a sample comprising:
    a. a plurality of target probes, wherein each target probe in the plurality comprises:
        i. a target-binding molecule that specifically binds to a distinct target molecule in the sample;

ii. an identification nucleotide sequence that identifies the target-binding molecule; and

iii. a cleavable, non-hybridizable linker between the target-binding molecule and the identification nucleotide sequence;

b. a plurality of reporter probes, wherein each reporter probe comprises:

i. a first target probe-specific region that is capable of binding a first portion of the identification nucleotide sequence; and

ii. a detectable label that identifies the reporter probe.

50. The kit of paragraph 49, further comprising a plurality of capture probes, wherein each capture probe comprises (i) a second target probe-specific region that is capable of binding a second portion of the identification nucleotide sequence; and (ii) an affinity tag.

51. The kit of paragraph 49 or 50, wherein the detectable label of the reporter probes comprises one or more labeling molecules that create a unique signal for each reporter probe.

52. The kit of paragraph 51, wherein the unique signal is an optical signal.

53. The kit of paragraph 52, wherein the optical signal comprises a sequence of light-emitting signals.

54. The kit of any of paragraphs 51-53, wherein the one or more labeling molecules are selected from the group consisting of a fluorochrome moiety, a fluorescent moiety, a dye moiety, a chemiluminescent moiety, and any combinations thereof.

55. The kit of any of paragraphs 49-54, wherein the target-binding molecule comprises proteins, peptides, metabolites, lipids, carbohydrates, toxins, growth factors, hormones, cytokines, cells, and any combination thereof.

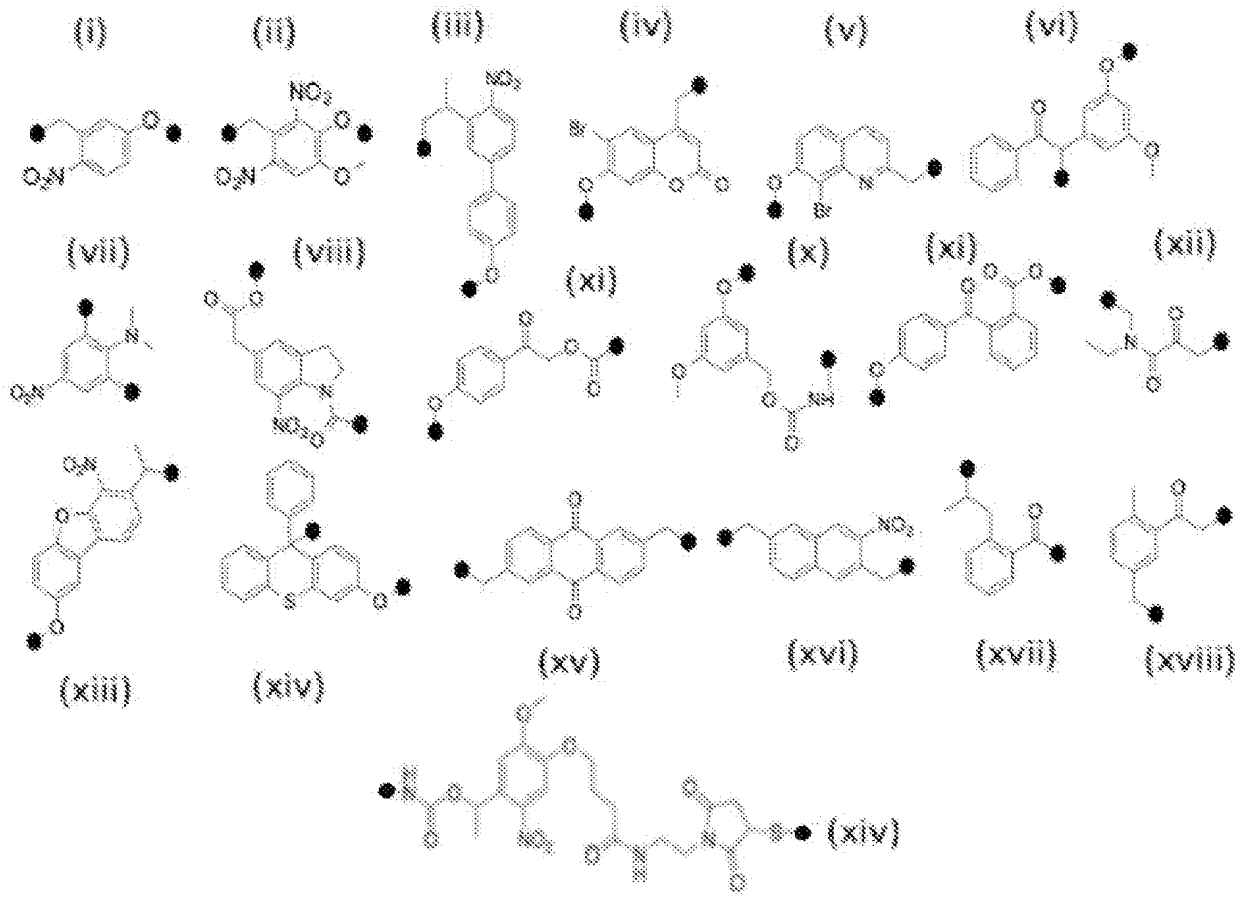
56. The kit of any of paragraphs 49-55, wherein the cleavable, non-hybridizable linker is sensitive to an enzyme, pH, temperature, light, shear stress, sonication, a chemical agent (e.g., dithiothreitol), or any combination thereof.

57. The kit of any of paragraphs 49-56, wherein the cleavable, non-hybridizable linkers are selected from the group consisting of hydrolyzable linkers, redox cleavable linkers, phosphate-based cleavable linkers, acid cleavable linkers, ester-based cleavable linkers, peptide-based cleavable linkers, photocleavable linkers, and any combinations thereof.

58. The kit of any of paragraphs 49-57, wherein the cleavable, non-hybridizable linker comprises a disulfide bond, a tetrazine-trans-cyclooctene group, a sulfhydryl group, a nitrobenzyl group, a nitoindoline group, a bromo hydroxycoumarin group, a bromo hydroxyquinoline group, a hydroxyphenacyl group, a dimethozybenzoin group, or any combinations thereof.

59. The kit of any of paragraphs 49-58, wherein the cleavable, non-hybridizable linker comprises a photocleavable linker.

60. The kit of paragraph 59, wherein the photocleavable linker is selected from the group consisting of molecules (i)-(xiv) and any combinations thereof, wherein the chemical structures of the molecules (i)-(xiv) are shown as follows:



wherein each of the black dots in each molecule represents a connecting or coupling point that connects, directly or indirectly, to the target-binding molecule or the indentification nucleotide sequence.

61. The kit of paragraph 59, wherein the photocleavable linker comprises the molecule (xiv).

62. The kit of any of paragraphs 49-61, further comprising a plurality of control probes, wherein each control probe in the plurality comprises:
    i. a control-binding molecule that specifically binds to one control molecule in the sample;
    ii. an identification control sequence that identifies the control-binding molecule; and
    iii. a cleavable linker between the control-binding molecule and the identification control sequence.

63. The kit of paragraph 62, wherein the control-binding molecule binds to a control protein.

64. The kit of paragraph 63, wherein the control protein is selected from the group consisting of housekeeping proteins, control IgG isotypes, mutant non-functional or non-binding proteins, and any combinations thereof.
65. The kit of any of paragraphs 49-64, further comprising a reagent selected from the group consisting of a hybridization reagent, a purification reagent, an immobilization reagent, an imaging agent, a cell permeabilization agent, a blocking agent, a cleaving agent for the cleavable linker, and any combinations thereof.
66. The kit of any of paragraphs 49-65, further comprising a device, wherein the device comprises a surface for immobilization of the capture probes upon coupling to the identification nucleotide sequences.

Some selected definitions

[00310] It should be understood that this invention is not limited to the particular methodology, protocols, and reagents, etc., described herein and as such may vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims.

[00311] As used herein and in the claims, the singular forms include the plural reference and vice versa unless the context clearly indicates otherwise. The term “or” is inclusive unless modified, for example, by “either.” Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities of ingredients or reaction conditions used herein should be understood as modified in all instances by the term “about.” The term “about” with respect to numerical values means within 5%.

[00312] As used herein, the term "comprising" or "comprise(s)" is used in reference to compositions, methods, and respective component(s) thereof, that are essential to the invention, yet open to the inclusion of unspecified elements, whether essential or not.

[00313] As used herein, the term "consisting essentially of" or "consist(s) essentially of" refers to those elements required for a given embodiment. The term permits the presence of additional elements that do not materially affect the basic and novel or functional characteristic(s) of that embodiment of the invention.

[00314] As used herein, the term "consisting of" or "consist(s) of" refers to compositions, methods, and respective components thereof as described herein, which are exclusive of any element not recited in that description of the embodiment.

[00315] The term “multiplexed detection” refers to detection of a plurality of target molecules from a single sample in a single assay. In some embodiments, multiplexed detection refers to simultaneous measurements of at least 2, at least 3, at least 4, at least 5, at least 10, at least

20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250, at least 300, at least 350, at least 400, at least 450, at least 500 or more different target molecules from a single sample.

[00316] As used herein, the term "fixed cell or tissue sample" refers to a sample obtained from a cell or tissue which has been previously fixed in a cell- or tissue-fixing solution and optionally afterwards embedded in a solid substrate. Various cell- or tissue-fixing solutions are known in the art, including, but not limited to aldehydes (e.g., but not limited to formaldehyde, formalin), alcohols (e.g., but not limited to, ethanol, methanol, and/or acetone), oxidizing agents (e.g., but not limited to, osmium tetroxide, potassium dichromate, chromic acid, and/or potassium permanganate), picrates, mercurial (e.g., but not limited to, B-5 and/or Zenker's fixative), Hepes-glutamic acid buffer-mediated Organic solvent Protection Effect (HOPE) fixative. In some embodiments, a fixed cell or tissue sample also encompasses a frozen cell or tissue sample.

[00317] As used herein, the term "alien or foreign DNA barcode" refers to a DNA sequence used as a barcode or tag for identification of a target molecule in a sample of an organism, wherein the DNA sequence is an alien or foreign sequence relative to the genomes of the organism from which the sample is derived or obtained. As used herein, the term "alien or foreign" refers to a nucleotide sequence that shows no or little homology against an organism (from which a sample is derived or obtained) and/or other major organisms, e.g., in the NCBI Reference Sequence (RefSeq) database. In some embodiments, a nucleotide sequence is alien or foreign when it shares a homology (sequence identity) with that of the organism by no more than 50% or less, including, e.g., no more than 40%, no more than 30%, no more than 20%, no more than 10% or less. In some embodiments, the identification nucleotide sequences described herein are alien or foreign DNA barcodes.

[00318] The term "antibody" as used herein refers to a full length antibody or immunoglobulin, IgG, IgM, IgA, IgD or IgE molecules, or a protein portion thereof that comprises only a portion of an intact antibody, generally including an antigen binding site of the intact antibody and thus retaining the ability to bind a target, such as an epitope or antigen. Examples of portions of antibodies or epitope-binding proteins encompassed by the present definition include: (i) the Fab fragment, having VL, CL, VH and CH1 domains; (ii) the Fab' fragment, which is a Fab fragment having one or more cysteine residues at the C-terminus of the CH1 domain; (iii) the Fd fragment having VH and CH1 domains; (iv) the Fd' fragment having VH and CH1 domains and one or more cysteine residues at the C terminus of the CH1 domain; (v) the Fv fragment having the VL and VH domains of a single arm of an antibody; (vi) the dAb fragment (Ward et al., 341 Nature 544 (1989)) which consists of a VH domain or a VL domain that binds antigen; (vii) isolated CDR regions or isolated CDR regions presented in a functional framework; (viii) F(ab')2 fragments which are bivalent fragments including two Fab' fragments linked by a disulphide bridge at the hinge region;

(ix) single chain antibody molecules (e.g., single chain Fv; scFv) (Bird et al., 242 Science 423 (1988); and Huston et al., 85 PNAS 5879 (1988)); (x) "diabodies" with two antigen binding sites, comprising a heavy chain variable domain (VH) connected to a light chain variable domain (VL) in the same polypeptide chain (see, e.g., EP 404,097; WO 93/11161; Hollinger et al., 90 PNAS 6444 (1993)); (xi) "linear antibodies" comprising a pair of tandem Fd segments (VH-CH1-VH-CH1) which, together with complementary light chain polypeptides, form a pair of antigen binding regions (Zapata et al., 8 Protein Eng. 1057 (1995); and U.S. Patent No. 5,641,870).

[00319] "Antibodies" include antigen-binding portions of antibodies such as epitope- or antigen-binding peptides, paratopes, functional CDRs; recombinant antibodies; chimeric antibodies; tribodies; midibodies; or antigen-binding derivatives, analogs, variants, portions, or fragments thereof.

[00320] The term "aptamer" refers to a nucleic acid molecule that is capable of binding to a target molecule, such as a polypeptide. For example, an aptamer of the invention can specifically bind to a target molecule, or to a molecule in a signaling pathway that modulates the expression and/or activity of a target molecule. The generation and therapeutic use of aptamers are well established in the art. See, e.g., U.S. Pat. No. 5,475,096.

[00321] All patents and other publications identified are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the present invention. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents is based on the information available to the applicants and does not constitute any admission as to the correctness of the dates or contents of these documents.

[00322] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as those commonly understood to one of ordinary skill in the art to which this invention pertains. Although any known methods, devices, and materials may be used in the practice or testing of the invention, the methods, devices, and materials in this regard are described herein.

EXAMPLES

Example 1. Optimization for methods for multiplex detection of target molecules from a sample

[00323] ***Evaluation of different cleavable linkers:*** In various embodiments described herein, target-binding molecules can be conjugated to identification nucleotide sequences via any cleavable linker(s) known in the art. In this Example, three alternative methods of conjugating target molecules

(e.g., antibodies) to identification nucleotide sequences (e.g., DNA) via a cleavable linker were evaluated using exemplary procedures detailed below.

[00324] In the first method, target-binding molecules (e.g., antibodies) were modified with (E)-Cyclooct-4-enyl 2,5-dioxopyrro-lidin-1-yl carbonate (trans-cyclooctene N-hydroxy-succinimidyl ester; TCO-NHS) and synthesized as previously reported in Ref. 26. If present, sodium azide was removed using a 2 ml Zeba desalting column (7 K MWCO). The reaction was performed using 1000 molar equivalents of TCO-NHS in PBS containing 10% (v/v) DMF and 10 mM sodium bicarbonate for 4 h at RT. At the same time, a photocleavable Tz-NHS was reacted with an amine group on the 5' end of the 70mer DNA strand (15 molar excess) for 4 h at RT. After the reactions concluded, the target-binding molecule-TCO (e.g., Ab-TCO) conjugate was purified using a Zeba column (7000 MWCO), and the DNA-Tz conjugate was purified using a 3 K MWCO Amicon filter followed by three washes with PBS. Next, the target-binding molecule-TCO (e.g., Ab-TCO) and Tz-DNA were combined via click chemistry (26) for two hours at RT. The final target probe (e.g., antibody-DNA conjugate) was purified by size separation using Amicon 100 K MWCO filters followed by washes with PBS.

[00325] In the second method, the photocleavable bifunctional linker (**Fig. 2**) reacted (10 molar excess) with the amine group on the 5' end of the 70mer single-stranded DNA (IDT) for 4 h at RT. Three hours after the DNA reaction began, the target-binding molecules (e.g., antibody) was reacted with 2-iminothiolane (Traut's reagent, 10 molar excess, Thermo Scientific) to convert amine groups to sulfydryl (-SH) groups in PBS with 2 mM EDTA for 1 h at RT. When the reactions concluded, the thiolated target-binding molecules (e.g., antibody) was separated from excess Traut's Reagent using a Zeba desalting column (7000 MWCO) that had been equilibrated with PBS containing 2 mM EDTA. The excess photocleavable (PC) bifunctional linker was purified from the DNA with a 3 K MWCO Amcion filter. Then the target-binding molecule-SH (e.g., antibody-SH) and the DNA-PC-linker (~15 molar excess) were reacted overnight at 4°C. The final target probe (e.g., antibody-DNA conjugate) was purified by size separation using Amicon 100 K MWCO filters followed by washes with PBS.

[00326] In the third method, an amine to sulfhydryl linker, sulfosuccinimidyl 6-[3′(2-pyridyldithio)-propionamido] hexanoate (sulfo-LC-SPDP, Thermo Scientific), was reacted with a target-binding molecule (e.g., antibody) in PBS-EDTA at 50 molar excess and aged for 1 h at RT. At the end of the reaction, excess sulfo-LC-SPDP was removed using a Zeba desalting column (7000 MWCO). The thiolyated DNA was reduced with DTT and purified via a NAP-5 column, as previously described in the DNA-antibody conjugation in the "Exemplary materials and methods" section below. Once excess sulfo-LC-SPDP was purified using a Zeba column, the target-binding molecule (e.g., antibody) was reacted with the reduced thiolyated DNA (~15 molar excess)

overnight at 4°C. The final target probe (e.g., antibody-DNA conjugate) was purified by size separation using Amicon 100 K MWCO filters followed by washes with PBS.

[00327] The three UV-cleavable target-binding molecule-DNA (e.g., Ab-DNA) linker methods were compared by first labeling A431 cells with EGFR and EPCAM DNA conjugates and then determining which method resulted in the highest signal to noise ratio (SNR), e.g., via Nanostring. The conjugation of target molecules (e.g., antibodies) with the bifunctional photocleavable linker described in **Fig. 2A** gave the highest SNR. This target probe (e.g., antibody-conjugate) was then compared to the target probe (e.g., antibody-DNA conjugate containing the DTT cleavable disulfide bond. SKOV3 cells (5×10^5 cells) were labeled with Herceptin-DNA conjugates (1 μg). After 30 minutes the cells were spun down at 400 × g for 3 minutes, and the excess Herceptin was removed with two SB+ washes. The Herceptin-DNA conjugate with the disulfide linker was then cleaved by adding DTT (50 mM) for 15 minutes at 37°C. At the same time, the Herceptin-DNA conjugate with the photocleavable linker was exposed to UV light (wavelength) for 15 minutes. After the 15-minute cleavage step, the cells were spun down at 400 × g for 5 minutes, and the supernatant was removed. The DNA in the supernatant was measured using the single-stranded Qubit assay to determine the amount of DNA cleaved from the antibody. The UV photocleavable linker had 2.4-fold more DNA than the disulfide linker.

[00328] ***Optimization of lysis conditions:*** Four different lysis conditions were evaluated to determine which was the most efficient (**Fig. 3**): (A) Proteinase K with buffer PKD (Qiagen) and UV; (B) Proteinase K with buffer ATL (Qiagen) and UV; (C) ATL buffer with UV; and (D) UV. Based on the tested methods, method (B) showed a 20% increase in signal over methods (A) and (C).

Example 2. Development and validation of methods for multiplex detection of target molecules from a sample

[00329] In this Example, an antibody barcoding with photocleavable DNA (ABCD) platform was designed to perform multiplexed protein measurements and system-wide profiling on small amounts of clinical sample material (e.g., ~100 cells). The method was designed to preserve genetic material and to enable specific isolation of rare single cells. This approach interrogates single cells by tagging antibodies of interest with short (~70-mer) DNA "barcodes"-with each antibody having a unique sequence- using a stable photocleavable linker. Photocleavable linkers known in the art (e.g., Ref. 9) can be used herein. After antibody binding to the cells, the photocleavable linker releases the unique DNA barcode, which can then be detected by various means. In some embodiments, different DNA barcodes can be identified based on size using gel electrophoresis. However, this method had limited multiplexing (8 to 12 markers) and was only semiquantitative (9). Other quantitative

methods, such as sequencing and quantitative polymerase chain reaction (qPCR), are reliable and can be used to detect the released DNA barcodes, but may introduce bias during amplification steps, require prolonged processing time, or are not cost-effective. Multiplexed qPCR only measures a maximum of five markers at a time. Thus, a fluorescence hybridization technology, which have been traditionally used for multiplexed quantitation (16,384 barcodes) of femtomolar amounts of DNA and RNA (10, 11), was selected to detect the released DNA barcodes. While the fluorescence hybridization technology has been used to quantify DNA and RNA, it had not been previously extended to measure proteins within cells or clinical samples. This Example and subsequent Examples show application and validation of the ABCD platform in cell lines and human clinical specimens, as well as evaluating drug treatment response and inter- and intrapatient heterogeneity in lung cancer.

[00330] Cells were first harvested from fine-needle aspirates (FNAs) from a given patient (**Fig. 1A**). To better isolate cancer cells from their heterogeneous cellular milieu, aspirates were labeled with antibodies directed against established markers (for example, CD45 to deplete tumor-infiltrating leukocytes from the sample). The antibody was tagged with magnetic nanoparticles and passed through a microfluidic device containing a self-assembled magnetic layer to deplete tagged cells [12]. The purified cancer cell population was retrieved from the device and stained with a mixture comprising a plurality of one or more embodiments of target probes as described herein. In this Example, the purified cancer cell population were stained with a mixture of target probes each containing an antibody and a unique barcode attached via a photocleavable linker (referred to as "antibody conjugate" or "antibody-DNA conjugate" herein) (**Fig. 1B** and **Fig. 2**). Example antibodies for use in the antibody conjugates are listed in **Table 1** below. In this Example, more than 90 antibodies in the cocktail were chosen and used to demonstrate that bulk labeling yielded similar results to single antibody labeling. The 90 antibody-DNA conjugates were specially designed to tag an alien DNA sequence that would not cross-react with the human genome. Target markers were selected to cover hallmark pathways in cancer (e.g., apoptosis, epigenetic, and DNA damage), cancer diagnostic markers known in the art, e.g., those commonly used in the clinic, and housekeeping and control proteins. Before labeling, antibody-DNA conjugates were isolated via immunoglobulin G (IgG)-specific pull-down and pooled together into a cocktail. After cell blocking, permeabilization and labeling, and washing, the DNA was released from the cells of interest with both proteolytic cleavage and photocleavage to increase yield and, by extension, sensitivity (**Fig. 1C**).

Table 1: List of example antibodies.

Antibody	Species	Catalog	Vendor
GAPDH (14C10)	Rabbit	2118BF	Cell Signaling

β-Tubulin (9F3)	Rabbit	2128BF	Cell Signaling
Ku80 (C48E7)	Rabbit	2180BF	Cell Signaling
Phospho-Chk2 (Thr68) (C13C1)	Rabbit	2197BF	Cell Signaling
S6 ribosomal protein (54D2)	Mouse	2317BF	Cell Signaling
Phospho-Chk1 (Ser345) (133D3)	Rabbit	2348BF	Cell Signaling
VE-cadherin (D87F2)	Rabbit	2500BF	Cell Signaling
p53 (7F5)	Rabbit	2527BF	Cell Signaling
Phospho-53BP1 (Ser1778)	Rabbit	2675BF	Cell Signaling
Phospho-(Ser/Thr) ATM/ATR Substrate	Rabbit	2851BF	Cell Signaling
Phospho-4E-BP1 (Thr37/46) (236B4)	Rabbit	2855BF	Cell Signaling
Bim (C34C5)	Rabbit	2933BF	Cell Signaling
Cyclin D3 (DCS22	Mouse	2936BF	Cell Signaling
Cyclin D1 (92G2)	Rabbit	2978BF	Cell Signaling
mTOR (7C10)	Rabbit	2983BF	Cell Signaling
Phospho-cyclin D1 (Thr286) (D29B3)	Rabbit	3300BF	Cell Signaling
Phospho-histone H3 (Ser10) (D2C8)	Rabbit	3377BF	Cell Signaling
ALK (D5F3)	Rabbit	3633BF	Cell Signaling
Phospho-EGF Receptor (Tyr1068) (D7A5)	Rabbit	3777BF	Cell Signaling
Phospho-Akt (Ser473) (D9E)	Rabbit	4060BF	Cell Signaling
CDCP1 Antibody	Rabbit	4115BF	Cell Signaling
Cyclin E1 (HE12)	Mouse	4129BF	Cell Signaling
Phospho-cyclin E1 (Thr62)	Rabbit	4136BF	Cell Signaling
Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (D13.14.4E)	Rabbit	4370BF	Cell Signaling
Keratin 7 (D1E4)	Rabbit	4465BF	Cell Signaling
Histone H3 (D1H2)	Rabbit	4499BF	Cell Signaling
Phospho-p38 MAPK (Thr180/Tyr182) (D3F9)	Rabbit	4511BF	Cell Signaling
Phospho-SEK1/MKK4 (Ser257) (C36C11)	Rabbit	4514BF	Cell Signaling
Pan-keratin (C11	Mouse	4545BF	Cell Signaling
Keratin 8/18 (C51)	Mouse	4546BF	Cell Signaling
Keratin 18 (DC10)	Mouse	4548BF	Cell Signaling
Akt (pan) (C67E7)	Rabbit	4691BF	Cell Signaling
p44/42 MAPK (Erk1/2) (137F5	Rabbit	4695BF	Cell Signaling
COX IV (3E11	Rabbit	4850BF	Cell Signaling
Phospho-S6 ribosomal protein (Ser235/236)	Rabbit	4858BF	Cell Signaling
53BP1	Rabbit	4937BF	Cell Signaling
β-Actin (13E5)	Rabbit	4970BF	Cell Signaling
Akt2 (L79B2)	Mouse	5239BF	Cell Signaling
Phospho-mTOR (Ser2448) (D9C2)	Rabbit	5536BF	Cell Signaling
Cleaved PARP (Asp214) (D64E10)	Rabbit	5625BF	Cell Signaling
Vimentin (D21H3)	Rabbit	5741BF	Cell Signaling

Cleaved caspase-9 (Asp330) (D2D4)	Rabbit	7237BF	Cell Signaling
Met (D1C2)	Rabbit	8198BF	Cell Signaling
FGF receptor 4 (D3B12)	Rabbit	8562BF	Cell Signaling
Axl (C89E7)	Rabbit	8661BF	Cell Signaling
p38 MAPK (D13E1)	Rabbit	8690BF	Cell Signaling
BRCA1 (D54A8)	Rabbit	9025BF	Cell Signaling
Phospho-Stat3 (Tyr705) (D3A7)	Rabbit	9145BF	Cell Signaling
Cleaved caspase-7 (Asp198)	Rabbit	9491BF	Cell Signaling
Cleaved caspase-8 (Asp391) (18C8)	Rabbit	9496BF	Cell Signaling
Cleaved caspase-9 (Asp315)	Rabbit	9505BF	Cell Signaling
PARP (46D11)	Rabbit	9532BF	Cell Signaling
4E-BP1 (53H11)	Rabbit	9644BF	Cell Signaling
Cleaved caspase-3 (Asp175)	Rabbit	9661BF	Cell Signaling
Phospho-histone H2A.X (Ser139) (20E3)	Rabbit	9718BF	Cell Signaling
FGF receptor 1 (D8E4)	Rabbit	9740BF	Cell Signaling
Caspase-8 (1C12)	Mouse	9746BF	Cell Signaling
Caspase-9	Rabbit	9502 BF	Cell Signaling
Phospho-β-Catenin (Ser675) (D2F1)	Rabbit	4176BF	Cell Signaling
Phospho-GSK-3β (Ser9) (D85E12)	Rabbit	5558BF	Cell Signaling
Dimethyl-Histone H3 (Lys9) (D85B4)	Rabbit	4658BF	Cell Signaling
Dimethyl-Histone H3 (Lys4) (C64G9)	Rabbit	9725BF	Cell Signaling
Dimethyl-Histone H3 (Lys36) (C75H12)	Rabbit	2901BF	Cell Signaling
Dimethyl-Histone H3 (Lys27)	Rabbit	9755BF	Cell Signaling
Dimethyl-Histone H3 (Lys79)	Rabbit	9757BF	Cell Signaling
Acetyl-histone H3 (Lys9) (C5B11)	Rabbit	9649BF	Cell Signaling
Acetyl-histone H3 (Lys14)	Rabbit	4318BF	Cell Signaling
Acetyl-histone H3 (Lys27)	Rabbit	4353BF	Cell Signaling
Acetyl-histone H3 (Lys56)	Rabbit	4243BF	Cell Signaling
Acetyl-histone H3 (Lys18)	Rabbit	9675BF	Cell Signaling
LC3A (D50G8)	Rabbit	4599BF	Cell Signaling
LC3B (D11)	Rabbit	3868BF	Cell Signaling
p21waf1/cip1	Rabbit	2947BF	Cell Signaling
Beclin-1 (D40C5)	Rabbit	3495BF	Cell Signaling
β-Catenin (6B3)	Rabbit	9582BF	Cell Signaling
Slug (C19G7)	Rabbit	9585BF	Cell Signaling
Snail (C15D3)	Rabbit	3897BF	Cell Signaling
TCF8/ZEB1 (D80D3)	Rabbit	3396BF	Cell Signaling
c-Myc (D84C12)	Rabbit	5605BF	Cell Signaling
Met (D1C2)	Rabbit	8198BF	Cell Signaling
Phospho-Src family (Tyr416)	Rabbit	6943BF	Cell Signaling
Phospho-Jak2 (Tyr1007)	Rabbit	4406BF	Cell Signaling

Phospho-Jak3 (Tyr980/981)	Rabbit	5031BF	Cell Signaling
Phospho-PLCγ1 (Tyr783)	Rabbit	2821BF	Cell Signaling
Bcl-2 (D55G8)	Rabbit	4223BF	Cell Signaling
Bcl-xL (54H6) rabbit mAb #2764	Rabbit	2764BF	Cell Signaling
Control mouse IgG1	Mouse	400102	Biolegend
Control mouse IgG2a	Mouse	400202	Biolegend
Control mouse IgG2b	Mouse	401202	Biolegend
Control rabbit	Rabbit	550875	BD Bioscience
Control rat IgG2b	Rat	553986	BD Bioscience
Her2	Human/Mouse	Herceptin	Genentech
EGFR	Human/Mouse	Cetuximab	Bristol-Meyers
EpCAM	Mouse	MAB9601	R&D
MUC1	Mouse	M01102909	Fitzgerald
MUC16	Mouse	ab1107	abcam
EpHA2	Mouse	MAB3035	R&D
FOLR1	Mouse	MAB5646	R&D
FSHR	Mouse	GTX71451	Genetex
TSPAN8	Mouse	MAB4734	R&D
Claudin-3	Mouse	MAB4620	R&D
Transferin	Mouse	MAB2474	R&D
CD44s	Mouse	BBA10	R&D
CD44	Mouse	103002	Biolegend
E-Cadherin	Mouse	324102	Biolegend
CEA		10-C10C	Fitzgerald
B7-H3		MAB1027	R&D
EMMPRIN	Mouse	MAB972	R&D
CD45	Mouse	304002	Biolegend
Calretinin	Mouse	sc-135853	Santa Cruz biotechnology
Ki67	Mouse	556003	BD Bioscience
Control mouse IgG	Mouse	5414BF	Cell Signaling
Control rabbit IgG	Rabbit	3900BF	Cell Signaling

[00331] The antibody-DNA conjugates were first evaluated in MDA-MB-231 (human breast cancer) cells. Cells were blocked to prevent nonspecific DNA or antibody labeling and then "stained" with the pooled cocktail following techniques akin to standard flow cytometry staining known in the art. Next, DNA was released with a light pulse, hybridized to fluorescent barcodes, and imaged on a cartridge via a charge-coupled imaging device (CCD) (NanoString Technologies).

[00332] Several DNA conjugation using various cleavable linkers and corresponding release methods were evaluated and optimized (**Figs. 2** and **3**) Among the tested cleavable linkers, the

photo-cleavable linker was selected for its superior performance (**Fig. 2**). Probe quantification translated into proteomic sample profiling (**Fig. 1C**) by normalizing according to DNA per antibody and housekeeping proteins (**Fig. 4**). On average, there were about three to five DNA fragments per antibody; markers were thresholded on the basis of nonspecific binding of IgG controls.

[00333] Repeated analyses showed consistent results across different batches of cells analyzed on different days and over time (**Fig. 5**). In subsequent studies, antibodies that did not fall above 1.2-fold control IgG threshold were not included [for example, dimethyl-histone H3 (Lys^4)]. Excluding these outliers, the median SE across all antibodies was 6%. A profile of the human MDA-MB-231 line was derived from about 50 cells and showed, for example, high expression of keratin 7 and epidermal growth factor receptor (EGFR), two diagnostic markers commonly used in pathology laboratories to identify cancer subtypes. Epigenetic and phosphoproteomic markers have lower expression because these naturally occur at lower abundance in cells relative to extracellular markers. Intracellular markers such as phospho-Src (pSrc) and phospho-glycogen synthase kinase 3β (pGSK3β) could also be detected, e.g., using the optimized permeabilization method (**Figs. 6A-6B**).

[00334] Additional benchmarking experiments were performed to demonstrate assay consistency and reproducibility. Conjugated antibodies behaved similarly to native, unmodified antibodies as evidenced by head-to-head comparison on flow cytometry (**Fig. 7A**). Similar results were found when testing intracellular antibodies such as p53 and phospho-S6 ribosomal protein (pS6RP) with dot blots and immunoblotting (**Fig. 7B**). Antibody-DNA conjugates generated equal or stronger signals compared to native antibodies on dot blots. Furthermore, the DNA-modified antibodies showed similar expression patterns across cell lysates when compared to native antibody. To assess reproducibility, two DNA-modified antibody clones specific to the same target [e.g., epithelial cell adhesion molecule (EpCAM)] were shown to give nearly identical expression levels ($R^2 = 0.99$) across multiple cell lines and clinical samples (**Fig. 8A**). Antibody staining was evaluated using both a cocktail of 60+ antibodies and as single agents; expression levels from both methods, as measured by an antibody barcoding with photocleavable DNA (ABCD) platform as described herein, showed high, linear correlation ($R^2 = 0.93$; **Fig. 8B**). Protein marker changes measured with the ABCD platform linearly correlated to expression changes measured by independent immunofluorescence studies in taxol-treated HT1080 fibrosarcoma cells (**Fig. 8C**). Flow cytometry measurements across eight cell lines and six different markers showed linear correlations ($R^2 = 0.92$ to 0.99) (**Fig. 9**).

Example 3. Single-cell sensitivity of one embodiment of the methods for multiplex detection of target molecules from a sample (antibody barcoding with photocleavable DNA (ABCD) platform)

[00335] The sensitivity of the ABCD platform was assessed by detecting across varying cell numbers (50, 15, 10 or 5 cells) from a bulk sample of 500,000 cells, in multiple repeats, by serial dilution (**Fig. 10A**). The correlations between bulk and diluted DNA counts were linear, with correlation coefficients >0.9 (**Fig. 10B**). Additional experiments were performed to validate the ABCD platform in single human A431 cells. **Figure 10C** displays the expression levels of 90 analyzed proteins for four randomly chosen single cells and in bulk samples. Consistent with previous reports (13), there were some intercellular heterogeneity, but generally, single-cell profiles matched their respective bulk profiles with correlations as high as 0.96 and as low as 0.63. Multinucleated cells were excluded; cells were otherwise selected at random.

[00336] To demonstrate biological variation at the single-cell level, untreated single human A431 cells were compared to cells treated with gefitinib- a selective tyrosine kinase inhibitor of the EGFR. Unsupervised clustering of single cells showed unique patterns for treated and untreated groups (**Fig. 11A**). A431 cell lines overexpress EGFR and are highly sensitive to gefitinib [median inhibitory concentration (IC_{50}) = 100 nM], as shown by widespread pathway inhibition in gefitinib-treated A431 cells. A threshold was applied at the single-cell level to ensure that marker expression levels were detectable above all six IgG controls for all cell lines. The majority of the panel was still detectable, although some markers such as phospho-EGFR fell below threshold levels in some cells, and thus were not included for hierarchical clustering. Nevertheless, pairwise comparisons between the two cohorts showed significant changes in key markers (**Fig. 11B**) such as pS6RP, Ku80, and phospho-histone H3 (pH3). These changes in the markers were also consistent with previous reports (14, 15). Unlike most signaling inhibition studies, the untreated cell line was not prestimulated with EGF before treatment. Hence, the assay conditions mimicked natural signaling variability to better approximate patient samples.

Example 4. Measuring inter- and intratumoral heterogeneity in clinical samples using the ABCD platform

[00337] To demonstrate the clinical capabilities of ABCD and explore single-cell heterogeneity, FNAs were obtained from patients with lung adenocarcinoma. Single-pass FNA samples were initially processed using antibody-mediated magnetic selection to isolate EpCAM-positive cells. Single cells for subsequent analyses were harvested via micromanipulation, whereas other sample debris was removed. In one representative patient, protein marker expression in 11 single cells (EpCAM+/DAPI+/CD45-) correlated with bulk measurement (about 100 remaining cells from FNAs) (**Fig. 12A**). Yet overall, correlation between patient cells and bulk FNAs was lower and varied compared to single cells from cell lines and their respective bulk in **Figs. 10A-10C**. The

highest correlation with the bulk measurement was 0.79 (cell culture showed R = 0.96), whereas the lowest value was 0.43 (**Fig. 12B**).

[00338] Interpatient heterogeneity in bulk samples was next determined from six patients with biopsy-proven lung adenocarcinoma (**Fig. 13**). Although these cancers harbored identical histopathology, proteomic profiling revealed clear differences, even in this small cohort. Marker panels were chosen to evaluate protein heterogeneity across a broad range of functional protein networks (16) relevant for therapy assessment. **Fig. 13** shows visual similarity among patients 1, 2, and 5 (Spearman $R_{1,2} = 0.94$, $R_{1,5} = 0.96$, $R_{2,5} = 0.95$). This partially concurred with genotyping because both patients 1 and 2 had EGFR T790M mutations, whereas patient 5 had a KRAS mutation (KRAS 35G>T). This indicates that different genotypes may still yield similar proteomic phenotypes. Patients 3, 4, and 6 harbored distinct proteomic profiles and differing mutations (**Fig. 13**). Patient 3 had an exon 20 EGFR mutation, whereas patient 4 had an EGFR L858R mutation and an additional BRAF mutation. Patient 6 was noted to have an EML4-ALK translocation.

[00339] Protein clustering also showed possible personalized targets (**Fig. 13**). For example, patient 4 (EGFR/BRAF mutant) had high phospho-extracellular signal-regulated kinase 1/2 (pERK1/2) and pS6RP, as expected for a patient with an EGFR L588R mutation; however, this patient also showed a high level of the DNA repair/damage markers poly(adenosine diphosphate-ribose) polymerase (PARP), Ku80, and phospho-histone H2A.X (pH2A.X) expression, indicating that PARP inhibitors or DNA-damaging agents (for example, cisplatin) could be effective for this patient. Thus, such information determined by methods for detecting a plurality of target molecules as described herein (e.g., ABCD platform) can be used to complement pharmacogenomics.

Example 5. In vitro discrimination of pathway analyses during treatment using ABCD platform

[00340] Having established feasibility of inter- and intrapatient analyses in clinical samples, it was next sought to explore the feasibility of monitoring cancer treatment over time. To this end, it was first sought to determine if known pathway responses to different drug treatments could be discriminated. **Fig. 14A** shows the validation that triple-negative breast cancer cells (MDA-MB-436) treated with kinase inhibitors (gefitinib and PKI-587), antibody drugs (cetuximab), and DNA-damaging drugs (olaparib and cisplatin) showed profiles that clustered according to drug mechanism of action. As a control study, cell lines treated with cetuximab resulted in expected drug inhibition (**Fig. 15B**). Expected protein inhibition in drug-sensitive human cancer cell lines using optimized drug doses and incubation times was demonstrated using the ABCD platform. Notable examples include pS6RP for targeted treatments, and pH2A.X, pATM/ATR (phospho-ataxia telangiectasia mutated/ATM- and Rad3-related) substrate, and cleaved PARP for DNA-damaging agents. Unexpected results, such as epigenetic histone modifications after treatment with a

phosphatidylinositol 3-kinase inhibitor (PI3Ki) was also found (**Fig. 15E**). For additional in vitro validation of treatment, HT1080 fibrosarcoma cell lines were treated with four different doses of taxol. Several panel markers displayed dose-response changes to taxol treatment, including pERK and phospho-cyclin D.

[00341] Proteomic profiling of olaparib and cisplatin treatments was performed for four human cancer cell lines, showing varying drug sensitivities as measured by viability assays (**Figs. 14A-14B** and **Fig. 15A**). The degree of change in protein profiles was quantified by calculating the number of markers that were significantly different from the untreated condition using pairwise t test [false discovery rate (FDR) = 0.1]. This profiling indicated that global pharmacodynamic changes correlated with treatment sensitivity: As IC_{50} values decreased, the number of protein markers with significant changes increased (**Fig. 14B**). For resistant cell lines (for example, OVCA429), no significant changes were detected. Expected changes in DNA damage and apoptosis markers, such as degradation of Bim and up-regulation of pERK (**Figs. 15C-15D**) were also detected, indicating previous studies of DNA damage response to cisplatin treatment (17).

[00342] To evaluate the assay's ability to measure even small marker changes, HT1080 human fibrosarcoma cells were treated with taxol at five different doses. Marker changes at high doses were compared to marker changes quantified by an independent immunofluorescence screen (**Fig. 16A**). Several protein markers showed dose-response curves, including CDCP1, phospho-cyclin D, cyclin E1, fibroblast growth factor 4 (FGF4), BRCA2, and pERK1/2. These in vitro studies established that the marker panel could indeed measure pathway changes in response to varying drug mechanisms; furthermore, these changes could be detected in a sensitive, dose-dependent manner. Additionally, pairwise t tests between the dosed and untreated cells showed an increase in significant marker changes at the highest dose (700 nM taxol) compared to the lower 70 nM dose (**Figs. 16B** and **16C**).

Example 6. Monitoring PI3Ki treatment response in cancer patients

[00343] In some embodiments, it is desirable to translate these pathway analyses to patient samples, e.g., to analyze serial biopsies in early-phase clinical trials with the goal to better assess drug efficacy and dosage. However, such invasive procedures can introduce risk of morbidity and high costs. The ability to analyze small numbers of cells from alternative sources (for example, FNAs) becomes paramount when responsive tumors shrink after treatment, making repeat biopsies difficult. As proof of concept, scant cell analyses were performed in four patients before and after PI3Ki treatment during phase 1 dose escalation trials (**Fig. 17A**). Pretreatment samples were collected the day before the first drug dose; post treatment samples were collected at the end of the second treatment cycle. Collection and processing occurred over the course of a year to correlate

profiles to patient response. All four patients had metastatic cancers of various subtypes and were selected on the basis of genetic PI3K mutations that could predispose their tumors to pathway inhibition using PI3Ki treatment. In all, two patients responded and two progressed. Data analysis was performed in a blinded manner. Unsupervised clustering separated out two groups of responders versus non-responders (**Fig. 17A**). Among the two responders, one patient showed larger fold changes across the marker panel. Subsequent unblinding revealed that this patient received a higher dose of the drug during phase 1dose escalation than did the other responding patient. Additional patient samples can be used to measure ABCD platform's clinical impact during drug dosing pathway studies.

[00344] In some embodiments, the screen performed by the ABCD platform could help predict clinical outcome or identify promising markers of treatment response. To demonstrate this, five drug-naïve patients, all with various PI3K mutations, who eventually received small-molecule PI3Ki treatment were profiled. Patients were categorized as nonresponders or responders (**Fig. 17B**) and a marker-ranking algorithm was used to determine top differential markers. The top marker, di-methylation of histone H3 at Lys79 (H3K79me2), clustered with several markers: pS6RP (a known downstream target of PI3K and an emerging key biomarker of treatment response) (14), pH2A.X, and PARP. According to canonical pathway signaling, selecting epigenetic or DNA damage markers as readouts of PI3K treatment response would not be an intuitive decision. DNA damage and epigenetic marker changes were also identified by in vitro profiling of a PI3Ki (**Figs. 15C-15E**). This cluster covered diverse proteins across various pathways: epigenetic changes, DNA damage, and growth and survival pathways [PI3K and mitogen-activated protein kinase (MAPK)], indicating the potential value of system-wide profiling for developing better companion diagnostics during treatment.

Discussion based on Examples 1-6

[00345] In some embodiments, presented herein is an amplification-free method capable of sensing hundreds of proteins in human cells by using one or more embodiments of target probes described herein (e.g., DNA-barcoded antibodies) coupled with highly sensitive optical readouts. Cell labeling, washing, and analysis can be completed within hours, making same-day protein analysis possible. The method measures more markers on limited material than immunohistochemistry and preserves genetic material from samples, which is not possible with traditional tools like multiplexed cytometry (18). The protein coverage and/or methods described herein can be extended to include additional protein targets and/or other target molecules through conjugation of target molecules to identification nucleotide sequences (e.g., antibody-DNA conjugations), resulting in a scalable, multiplexed target molecule (e.g., protein) screening platform.

[00346] In general, the method can provide analyses of protein expression levels for both single and bulk cell populations. The in vitro studies as shown in the Examples showed that single cells from cell lines showed higher correlation to bulk measurements than those isolated from patient tumors. In FNAs, the single cells also showed higher correlations with each other than with the bulk population. This could be, for example, because an averaged bulk measurement is less likely to correlate strongly with a single clonal phenotype.

[00347] The findings presented herein showed that the methods described herein can be used to detect extracellular proteins (e.g., but not limited to, CD44, EGFR), intracellular/cytosolic proteins (e.g., but not limited to, p-S6RP), and/or intracellular/nuclear proteins (e.g., 53BP1) in a sample.

[00348] The findings presented herein also showed that current cell culture models are an insufficient estimate of proteomic heterogeneity in clinical samples. The methods for detecting a plurality of target molecules from a sample described herein (e.g., ABCD platform tool) are therefore useful for its ability to study rare single cells in clinical samples, such as circulating tumor cells, stem cells, and immune cell populations. As shown herein, even scarce proteins, such as 53BP1 and pH2A.X, could be detected at the single-cell level. Large-scale protein mapping of isolated, rare cells and clonal populations could shed insight into cancer heterogeneity, drug resistance, and the clinical utility of circulating tumor cells. Intratumoral heterogeneity may itself be a biomarker of poor clinical outcome (19). Thus, the methods described herein (e.g., ABCD platform) can be used to determine intratumoral heterogeneity, which can be used as a biomarker for diagnosis and/or prognosis. Establishing causal and reactive correlations between diseases and altered biomarkers could also radically improve physicians' abilities to diagnose and treat patients (20, 21). In some embodiments, the methods described herein (e.g., ABCD platform) can be used to determine causal and reactive correlations between diseases and altered biomarkers in order to improve physicians' abilities to diagnose and/or treat patients.

[00349] The inventors have demonstrated the ABCD method's ease of use, reproducibility, compatibility with clinical applications, such as profiling of FNA cancer samples, and its translational potential to monitor cancer treatment as demonstrated in four patients. The findings showed that broader profiling can improve understanding about potentially useful companion diagnostic biomarkers and help explore how drug dosing corresponds to cellular pharmacodynamics. Smarter protein marker selection, as demonstrated by the ABCD platform, could markedly reduce drug development costs, narrow patient cohorts, and improve clinical trial design.

[00350] The methods described herein (e.g., ABCD platform) could complement other art-recognized single-cell proteomic techniques, such as mass cytometry and fluorophore-inactivated multiplexed immunofluorescence (8, 22). One of the advantages of the methods described herein

(e.g., ABCD platform) is that both genetic material and protein barcodes can be concurrently extracted from a single sample, thus paving the way for more biologically relevant analyses of protein-DNA-RNA interrelationships. Such integrative measurements could explain "missing pieces" in genomics associated with various diseases or disorders, e.g., cancer genomics. For example, in the Examples presented herein, not all patients with PIK3CA DNA mutations responded to a given PI3Ki; this is consistent with clinical experience (23, 24). However, proteomic biomarkers revealed differential changes between responding and nonresponding cohorts. The Examples indicate that protein profiling will help complement genotyping to shape therapeutic advances for cancer and other diseases.

[00351] The Examples presented herein demonstrated proof of principle that the technology described herein can work in clinical samples with a wide range of applications, including rare cell profiling and companion diagnostics within cancer clinical trials.

[00352] In some embodiments, the technology described herein (e.g., ABCD platform) can be modified to suit the needs of various applications. For example, the methods described herein (e.g., ABCD platform) can be adapted to work with both whole cells and/or cell lysates, and DNA can be quantified with other readouts (for example, sequencing) to perform simultaneous measurement of RNA, DNA, epigenetic, and protein expression. In some embodiments, the methods described herein (e.g., ABCD platform) can include a module to rapidly isolate and measure entire populations of single cells. For example, additional components and wells can be added to microfluidic devices such as the one described in the Examples to increase the throughput of single-cell analysis.

[00353] Single-cell studies can be validated with a higher-throughput device. For example, larger numbers of cells can be used to compare population differences and spreads between the methods described herein and other gold standards (for example, flow cytometry). In some embodiments, the methods described herein (e.g., ABCD platform) can be used to identify novel companion diagnostic markers or specific pathway markers for diagnosis of a disease or disorder (e.g., cancer subtypes) and/or monitoring patients' response therapeutics.

[00354] The methods described herein (e.g., ABCD platform) can enable larger-scale studies to yield mechanistic insights into existing and/or novel therapeutic strategies. Moreover, the methods described herein (e.g., ABCD platform) can also be used for rare, single-cell (for example, but not limited to circulating tumor cells) profiling to derive further understanding of their biological and clinical relevance. Because genetic material from samples is preserved, the methods described herein (e.g., ABCD platform) can be adapted to study proteins that interact with genetic regulatory elements such as microRNAs. The methods described herein (e.g., ABCD platform) can be used for

various applications in research laboratories, academic hospitals, and pharmaceutical companies to help propel drug trials and biological investigation.

Exemplary materials and methods for Examples 1-6

[00355] **Study design.** In order to determine if protein networks (as opposed to single biomarkers) will reveal clinical or biological insights into how a disease or condition (e.g., cancers) evolves and responds to drugs, a multiplexed platform for detecting protein expression, e.g., in clinical samples and in cell lines, was developed. The Examples herein demonstrate the use of the methods described herein (e.g., ABCD platform) in understanding treatment response in cancer

[00356] Clinical studies were performed on limited cohorts of patients for proof of principle. The number of patients was selected based on a 1-year enrollment cycle (March 2012 to March 2013). All protein measurements were included as long as their signals were above a pre-determined threshold. In one embodiment, the threshold was ~1.2-fold higher than that of its corresponding nonspecific IgG isotype. This threshold was set to be over three times the median SE from the antibody cohorts pooled. Only antibodies that were validated (via flow cytometry measurements on cell lines) were included. All in vitro studies were performed in replicates (n = 3, unless otherwise specified). After optimization, studies with the final protocol were repeated multiple times on different days to ensure consistency and reproducibility. All experiments on clinical studies were performed blinded during experimental procedures and raw data analysis.

[00357] **Cell lines.** Validation experiments were performed in the following cell lines, which were purchased from the American Type Culture Collection (ATCC): SKOV3, ES-2, OVCA429, UCI-107, UCI-101, TOV-112D, TOV-21G, A2780, MDA-MB-231, MDA-MB-436, A431, and HT1080. Cells were passaged in Dulbecco's modified Eagle's medium (Cellgro) or RPMI (Cellgro) as recommended by ATCC. cell lines were derived from ovarian surface epithelium (OSE) brushings cultured in 1:1 Medium 199/MCDB 105 (Sigma-Aldrich) with gentamicin (25 μg/ml) and 15% heat-inactivated serum. TIOSE6 cell lines were obtained by transfecting hTERT into NOSE cells maintained in 1:1 Medium 199/MCDB 105 with gentamicin (25 μg/ml), 15% heat-inactivated serum, and G418 (500 μg/ml) (25). After trypsinization, cells were immediately fixed with 1× Lyse/Fix buffer (BD Bioscience) for 10 min at 37°C and then washed twice with SB+ [phosphate-buffered saline (PBS) with 2% bovine serum albumin (BSA)]. The cells were aliquoted into tubes (~1 × 10^6 cells/ml) and stored at -20°C until labeling. Biological replicates were seeded in different wells and collected separately. Cultured cells were processed and stored under the exact same conditions as clinical samples. A total of 276 samples were prepared and analyzed in-dependently via the barcoding method.

[00358] **Clinical samples.** The study was approved by the Institutional Review Board at the Dana-Farber/Harvard Cancer Center, and informed consent was obtained from all subjects (n = 10).

Fourteen minimally invasive procedures were performed on the 10 enrolled patients. Six patients had primary lung adenocarcinomas. The four patients undergoing PI3Ki treatment with repeated biopsies had carcinomas of varying origins in the abdomen, all with underlying PI3K mutations. All pretreatment biopsies were collected in the week before the first cycle of treatment. All post treatment biopsies were collected after a cycle was completed, typically after several weeks to months. Image-guided FNAs with a 22-gauge needle were obtained before routine core biopsies. Correct needle location was confirmed by computed tomography imaging and real-time readout by cytopathology. FNA samples were processed immediately by centrifugation and removal of excess PBS. If there were visual clumps present before the fixation step, collagenase (Sigma-Aldrich) was added at 0.2 mg/ml. Cells were fixed with Lyse/Fix buffer (BD Biosciences) for 10 min at 37°C and washed twice with PBS with 2% BSA. All centrifugations were performed at 300g for 5 min. Clinical samples were stored at -20°C. A total of 24 samples were prepared and analyzed independently via the barcoding method.

[00359] **Drug treatments of cell lines.** To test the effect of drug treatment on protein expression levels, the cell lines were treated with a number of different chemotherapeutic or molecularly targeted drugs. A431 cell lines were dosed with gefitinib (Selleck Chemicals) in medium with 1% dimethyl sulfoxide (DMSO) for 12 hours at a concentration of 10 μM. The triple-negative human breast cancer MDA-MB-436 cell line was dosed with the PARP inhibitor olaparib (10 μM in 0.1% DMSO in medium), cisplatin (10 μM, 1% Hanks' balanced salt solution in medium), the PI3K/mTOR inhibitor PKI-587 (100 nM, 0.1% DMSO/medium), and the EGFR inhibitors cetuximab (75 μg/ml in medium) and gefitinib (10 μM in 0.1% DMSO/medium). All molecularly targeted agents (PKI-587, cetuximab, and gefitinib) were applied for 12 hours. DNA-damaging agents olaparib and cisplatin were applied to cells for 3 days. Changes in protein expression levels were compared to medium controls under identical conditions but without drug treatment.

[00360] **Flow cytometry.** Flow cytometry was used to validate protein expression levels in bulk samples. Fixed cells stored at -20°C were thawed and then permeabilized with a saponin-based buffer, PW+ (1× Perm/Wash Phosflow Buffer, BD Biosciences, with 2% BSA). About 200,000 cells per tube were incubated with primary antibodies for 1 hour at either 1 μg/ml or the appropriate dilution as recommended by Cell Signaling for flow cytometry applications. An example list of primary antibodies is shown in **Table 1** above. After one wash with PW+, the appropriate secondary antibodies targeting mouse, human, or rabbit IgG were applied. The specific secondary antibodies used were anti-rabbit IgG (H+L) $F(ab')_2$ Fragment Alexa Fluor 647 Conjugate (Cell Signaling #4414), anti-mouse IgG (H+L) $F(ab')_2$ Fragment Alexa Fluor 647 (Cell Signaling #4410), and anti-human FITC (Abcam ab98623). Expression levels for each protein were then calculated by

normalizing the geometric mean from each antibody with the appropriate control IgG. These values were then correlated to the expression values derived from the DNA barcoding technique.

[00361] **Synthesis of photocleavable DNA-antibody bifunctional linker.** The photocleavable linker was synthesized as previously described in Ref. 9. For example, compound 1 (**Fig. 2B**, ~0.100 g, 0.334 mmol) was dispersed in 5 ml of dry dichloromethane (DCM) in a round bottom flask under argon atmosphere. The flask was cooled to 0°C by placing it on an ice bath. 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) (0.139 g, 0.368 mmol) and triethylamine (TEA) (109 µl, 0.835 mmol) were added to the solution. The reaction mixture was stirred at 0°C for 5 min, and N-(2-aminoethyl)maleimide trifluoroacetate salt (0.093 mg, 0.368 mmol) was subsequently added. After stirring at 0°C for 15 min, the reaction mixture was allowed to equilibrate to room temperature while being stirred for 18 h. After the reaction mixture was diluted with DCM (45 ml), the organic phase was washed with water and saturated NaCl solution, then dried over sodium sulfate. The organic layer was concentrated under reduced pressure and charged to a SiO_2 column (eluent: 100% DCM to 3% methanol in DCM, v/v) for purification. The yield of compound 2 was approximately 60%. ^{1}H NMR (400 MHz, CD_3OD): 7.58 (s, 1H), 7.37 (s, 1H), 6.77 (s, 2H), 5.44 (q, ^{4}J = 6 Hz, 1H), 4.03 (t, ^{3}J = 6.4 Hz, 2H), 3.94 (s, 3H), 3.61 (t, 3J = 5.6 Hz, 2H), 3.35 (t, 2H, overlapping with the solvent residual peak), 2.32 (t, ^{3}J = 7.2 Hz, 2H), 2.05 (m, ^{3}H), 1.46 (d, ^{2}J = 6.4 Hz, 3H). MS (electrospray ionization mass spectrometry: ESI-MS) calculated: 421.15, found: 466.18 $\{M+HCOO\}^-$.

[00362] Compound 2 (0.010 g, 0.024 mmol) was dissolved in anhydrous dimethylformamide (DMF) (1 ml). N,N'-Disuccinimidyl carbonate (DSC;0.018 mg, 0.071 mmol) and TEA (12.5 µl, 0.096 mmol) were successively added to the solution. The reaction mixture was stirred at RT for 18 h. The reaction mixture was directly loaded onto a C18 reverse phase column for purification (eluent: 5% acetonitrile in water to 95% acetonitrile in water, v/v). The yield of the photocleavable bifunctional linker product was approximately 70%. ^{1}H NMR (400 MHz, $CDCl_3$): 7.63 (s, 1H), 7.05 (s, 1H), 6.67 (s, 2H), 6.48 (q, ^{4}J = 6.4 Hz, 1H), 6.03 (br, 1H), 4.08 (t, ^{3}J = 5.8 Hz, 2H), 4.02 (s, 3H), 3.68 (m, 2H), 3.45 (m, 2H), 2.79 (s, 4H), 2.36 (t, 3J = 7 Hz, 2H), 2.15 (m, 3H), 1.75 (d, ^{2}J = 6.4 Hz, 3H). ESI-MS calculated: 562.15, found: 607.22 $\{M+HCOO\}^-$.

[00363] **DNA-antibody conjugations.** Antibodies (e.g., listed in **Table 1**) were conjugated to specially designed alien DNA sequences derived from the potato genome (exemplary sequences shown in **Table 2**).

Table 2. List of example 70mer alien sequences used for barcoding a target-binding molecule

Target sequence	T_m capture probe (°C)	T_m reporter probe (°C)

GCTAAGTTTGGAATTAAGAAAGGAGTTGCTGGAGGTCCTTTCCAGCATAAG AACCAGCCATATTGCTTAA (SEQ ID NO: 1)	79	79
TGCCTTCTGAAAGAGACGTTATTGTTGAAGCAAGAGATAGCTTAGTAACAA ATGCTATAGCTCAGGCAGG (SEQ ID NO: 2)	80	78
CCTGATCATGCTTTGTCAGCAGACCCAGAAGAATTCATCACAATCACTGGA AGATTGAGCTTAGGAAAGT (SEQ ID NO: 3)	82	78
GAGCGGATGTTATTGAGAAGCACTTTACCTTAGATTTCTAAAGCTCTCTTCC TCCTCTCTTCTCCGCTCA (SEQ ID NO: 4)	78	82
ATCGGCTGTGCGATTGCTATTGATGTGTTAAGAAATTTGGTTTGTGATTGGC AAATCTCTCCTCCAACTC (SEQ ID NO: 5)	80	81
ATTTGGATGAAGTCGGCTTTATGGTGACACAAATCATGATGAGCTGAGGTT CTGACAGCAAATACGCTCA (SEQ ID NO: 6)	79	82
ATAGAACCATTTGCTGATGAGGTGACAACAGATCGTTGCACTTATGCTATC CCGTTAGACTATCTGCTAT (SEQ ID NO: 7)	80	78
ACTACCATGTACTGCGCGAGACTAGCCTATCATTGGATTGCAGCGATGACT ATATCTGAGCACCTGTGAC (SEQ ID NO: 8)	82	81
ATATGAGACGACTAGCACGCCATAGCGTTACATACGTGTCGATCCGAGAAC ATCACTCTAATGACGAGTG (SEQ ID NO: 9)	80	81
CATCATCGACAGTTCGCAGCCCTATAACATGATACTAGATAACGATGCTCC ATGTTAGTGAATGCGAGTC (SEQ ID NO: 10)	80	79
ACTCACACATAGTACTGACACGTAAGATAGGATGCTATATGGTCATTGGTC ACCCGAGTTACGATCAAAT (SEQ ID NO: 11)	79	79
CAGATAGACTCACCTCGATATACAGGGAGCCACGACTTAGGACTATGGATA AGTCATCTAAAGCGTCCGA (SEQ ID NO: 12)	82	78
CACTGTCTATACATGGACGACACTTTGCACATCATTACCAAAGAGCGCAAC GTATCTAGGATTGAGCAGT (SEQ ID NO: 13)	80	81
AGACTAATTGATCGGACCGATGACAGTTCACAGAGGGATACACTGTTGAGC CGACCCTATTAGCTGATAT (SEQ ID NO: 14)	80	80
TGATCCACACTGACGAATCATGTACTCACTCGATCGCCACTTCACACAAGA ACACAAATTTGGAGTATTG (SEQ ID NO: 15)	80	79
CTCGAGAATCACACACAGTCGTCTAAGACACGACAAGTGCAACAGCAATC CACATCTTAGATGAGATTAG (SEQ ID NO: 16)	81	78
CGATTACAAGGCGTGGTCAGATATTAGACTCCAGGGGATTTAATGCCAGTC CAAGCTCTCTTCCACATTC (SEQ ID NO: 17)	81	81
ATCTGCATGAACGGGAAAGGAGTTCGATGAGACTTTCAAACCAACATAATG TCTCTCCAACCTCAGGAAG (SEQ ID NO: 18)	82	80
ATAGTCTTTAGAGCCTCAGAATAGGCTGTGACGCGGAAGATAACTCATAAG TGCCTCCCTCGGTAATTTG (SEQ ID NO: 19)	82	79
GCCAGGTATGCCGTGAACGAGTTCTTCATTAACTGTTATGTCTCGGGAGTCT GATATTGGTACTTCTCCC (SEQ ID NO: 20)	82	80
TTAGCACCGATATCAATACTGATGATGTCACCGTCGAGCTCGTGTTGAACC CTTCAAGTAACAACCTGAC (SEQ ID NO: 21)	79	82
ACTTGTTCGACTGACAGTTTAACGCCTGACATGAACGGCTTGCTTATAATGA CTGGCAGGGTTATGAATG (SEQ ID NO: 22)	81	81
AAACTGACCGTACCGTTAGAAGAGAGTTCCGCTTCTCTCATGATGTGCGCA TCTCCCACATTATTTGACC (SEQ ID NO: 23)	82	81
TGATGACAGTGACAATTGACCGAATTGCCTGATCATTACCTTACAGTGCGC AGATTGGGATAATCGATTT (SEQ ID NO: 24)	81	79
TAGGCGTTGAGGCTTTGTTTCTTTGCCTCTATTGTAAGACTCATTCTGACGG CCTCTAGTCGTTGATATG (SEQ ID NO: 25)	81	80
AAGGACATTCTTTCGAATGCAAGTTCAAGGCACATTTTCTATATCAGCCAC CATGGGAGTGACATTTCTT (SEQ ID NO: 26)	80	79

CAATAGCTCCAGTAGTAATTGTTGTCGCTCCGCTGAGCAGTTAATCCTTATG TCAACAACCTCAGCATAG (SEQ ID NO: 27)	82	78
TTCACCAAGCTGAACAGGGTTGCGCTGAATAAATTTTACAGGATACTATGG ACAGGTTCAGAATCCTCGA (SEQ ID NO: 28)	82	79
GGAATGAATCCATTGCATTTCCATGAGAATGCAGACTTAATCGGACGTATC GACTTTGGGTCCACGATAT (SEQ ID NO: 29)	79	80
GAGGTCTTGTTTCATCTAAACCGAGCAGGATGATAAGCCATAATTCGTAAC CCGAGGGTATAATTCGTTA (SEQ ID NO: 30)	79	79
GTCCTTCTGCTTATGACATTCCGTGCATTCCGTAGCTACGTCAAGCGTTACA TAGTGACGGAACTGTTAG (SEQ ID NO: 31)	82	80
TCTGTACCTTGGCACTCCATCTGGTAAGTCACTTATAGTTGTATGGTTTCAG ATGAGGGAACGTGTAGGA (SEQ ID NO: 32)	81	80
AATTTCTGAGATTGTTGGTAGAGGGAGAAATGGGAAGGACATGTTTCAACA ATCACCGGATTAAAGCCTT (SEQ ID NO: 33)	79	80
TGTGGAAGGACTGTGATAAACCAATAGGGTGTCAAGATCTGTAAGTATGGG ATTAGGGATGTTCTGCCAG (SEQ ID NO: 34)	80	80
GCCGTCGGACATAACCACTTGGATATATACGTAGTTCATCAACCTTAACTC CCTCTGGGTTCATTGGGAG (SEQ ID NO: 35)	80	82
GCTATTGCAGCAAAGAGAACAGACGCTTTAACTGGTATCGAGCGCTTAGAT GGCTATATGGTCTACTAGA (SEQ ID NO: 36)	81	78
GAAATCAGATCAGTTCTACATTCGGTGGGAGCCCTCTATATGATTAGATCCT GCAGCCGTACTTCCGTCA (SEQ ID NO: 37)	82	80
GGTGGCTTGATTTAACTGAATCAGGCCCTAACCATTTGTATTGTGTCTACAC TGGTCCGTTCTTAGACGC (SEQ ID NO: 38)	82	81
GTTGTTTACCTTGTAGATCGACTTCACATCAGCGGCAGAAGGCCCTCAACG TAAATCTGCTCCACATTTA (SEQ ID NO: 39)	80	81
TGTTGACATCCGCAACAATGTACCTTATATCGGCATATGGATCTCTTGATCG AGCGAACCTCCCTTTAAC (SEQ ID NO: 40)	81	80
AAGGTGATTCACTAACCAGCTCTTACTCCTCGTTCGGTAGCAAATGAAATG CCGGATGCTGTTGAAGTAG (SEQ ID NO: 41)	80	81
CGCATAACTCGAACCACAGTTACTATCAGTCGACATCCCACCAGAGAAATT GAAGGATATTGTTGAAGCA (SEQ ID NO: 42)	80	79
GAATCTTGGAAGGTTTCCAGTTAAATAGGGCGTGCGAAGATTCCAGGCAGA TTTCTCAGGAATTCAGTCA (SEQ ID NO: 43)	81	80
CTGCTAATGCTGATGGCCCACCTTCTCTATTTGTCGCCATTATATGCGTTGA GGTTAGTTCAAGCAATAC (SEQ ID NO: 44)	82	78
GAACAGCTTTCCTTGCTCCCTCTAAATCACCATTTCCATTAGATGAAACCGA CTTCATTCCAGACTCAAT (SEQ ID NO: 45)	80	78
AATGCATTTGCCAATGTAGCCATTGTATAACCAGATACACTAGTCCAATGT CTCAACCAGGGATACCACA (SEQ ID NO: 46)	79	81
CTCAGAGCTTCAAATCTATCCTCTGGAATCTCTGTATAAGCCCTCGAATACA ACTTGAGGTATCCCGCAT (SEQ ID NO: 47)	79	81
CTCTTCTGCCCTACATCACTATCGACTATAGCAACATATCTTTCTCGGGTAA AGATTAGGCGTCCGATAT (SEQ ID NO: 48)	79	78
GTAACCGTAGTCGCGCAAACCGTTATATTACGGATATGATCCAAGTTATAT ACATTAGGACGCGGTTGCT (SEQ ID NO: 49)	81	79
ATGGTTAGTAAACAGCTTTGATTTCTACATCCGCCTAGCAAACCCATAGTTC TGCAGTAGATTCACAGCG (SEQ ID NO: 50)	79	81
TTCAGTTATAATGTGTCCAGCAGAAGCAGGAATTGAATTACCCAAGTTGCA AGTGGAAGATTTGGAGTTA (SEQ ID NO: 51)	79	78
TTGCAGAAGCATTCCCAATATGGGTTTCAAGAGTTTAAAGAATGTGGAACA TTCATGGGAACTGGTGAAG (SEQ ID NO: 52)	80	79

GCAACAACCTCATCTATACTGTGAATAGTCCCTCCGCTGTCTATATTGGAAC TGCTGCAATGGTTGCTCT (SEQ ID NO: 53)	80	82
CCGCAGATTATCGTTTACGATGCATCCATGGTCTCCGACCCATTGAGAGAG CCAATGGAATTAAGAACTT (SEQ ID NO: 54)	82	80
CACCATTCAGCCTGATATTGCGTTTGGTGTTGATGTGGCAACTGCATACTGA ATAACTCCCTGAAATAGC (SEQ ID NO: 55)	81	80
CGTTACATACTCAGCCATAGGCTTCGATAACAGCATTATTGGAACCTCTGG GACATTAACAGAGACAACA (SEQ ID NO: 56)	81	79
AGCGTACTAGGCATCTATTGGCTGAACTACCATGTAATTAGTGGTGTTCCA GCCTCTAAGATGATGTGGT (SEQ ID NO: 57)	81	80
GATAGGATGCGACTGCGTATCATATAGGCTGCACATTAGCTGTTGCTTCAA ATGCCAATCTTACCTCAAC (SEQ ID NO: 58)	82	79
AATGTATGAGCGGACACTATGCTAAGAGAGACTCCATCAATCCCTCTATGC AAGATAACAACATCTGGCT (SEQ ID NO: 59)	80	79
TGCACATCATAGTGCGACGTTGATCCAGATAGACTATAAGACGGCTTGGCA TTTACCCTAGTCACTATCT (SEQ ID NO: 60)	81	80
AATGTGTCAGCGGCCTAACTGTAATTGATCCACACCTTAGTTCGGGAGCTA CCGATCTAATCAACCGTTT (SEQ ID NO: 61)	82	80
AGACTCCAGGTCGATCATTGGATAACCAACCAGTCGGTTATCCATGACGAG TGAATAATCTTACCGCAGG (SEQ ID NO: 62)	82	80
TTTAGATCCTAAGAATGCGAAATGCCGATTCCCGCATATTTCGTAAGCTCGT TCGGGACTTTGTATCGGC (SEQ ID NO: 63)	82	81
GAGTGATAGGATCACTCTAAGATCGGCCACTATACGACGCTGAGGTTTATA TGAACGGCCGCAATTATGA (SEQ ID NO: 64)	79	81
TCTTGACCAACACCATGTCCGACATACTCCCTAACATGGGTACGGCGACTA CTGAATCGTTCTTTGAGAG (SEQ ID NO: 65)	82	82
TGTGTAAATGAAAGCATCTGACTCAACAGGCATCAGTAACGATAATGAGTA CAACGCCCAATGGTCATAG (SEQ ID NO: 66)	80	79
GCTTCAACGATTTCAATATACCCATTCGTCAGAGGAAGTAGTAGATCCCGC CGTCTTAGTCGGATTGAAA (SEQ ID NO: 67)	79	81
TGTGGTTCCGGTTGCGTATAGATCATGATTCTTTACCCACCTCTTGCTGTAA TGACCACAATCAACGTAG (SEQ ID NO: 68)	79	82
GTATCGGCGAACACGAAATCCTCTACTCTTGACAAACTCCCATTCCTACCTC TCCAAAGTTAGAGGAGAT (SEQ ID NO: 69)	81	80
TTGCATTACAATGGCCGATCAAGATAAGGACATTCATAATGGAGCTATAGA ATACAACACCAACGTCGCA (SEQ ID NO: 70)	79	79
TAATTCTTCCTTGATTCCGTGATTGGATGTCCCTCAGGAGTAGTAGTGTGGA TGTTGTTGTTAGACACTT (SEQ ID NO: 71)	79	78
TGGAGGGTCGTAACCGCTATAGATGTGATTCACTCCAACAACTTCCCTATCT TTAATCCTCTCACTCCAC (SEQ ID NO: 72)	81	78
TGAATAAATTCGTTGGCGCTGTAGAGATCGGAGTTCCGGATTCGTACTACT CGTTTACGGGATTTACAGA (SEQ ID NO: 73)	80	80
GCTAAAGGAGACTCCGGTTTAAACGTCATCGCAATCTTTGATGGGCAAGCG AGCACATAGATATGCGTTA (SEQ ID NO: 74)	81	82
AATATTCTCCGGCATGAATGGCGTGGGAATGAATCCGGCTTTGTGTTTATTG TACATAGACGTTGTCCCG (SEQ ID NO: 75)	82	81
GAGAACGAGCGGAGCAAGATAGCCTTTAACTGAATCGTCGTCTTATTCCCA GTACACATCATTCCAAATG (SEQ ID NO: 76)	81	79
ATATTCTGTACTCAGTGCCTATCCACCTAATAGGGACCTCAGCGACCTGTCC GTTACATTAATGAAACAT (SEQ ID NO: 77)	78	81
CATTCCGTAGAATTACTACACCGCGGGATCATTATAACGTCGAAGAGCTTC AGAGGTAAGTGAAACAAGG (SEQ ID NO: 78)	79	81

CCCGAAGGCATAATCAACATCCATTGTACATCCCTTGTTATAGCTCCAGGG CCAGAGATTAAAGGAATAG (SEQ ID NO: 79)	81	79
CTAGGATGTAACTTGCGTTAGTTGCAGATTCGCTATATTGCTTAAGCTCTGA GCTCCATGTCCAGTAATT (SEQ ID NO: 80)	79	79
TTCTCGCAGTTGTAAACTTATAGTGTCGCGCCTAGAAATTCATAGCCACAA ATTCTCTTTGGGCAGAGAT (SEQ ID NO: 81)	81	78
TATAGTTACCAAGTACTATGGGTTGGTGGAAGCCGAACGTCTGTCCAAATG GAGCTATAGTTAAGAGGGA (SEQ ID NO: 82)	80	80
AGACGCACACCGATAGAGGAGAGATCTTACATACCTGCTAAGGTTGTTAAT GGCATTGCAGATAGCTTAG (SEQ ID NO: 83)	81	79
CCAGAAAGGTACAGGGCCAATTAACACGTAATCGGCCTCCAACTCTGCCAT CTTTAAGCATTCTAAAGCT (SEQ ID NO: 84)	82	80
AATTCTCCGTCATGTGGTCGTCTGATGCCTAACTTTATCTGCTATCAATGTA GAGGATCGTGCATTACCG (SEQ ID NO: 85)	82	79
CGCGGGCTAAGTAGTAGGGTTCTAATGCTACTTTAAATACGCTCACAATCC AGGCTATATCGCTGTAGCT (SEQ ID NO: 86)	80	81
TAATCACTGTATTTGTTAATCATGGCTAGGCGGGTCCAATAGGGAAACTGA TACTAACGTAGGAGCACGC (SEQ ID NO: 87)	79	81
GTATTCTGGAGAACCTCGTGGCAATGGCAATTCTCCACGAGTGCTAAGATC TGAGCCGTTTACCAAAGAG (SEQ ID NO: 88)	82	81
ATAACCTGGTCTCCGGTTGATCGTTTACCTGAAACATGAGATTAGCAACGA CCCAAACATGCCACTTCAC (SEQ ID NO: 89)	81	82
CACAACATGCAGCAGGCAAGTAGGGTTTCTGATTATAAGCATCCAGCAATA AAGCCTCCTTCAAACCAAC (SEQ ID NO: 90)	81	81
CCCTAACCATGTTCTACGAGCGGTCACAGATTATATTCAACTACAAGTGTA AATGTACGAGCGCCGAGAT (SEQ ID NO: 91)	80	80
GAAAGGCATTTGACGGGAGCATTGACGAAGACATACGGTAATTTGTCGTCG CACGGACAATTAGTGAGTT (SEQ ID NO: 92)	82	81
TAATACTGGGTCACAAGATTAGATTCCAGCTGTGACGGCGATGAAGTCCGC GAGGATATGTTTCTATATC (SEQ ID NO: 93)	78	81
GGTTCATTGTCTCATCGTACGGCTAATGTAGATACGAGGTAGCCGAGTATG ACACACCACAGCAGTTAAT (SEQ ID NO: 94)	78	82
TTATGGATTCCGATGATCCTCCGCGTGGTACAAATGTTACCTTGATGCAATA GTCTCTGTATGCGATCGG (SEQ ID NO: 95)	82	80
AGCGGTACTAATATGCTATGAGCGAGTTCCCTAACGAGAGATAACGACCCT CTGTCGTAAGCACTTAAGG (SEQ ID NO: 96)	80	81
GAGGCATCTCTGCTAACTATATGCTGAACAGCTTTTCCACGATATAGGTAC ATTGGACGCTTACAGGATA (SEQ ID NO: 97)	80	79
TTTCGGCCCAACTTATATGCTCTCCGAATCTTGGAGCAGTCATCGTAACCTG ATAGCAATCTACGTCAAG (SEQ ID NO: 98)	82	80
ACTGCAGTGAGGGCAACCAATACAAATTAAATCTGCCTCCTATTGGGATAC CTCCCGTCCATTAAGTTAG (SEQ ID NO: 99)	79	80
TTGGAGAAACAACCATACAGGTGTCTTTAACTACCTGGAACTCTACCAATT GGAGCTTTCTTAGCTGTCT (SEQ ID NO: 100)	78	80
GCTATCAACTTCCCTATCCAAACCGTTGGATGAATTGAAAGCATAGATGTT CCTTGGAGAGGTTTCCCAG (SEQ ID NO: 101)	80	81
TGAGGAGTAAGTATACGACGCCTGCACTAGTCACTTGCTGGCTTTGAGCCA ATAGATGTGTTAATGGCTA (SEQ ID NO: 102)	82	81
CACAGCCAATCTCTTAGGACAGTACATGGTTAGTAACGTCTGTGGAAGTCA TGAGCACACGATCTGTAAG (SEQ ID NO: 103)	79	82
TGAGTATCTACAGGTGTTCTCATGGGATCGTAGTTGGTCTGTCCAACATGAC GTTATAGGCATAACTCCA (SEQ ID NO: 104)	79	81

TACCTTAAACTGCGCTGGTAACTTGGATCGTGTAGTCATTGGGAGCAAACC ATCTGTCTTTCGTATGGAG (SEQ ID NO: 105)	81	81
GTTAGGTTCAGCCTCATTCCCTAAGAATCCAACTCATAACTCAATCATGCGC GTCCAGCAAAGACAAATG (SEQ ID NO: 106)	80	81
ACTGTCTAATACAACCGGATTCTAAGACCACATGGTCTTAGACGCGCGTGC AATTCTGAACTATATGATT (SEQ ID NO: 107)	79	79
TGGCTATTGCCGCAGTAGATCAAAGATTGAGAGAGATATAGATTACTCCAT GATACACCCAAGCCTCGAC (SEQ ID NO: 108)	81	79
GCAACAAGTGATGCTGACGCAGTTGTTATAGATGGCCTTTGGCTCACGCTA ATTGAGTTACTGTAGGAAA (SEQ ID NO: 109)	81	80
GCTATCTCACCAGCTCCTCACCATGACATTTACTCTCCACATTTATCTGCGA CCTGTTTCGTAAACGATG (SEQ ID NO: 110)	81	80

[00364] The 70-mer sequence length was selected for optimal hybridization with the NanoString capture and reporter probes. Other sizes were evaluated as well. Shortening sequence length tended to improve signal but reduce hybridization capability. For example, although 50-mer sequences gave relatively higher signals when compared to controls, 30-mer sequences did not reliably hybridize. Thus, 70-mer sequences were selected for reliable hybridization. However, sequences that are longer or shorter than 70 nucleotides can also be used in the methods described herein.

[00365] Antibodies (e.g., listed in **Table 1**) can be purchased from commercial sources, and were initially purified from BSA and/or other contaminants with either a Zeba spin column or centrifugal filter. Antibodies were then incubated with photocleavable bifunctional linker in PBS (containing 5% N,N′-dimethylformamide and 10% 0.1 M $NaHCO_3$) at room temperature for 1.5 hours. Afterward, excess reagents were removed from maleimide-activated antibodies with a Zeba spin column [7000 molecular weight cutoff (MWCO), eluent: PBS].

[00366] Thiol-modified DNA oligos (from Integrated DNA Technologies) were reduced with dithiothreitol (DTT; 100 mM) in PBS (1 mM EDTA, pH 8.0) for 2 hours at room temperature. The reduced DNA oligos were then purified with NAP-5 column (GE Healthcare), with deionized water as the eluent. The fractions containing DTT (determined with the microBCA assay) were discarded. The remaining reduced DNA fractions were pooled and concentrated with a 3000 MWCO Amicon filter (Millipore).

[00367] The maleimide-activated antibodies were incubated with the reduced DNA oligos in PBS solution. In a typical conjugation process, 15-fold molar excess of DNA oligos was incubated with maleimide-activated antibodies. The conjugation reaction was allowed to proceed for 12 hours at 4°C. DNA barcode-antibody conjugates were purified with a Millipore 100K MWCO centrifugal filter followed by three washes with PBS. After the antibodies were mixed, a final purification of excess DNA was conducted with protein A/G-coated magnetic beads (Pierce/Thermo Scientific). The commercial protocol from Thermo for magnetic separation was only slightly modified to use a

tris-buffered saline (TBS)/0.1% Tween wash buffer and a Gentle Ag/Ab Elution Buffer (Thermo Scientific). Three elutions were performed for 20 min each. Solvent antibody was exchanged into pure TBS with a Zeba desalting column (7000 MWCO).

[00368] Antibody characterization. Antibodies were aliquoted and stored at concentrations of 0.25 mg/ml in PBS with BSA (0.15 mg/ml) at -20°C, with adequate usage for at least twelve experimental runs (the number of runs on each NanoString cartridge) to avoid freeze-thaw cycles. Various other storage methods were tested, including glycerol or 4°C storage, but aliquoting and freezing showed the most consistent, high-fidelity storage for up to 9 months. Antibody concentrations were determined via microBCA assay (Thermo Scientific). DNA concentrations were also independently determined using the Qubit ssDNA kit (Invitrogen) to quantify the relative number of DNA per antibody. To achieve relative DNA/Ab measurements with higher sensitivity across the cohort of antibodies, in some embodiments, the NanoString platform was used to add antibody cocktails under two conditions: (1) "Control": antibodies were added in their native forms with DNA still attached, and (2) "Released DNA": antibodies were treated with proteinase K and photocleaved. Under the control condition, the DNA was still attached to the antibody and thus could not simultaneously bind to the NanoString assay's reporter and capture probe. The difference in DNA readings between these two measurements thus revealed the relative number of DNA per antibody. This difference was divided by the isotype control measurement to account for possible inherent experimental error in protein concentration and/or antibody isolation (see **Fig. 6** for relative number of DNA:Ab ratio). Antibodies were rigorously tested and validated prior to use. Of 110 antibodies, 88 were selected for the final panel and all had been previously validated from specific vendors (primarily Cell Signaling Technologies, BioLegend; **Table 1**). Antibodies that did not work with DNA conjugates did not work in their native state either and were excluded; DNA conjugated antibodies worked as well as the parent antibody **(Figs. 7A-7B).**

[00369] Antibody staining and DNA collection for protein profiling. Prior to cell staining, antibodies were pooled into a cocktail with TBS, 0.1% Tween, and 0.2 mg/ml cysteine (to avoid DNA cross-reaction with other antibodies). Tubes were coated with serum blocking buffer overnight to prevent samples from non-specifically binding to tube walls. Cells were then incubated for a minimum of one hour with a blocking buffer at 37°C: 10% v/v Rabbit serum (Jackson Immuno Research Labs, 011-000-120), 2% BSA, 1 mg/ml SS salmon sperm DNA (Sigma Aldrich, D7656), 0.2 mg/ml cysteine (Sigma Aldrich), 20X Perm (BD Bioscience) or 0.1% Tween 20 (Sigma Aldrich)—all in PBS to minimize non-specific antibody or DNA binding. The antibody cocktail was then added to the fixed and permeabilized cells and incubated for one hour at RT with intermittent mixing.

[00370] After incubation, the cells were washed with PW+ with 0.05 mg/ml of DS sheared salmon sperm DNA (Life Technology, AM9680). Either two 15-ml washes in 15-ml tubes or four 1.5-ml washes in 1.7-ml microcentrifuge tubes were performed. Blocking and wash steps were desired for achieving low background even with femtomolar detection. All washes were performed on ice. Labeled cells could then be counted and selected for lysis/proteinase K/photocleavage to release the DNA. Lysis buffer was used on 10 μl of cells (with up to 50,000 cells), 34.2 μl of ATL lysis buffer (Qiagen) and 5.8 μl of Proteinase K (Qiagen). This reaction proceeded at 56°C for a minimum of 30 min. Photocleavage was then performed using long UV wavelength light (model) for 15 min. This resulted in a cell-lysis mix with released DNA. Samples were spun down at 14,000 × g for 10 min. Supernatant was collected, and serial dilutions were performed in nuclease-free water (Invitrogen, AM9937) to collect DNA equivalent to 50-100 cells to avoid saturating the read-out cartridge (Nanostring). This amount resulted in cartridge binding densities within the linear range of quantitation. Binding densities in the lower range (0.05- 0.2) were still linear and gave consistent protein profiles comparable to those in the higher range (1.5-2.5). At lower binding densities (for example single cells), the majority of markers could be measured, with the exception of low expression markers with weaker antibodies (pJAK2, pChk2).

[00371] **Immunofluorescence.** Immunofluorescence provided an independent measure and validated marker changes from paclitaxel (Taxol) treatment (**Figs. 16A-16B**). HT1080 cells were seeded at 4,000 cells per well in 96 well plates (Grenier), which were compatible with high resolution plates, and grown for 24h in DMEM media before either being treated with Paclitaxel at 100 nM or kept in control media. After 24 h, cells were fixed for 15 min at RT, then gently washed on a rocker with PBS/0.1% Tween for 5 min, repeated 3 times. All subsequent washes were also performed with this buffer, time duration, and repetition protocol.

[00372] Cells were then permeabilized with ice cold 90% methanol for 20 min. After washing, cells were blocked for 1 h at room temperature with blocking buffer (Odyssey). Primary antibodies (all from Cell Signaling; see **Table 1**) were then added in blocking buffer at prescribed dilutions, sealed, covered in foil and incubated overnight. The next day, after washing, anti-rabbit-FITC secondary antibodies, 1:500 Hoechst and 1:200 whole-cell stain blue (Cellomics) were added (all primary antibodies were rabbit IgGs) at 2 μg/ml and incubated for 2 h. Final wash steps were performed in PBS only, and the cells were subsequently imaged at 20X using an Olympus microscope (BX63) with a Delta Vision chamber and software. All images were taken in biological triplicate. Fluorescence intensity for each cell was determined using CellProfiler, which used Hoechst and whole-cell stain to delineate cell boundaries and size constraints to discount debris. Additional in-house MATLAB (Mathworks) code was then used to calculate marker signals for each condition and calculate the changes between treated and untreated cells.

[00373] Immunoblotting and dot blotting. OVCAR3, SKOV3, CAOV3, A2780, and OVCAR429 cell lines were plated in 6-well dishes and grown for 72 h prior to lysis for Western blot analysis. Cells were washed twice with ice-cold PBS, scraped into 200 µl per dish of radio immunoprecipitation assay buffer (RIPA buffer) (Cell Signaling Technology), containing HALT protease and phosphatase inhibitor cocktail (Pierce), and transferred to 2 ml microcentrifuge tubes. Lysates was passed through a 23-g syringe 5 times, and then incubated 5 min on ice with vortexing every min. Lysates were centrifuged 15 min at 14,000 × g (4°C). Supernatant was transferred to a new microcentrifuge tube and total protein was measured using the BCA assay. Equal total protein was prepared, boiled, and loaded on a Novex NuPAGE 4-12% Bis-Tris gel and then transferred to nitrocellulose.

[00374] Membranes were blocked for 1 h at room temperature in SuperBlock T20 (TBS) buffer (Pierce) and then washed briefly in tris-buffered saline (TBS) with 0.1% Tween-20 (TBST). Membranes were then incubated overnight at 4°C with rocking in p53 (1:1000, Cell Signaling), DNA-conjugated p53 (1:1000), pS6RP (1:1000, Cell Signaling), or DNA-conjugated pS6RP (1:1000) primary antibodies diluted in TBST with 10% SuperBlock. Membranes were washed three times, 5 min each in TBST and then incubated 1 h at room temperature in goat α-rabbit HRP conjugated secondary antibody diluted 1:1000 in TBST with 10% SuperBlock. Following washing, signal was detected using SuperSignal West Pico chemiluminescent substrate (Pierce). For the Ki67 antibodies, 1ug cell lysates (prior to denaturing) from above were loaded onto nitrocellulose a Bio-Dot microfiltration apparatus (Bio-Rad). Blots were then processed as above, using a Ki67 or DNA-conjugated Ki67 antibody (1:1000, BD Biosciences) and an α-mouse HRP conjugated secondary antibody diluted as above. Dot blots were detected as above.

[00375] Single-cell isolation and processing. After antibody staining, single cells were picked with a micromanipulator. Cells were stained with Hoechst 3342 (Molecular Probes), added to an open 10-cm dish, and imaged with a TE2000 microscope (Nikon). Single cells were placed directly into a PCR tube. Five microliters of lysis buffer/proteinase K was added (4.5 µl of ATL buffer and 0.5 µl of proteinase K). Lysis/enzymatic cleavage proceeded for 30 min at 56°C before photocleavage for 15 min. Reporter and capture probes (NanoString Technologies) were then directly added to this tube according to the manufacturer's recommendations.

[00376] Data analysis: Calculating proteomic expression profiles. Protein expression profiles were extracted from raw data as follows. First, raw DNA counts were normalized via the mean of the internal NanoString positive controls, which account for hybridization efficiency. These counts were then converted to antibody expression values using the relative DNA/antibody counts. Next, average background signal from control IgG was subtracted. Last, housekeeping genes were used for normalization that accounted for cell number variations. Signals were normalized via a

house-keeping protein, e.g., β-tubulin. For the taxol treatments, signals were normalized via the geometric mean of histone H3, GAPDH, and actin rather than tubulin, because tubulin is a primary target of taxol. Data were transformed into log2 scale as denoted in captions.

[00377] **Data analysis: Clustering.** Heat maps and clustergrams were plotted using MATLAB with a matrix input of marker expression values that were calculated as detailed above. All shown clustergrams were performed as a weighted linkage and were clustered using correlation values as a distance metric. Some clustergrams were normalized by row, as specified in captions, to highlight marker differences among different patients. If a marker was not detectable in one of the patients, it was removed from the matrix or heat map and is not displayed.

[00378] **Statistical analysis.** Raw data from NanoString DNA counts were normalized by first using the nSolver analysis software to account for hybridization differences on the cartridge. Only positive controls A to D on the NanoString software were used in normalization. DNA counts were within the linear range of detection and met all other criteria for inclusion as determined by the nSolver software (maximum fields of view, image quality, etc.). After determining an expression value by taking into account nonspecific IgG binding and housekeeping genes (cell count), data were log2-transformed.

[00379] Correlation between single-cell analysis and bulk measurement was calculated in GraphPad Prism. Spearman r values were calculated without assuming a normal, consistent distribution. Two-sided P values were calculated, where significant markers were identified by comparing two groups (for example, treated versus untreated) in Prism and performing pairwise t tests with an FDR of 0.2 for multiple test correction error. Significant marker changes and their P values between gefitinib-treated and untreated A431 single cells are shown in **Table 3** below. For heat maps, if any samples had markers below threshold, the entire marker row was removed (no imputed data values were used). To identify differentiating markers between responders and nonresponders, a multiclass sequential forward selection-ranking algorithm was used. The patients were classified as responders or nonresponders based on known data. Class separability was measured by the Bhattacharya distance.

[00380] **Table 3: Significant markers between A431 single cells with or without gefitinib treatment.** Six markers out of 49 markers showed significant difference between gefitinib-treated vs. untreated A431 single cells and the average expression values as calculated via Nanostring profiling for each cohort. Marker significance was determined by pairwise t-testing and corrected for multiple testing errors by using a false discovery rate of 0.2.

Proteins	P	Untreated	Treated
Phospho-S6RP	0.0067212	1171.3	58.4
Phospho-histone H3	0.0091305	4920.6	982.0

Ku80	0.0098001	770.2	120.6
FGFR4	0.0106319	914.9	114.1
CD56	0.0117795	1906.5	334.4
Dimethyl-histone H3 (Lys36)	0.0119939	695.7	86.9

References

1. M. Basik, A. Aguilar-Mahecha, C. Rousseau, Z. Diaz, S. Tejpar, A. Spatz, C. M. Greenwood, G. Batist, Biopsies: Next-generation biospecimens for tailoring therapy. Nat. Rev. Clin. Oncol. 10, 437-450 (2013).

2. B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz Jr., K. W. Kinzler, Cancer genome landscapes. Science 339, 1546-1558 (2013).

3. J. G. Paez, P. A. Jänne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman, T. J. Boggon, K. Naoki, H. Sasaki, V. Fujii, M. J. Eck, W. R. Sellers, B. E. Johnson, M. Meyerson, EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. Science 304, 1497-1500 (2004).

4. M. Soda, V. L. Choi, M. Enomoto, S. Takada, V. Vamashita, S. Ishikawa, S. Fujiwara, H. Watanabe, K. Kurashina, H. Hatanaka, M. Bando, S. Ohno, V. Ishikawa, H. Aburatani, T. Niki, V. Sohara, V. Sugiyama, H. Mano, Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 448, 561-566 (2007).

5. M. B. Vaffe, The scientific drunk and the lamppost: Massive sequencing efforts in cancer discovery and treatment. Sci. Signal. 6, pe13 (2013).

6. E. D. Hsi, A practical approach for evaluating new antibodies in the clinical immunohisto-chemistry laboratory. Arch. Pathol. Lab. Med. 125, 289-294 (2001).

7. E. J. Lanni, S. S. Rubakhin, J. V. Sweedler, Mass spectrometry imaging and profiling of single cells. J. Proteomics 75, 5036-5051 (2012).

8. S. C. Bendall, E. F. Simonds, P. Qiu, A. D. Amir el, P. O. Krutzik, R. Finck, R. V. Bruggner, R. Melamed, A. Trejo, O. I. Ornatsky, R. S. Balderas, S. K. Plevritis, K. Sachs, D. Pe'er, S. D. Tanner, G. P. Nolan, Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science 332, 687-696 (2011).

9. S. S. Agasti, M. Liong, V. M. Peterson, H. Lee, R. Weissleder, Photocleavable DNA barcode-antibody conjugates allow sensitive and multiplexed protein analysis in single cells. J. Am. Chem. Soc. 134, 18499-18502 (2012).

10. P. Fortina, S. Surrey, Digital mRNA profiling. Nat. Biotechnol. 26, 293-294 (2008).

11. G. K. Geiss, R. E. Bumgarner, B. Birditt, T. Dahl, N. Dowidar, D. L. Dunaway, H. P. Fell, S. Ferree, R. D. George, T. Grogan, J. J. James, M. Maysuria, J. D. Mitton, P. Oliveri, J. L. Osborn, T.

Peng, A. L. Ratcliffe, P. J. Webster, E. H. Davidson, L. Hood, K. Dimitrov, Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat. Biotechnol. 26, 317-325 (2008).

12. J. Chung, D. Issadore, A. Ullal, K. Lee, R. Weissleder, H. Lee, Rare cell isolation and profiling on a hybrid magnetic/size-sorting chip. Biomicrofluidics 7, 054107 (2013).

13. S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, P. K. Sorger, Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature 459, 428-432 (2009).

14. J. N. Andersen, S. Sathyanarayanan, A. Di Bacco, A. Chi, T. Zhang, A. H. Chen, B. Dolinski, M. Kraus, B. Roberts, W. Arthur, R. A. Klinghoffer, D. Gargano, L. Li, I. Feldman, B. Lynch, J. Rush, R. C. Hendrickson, P. Blume-Jensen, C. P. Paweletz, Pathway-based identification of biomarkers for targeted therapeutics: Personalized oncology with PI3K pathway inhibitors. Sci. Transl. Med. 2, 43ra55 (2010).

15. S. Shintani, C. Li, M. Mihara, N. Terakado, J. Vano, K. Nakashiro, H. Hamakawa, Enhancement of tumor radioresponse by combined treatment with gefitinib (Iressa, ZD1839), an epidermal growth factor receptor tyrosine kinase inhibitor, is accompanied by inhibition of DNA damage repair and cell growth in oral cancer. Int. J. Cancer 107, 1030-1037 (2003).

16. D. Hanahan, R. A. Weinberg, Hallmarks of cancer: The next generation. Cell 144, 646-674 (2011).

17. J. Wang, J. V. Zhou, G. S. Wu, Bim protein degradation contributes to cisplatin resistance. J. Biol. Chem. 286, 22384-22392 (2011).

18. C. Benoist, N. Hacohen, Immunology. Flow cytometry, amped up. Science 332, 677-678 (2011).

19. N. McGranahan, R. A. Burrell, D. Endesfelder, M. R. Novelli, C. Swanton, Cancer chromosomal instability: Therapeutic and diagnostic challenges. EMBO Rep. 13, 528-538 (2012).

20. J. Bousquet, J. M. Anto, P. J. Sterk, I. M. Adcock, K. F. Chung, J. Roca, A. Agusti, C. Brightling, A. Cambon-Thomsen, A. Cesario, S. Abdelhak, S. E. Antonarakis, A. Avignon, A. Ballabio, E. Baraldi, A. Baranov, T. Bieber, J. Bockaert, S. Brahmachari, C. Brambilla, J. Bringer, M. Dauzat, I. Ernberg, L. Fabbri, P. Froguel, D. Galas, T. Gojobori, P. Hunter, C. Jorgensen, F. Kauffmann, P. Kourilsky, M. L. Kowalski, D. Lancet, C. L. Pen, J. Mallet, B. Mayosi, J. Mercier, A. Metspalu, J. H. Nadeau, G. Ninot, D. Noble, M. Oztürk, S. Palkonen, C. Préfaut, K. Rabe, E. Renard, R. G. Roberts, B. Samolinski, H. J. Schünemann, H. U. Simon, M. B. Soares, G. Superti-Furga, J. Tegner, S. Verjovski-Almeida, P. Wellstead, O. Wolkenhauer, E. Wouters, R. Balling, A. J. Brookes, D. Charron, C. Pison, Z. Chen, L. Hood, C. Auffray, Systems medicine and integrated care to combat chronic noncommunicable diseases. Genome Med. 3, 43 (2011).

21. E. E. Schadt, Molecular networks as sensors and drivers of common human diseases. Nature 461, 218-223 (2009).

22. M. J. Gerdes, C. J. Sevinsky, A. Sood, S. Adak, M. O. Bello, A. Bordwell, A. Can, A. Corwin, S. Dinn, R. J. Filkins, D. Hollman, V. Kamath, S. Kaanumalle, K. Kenny, M. Larsen, M. Lazare, Q. Li, C. Lowes, C. C. McCulloch, E. McDonough, M. C. Montalto, Z. Pang, J. Rittscher, A. Santamaria-Pang, B. D. Sarachan, M. L. Seel, A. Seppo, K. Shaikh, Y. Sui, J. Zhang, F. Ginty, Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. Proc. Natl. Acad. Sci. U.S.A. 110, 11982-11987 (2013).

23. F. Janku, J. J. Wheler, S. N. Westin, S. L. Moulder, A. Naing, A. M. Tsimberidou, S. Fu, G. S. Falchook, D. S. Hong, I. Garrido-Laguna, R. Luthra, J. J. Lee, K. H. Lu, R. Kurzrock, PI3K/AKT/mTOR inhibitors in patients with breast and gynecologic malignancies harboring PIK3CA mutations. J. Clin. Oncol. 30, 777-782 (2012).

24. M. Elkabets, S. Vora, D. Juric, N. Morse, M. Mino-Kenudson, T. Muranen, J. Tao, A. B. Campos, J. Rodon, Y. H. Ibrahim, V. Serra, V. Rodrik-Outmezguine, S. Hazra, S. Singh, P. Kim, C. Quadt, M. Liu, A. Huang, N. Rosen, J. A. Engelman, M. Scaltriti, J. Baselga, mTORC1 inhibition is required for sensitivity to PI3K p110a inhibitors in PIK3CA-mutant breast cancer. Sci. Transl. Med. 5, 196ra99 (2013).

25. K. K. Zorn, A. A. Jazaeri, C. S. Awtrey, G. J. Gardner, S. C. Mok, J. Boyd, M. J. Birrer, Choice of normal ovarian control influences determination of differentially expressed genes in ovarian cancer expression profiling studies. Clin. Cancer Res. 9, 4811-4818 (2003).

26. N. K. Devaraj, R. Upadhyay, J. B. Haun, S. A. Hilderbrand, R. Weissleder, Fast and sensitive pretargeted labeling of cancer cells through a tetrazine/trans-cyclooctene cycloaddition. Angew. Chem. Int. Ed. Engl. 48, 7013-7016 (2009).

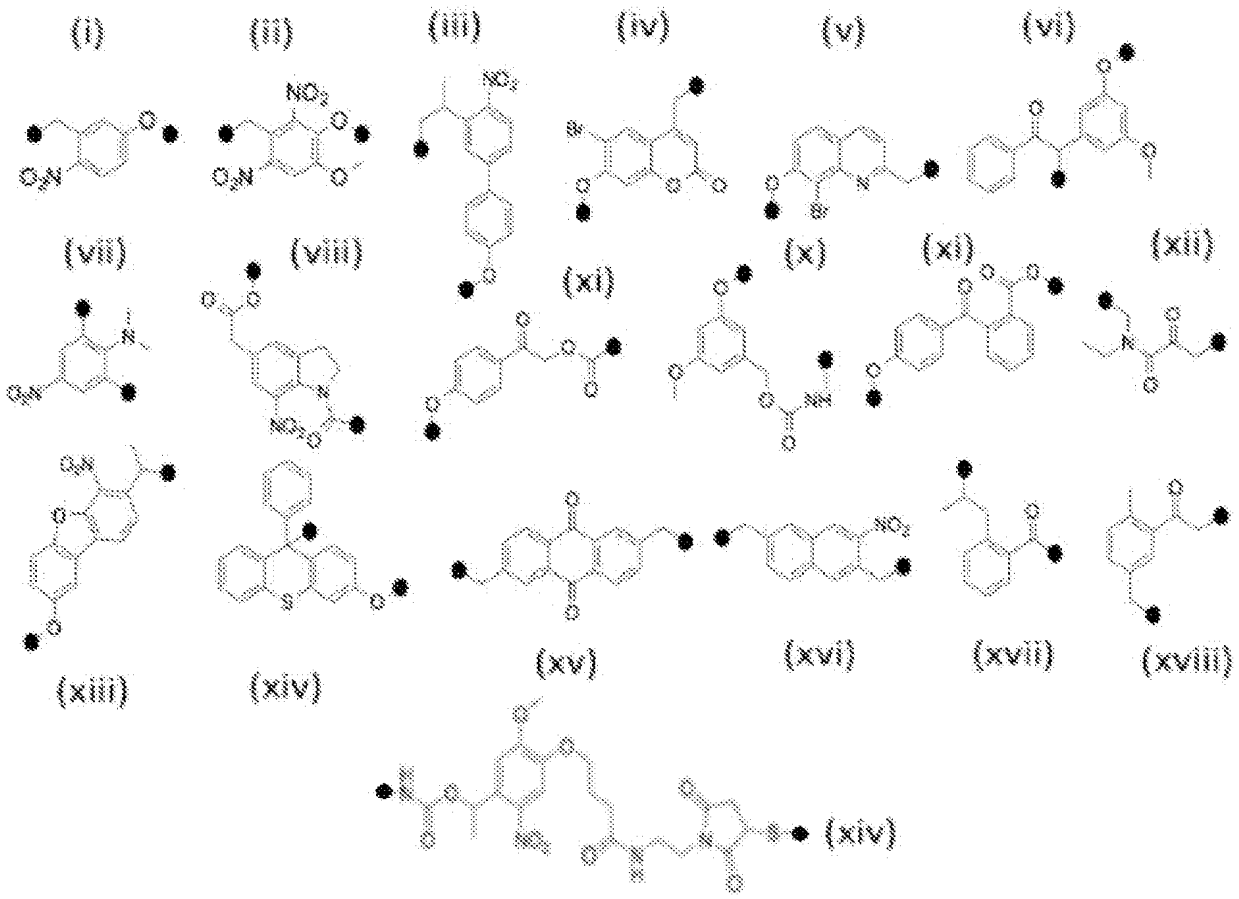
CLAIMS

What is claimed is:

1. A method for detecting a plurality of target molecules in a sample comprising:
    a. contacting a sample with a composition comprising a plurality of target probes, wherein each target probe in the plurality comprises:
        i. a target-binding molecule that specifically binds to a distinct target molecule in the sample;
        ii. an identification nucleotide sequence that identifies the target-binding molecule; and
        iii. a cleavable linker between the target-binding molecule and the identification nucleotide sequence;
    b. separating unbound target probes from a plurality of complexes in the sample, each complex having a target molecule and a single target probe bound thereto, wherein the complex does not have a second target probe binding to a different region of the target molecule;
    c. releasing the identification nucleotide sequences from the plurality of complexes; and
    d. detecting signals from the released identification nucleotide sequences based on a non-gel electrophoresis method, wherein the signals are distinguishable for the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of different target molecules in the sample.
2. The method of claim 1, wherein the non-gel electrophoresis method comprises sequencing, quantitative polymerase chain reaction (PCR), multiplexed (PCR), mass cytometry, fluorophore-inactivated multiplexed immunofluorescence, hybridization-based methods, fluorescence hybridization-based methods, or any combinations thereof.
3. The method of claim 1 or 2, further comprising, prior to the detecting step (d), coupling the released identification nucleotide sequences from the releasing step (c) to a detection composition comprising a plurality of reporter probes, wherein each reporter probe in the plurality comprises: a first target probe-specific region that is capable of binding a first portion of the identification nucleotide sequence; and a detectable label that identifies the reporter probe.

4. The method of claim 3, wherein the detecting comprises detecting signals from the respective detectable labels of the reporter probes that are coupled to the released identification nucleotide sequences, wherein the signals are distinguishable for the respective reporter probes and bound the identification nucleotide sequences, thereby identifying the corresponding target-binding molecules and detecting a plurality of target molecules in the sample.
5. The method of claim 3 or 4, wherein the detection composition further comprises a plurality of capture probes, wherein each capture probe comprises (i) a second target probe-specific region that is capable of binding a second portion of the identification nucleotide sequence; and (ii) an affinity tag.
6. The method of claim 5, wherein the affinity tag of the capture probe permits immobilization of the released identification nucleotide sequences onto a solid substrate, upon coupling to the detection composition.
7. The method of any of claims 3-6, wherein the detectable label of the reporter probes comprises one or more labeling molecules that create a unique signal for each reporter probe.
8. The method of claim 7, wherein the unique signal is an optical signal.
9. The method of claim 8, wherein the optical signal comprises a sequence of light-emitting signals.
10. The method of any of claims 7-9, wherein the one or more labeling molecules are selected from the group consisting of a fluorochrome moiety, a fluorescent moiety, a dye moiety, a chemiluminescent moiety, and any combinations thereof.
11. The method of any of claims 1-10, wherein the detecting step (d) comprises no amplification of the released identification nucleotide sequences.
12. The method of any of claims 3-11, wherein the detecting step (d) comprises no amplification of the first target probe-specific region, or the second target probe-specific region.
13. The method of any of claims 1-12, wherein the identification nucleotide sequences are selected such that they do not cross-react with a human genome.
14. The method of claim 13, wherein the identification nucleotide sequences are derived from a potato genome.
15. The method of any of claims 1-14, wherein the identification nucleotide sequences have a length of about 30-100 nucleotides.
16. The method of any of claims 1-15, wherein the identification nucleotide sequences have a length of about 70 nucleotides.
17. The method of claim 16, wherein the identification nucleotide sequences have a sequence selected from Table 2 (SEQ ID NO: 1 to SEQ ID NO: 110).

18. The method of any of claims 1-17, wherein the cleavable linker is a cleavable, non-hybridizable linker.
19. The method of claim 18, wherein the cleavable, non-hybridizable linker is sensitive to an enzyme, pH, temperature, light, shear stress, sonication, a chemical agent (e.g., dithiothreitol), or any combination thereof.
20. The method of claim 18 or 19, wherein the cleavable, non-hybridizable linkers are selected from the group consisting of hydrolyzable linkers, redox cleavable linkers, phosphate-based cleavable linkers, acid cleavable linkers, ester-based cleavable linkers, peptide-based cleavable linkers, photocleavable linkers, and any combinations thereof.
21. The method of any of claims 18-20, wherein the cleavable, non-hybridizable linker comprises a disulfide bond, a tetrazine-trans-cyclooctene group, a sulfhydryl group, a nitrobenzyl group, a nitoindoline group, a bromo hydroxycoumarin group, a bromo hydroxyquinoline group, a hydroxyphenacyl group, a dimethozybenzoin group, or any combinations thereof.
22. The method of any of claims 18-21, wherein the cleavable, non-hybridizable linker comprises a photocleavable linker.
23. The method of claim 22, wherein the photocleavable linker is selected from the group consisting of molecules (i)-(xiv) and any combinations thereof, wherein the chemical structures of the molecules (i)-(xiv) are shown as follows:



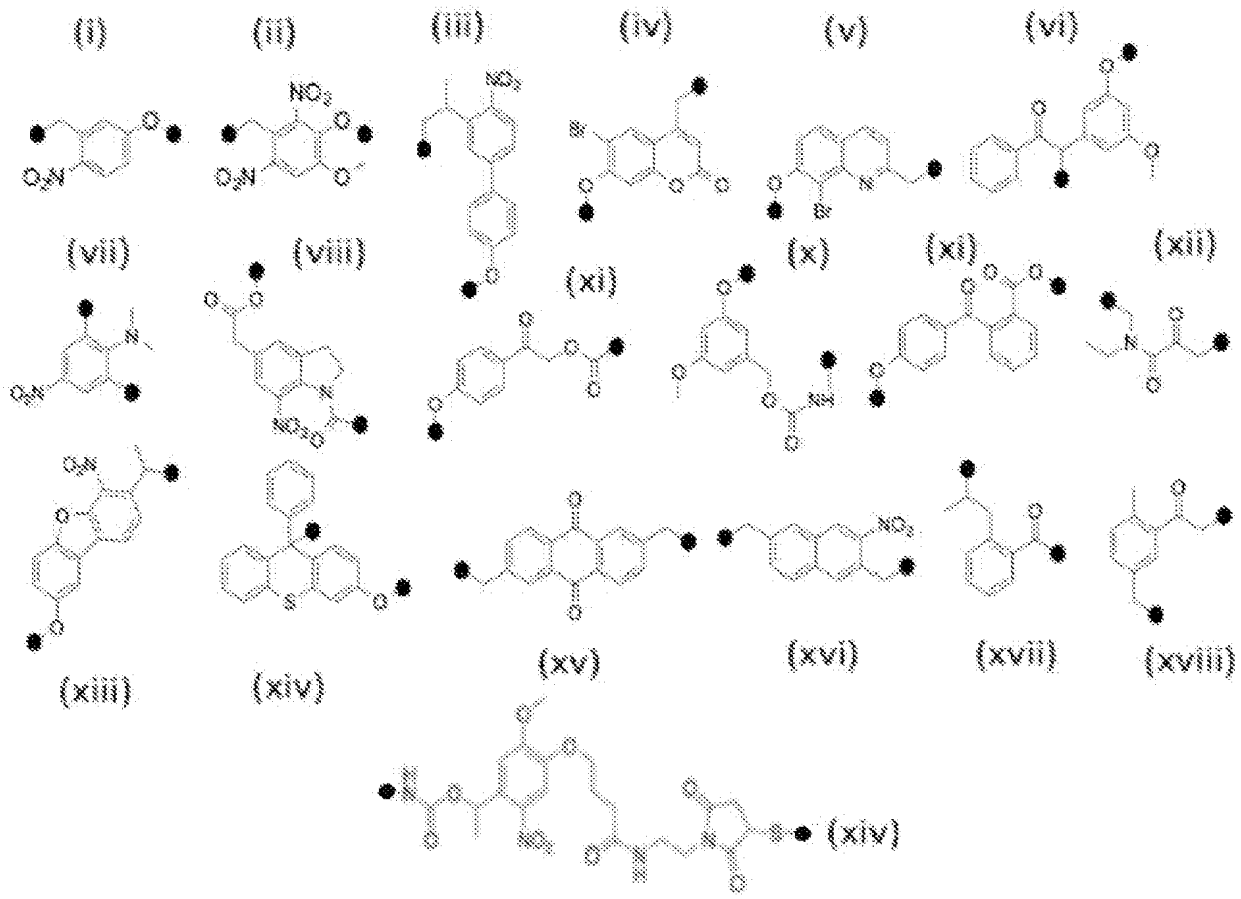
wherein each of the black dots in each molecule represents a connecting or coupling point that connects, directly or indirectly, to the target-binding molecule or the indentification nucleotide sequence.

24. The method of claim 22, wherein the photocleavable linker comprises the molecule (xiv).
25. The method of any of claims 22-24, wherein the releasing of the identification nucleotide sequences from the bound target probes comprises exposing the bound target probes to ultraviolet light.
26. The method of any of claims 1-25, wherein the sample comprises less than 500 cells.
27. The method of any of claims 1-26, wherein the sample is a single-cell sample.
28. The method of any of claims 1-27, wherein the sample comprises cells isolated from a fine-needle aspirate.
29. The method of any of claims 1-28, further comprising, prior to the contacting, separating target cells from interfering cells in the sample.
30. The method of claim 29, wherein the separating comprises labeling the interfering cells or target cells with magnetic particles and separating them from the sample by magnetic separation.

31. The method of claim 30, wherein the magnetic separation is performed in a microfluidic device.
32. The method of any of claims 29-31, wherein the target cells comprise rare cells.
33. The method of claim 32, wherein the rare cells are selected from the group consisting of circulating tumor cells, fetal cells, stem cells, immune cells, clonal cells, and any combination thereof.
34. The method of any of claims 29-33, wherein the target cells comprise tumor cells from a liquid biopsy (e.g., peritoneal, pleural, cerebrospinal fluid, and/or blood), a mucosal swap, a skin biopsy, a stool sample, or any combinations thereof.
35. The method of any of claims 1-34, wherein the target molecules comprise proteins, peptides, metabolites, lipids, carbohydrates, toxins, growth factors, hormones, cytokines, cells, and any combinations thereof.
36. The method of any of claims 1-35, further comprising permeabilizing the target cells in the sample.
37. The method of any of claims 1-36, wherein the composition further comprises a plurality of control probes, wherein each control probe in the plurality comprises:
    i. a control-binding molecule that specifically binds to one control molecule in the sample;
    ii. an identification control sequence that identifies the control-binding molecule; and
    iii. a cleavable linker between the control-binding molecule and the identification control sequence.
38. The method of claim 37, wherein the control-binding molecule binds to a control protein.
39. The method of claim 38, wherein the control protein is selected from the group consisting of housekeeping proteins, control IgG isotypes, mutant non-functional or non-binding proteins, and any combinations thereof.
40. The method of any of claims 1-39, further comprising thresholding the signals.
41. The method of claim 40, wherein the signals are thresholded on the basis of nonspecific binding.
42. The method of claim 41, wherein the threshold is greater than that of the signals from the non-specific binding.
43. The method of any of claims 40-42, wherein the threshold is determined by using standard deviation and measurement error from at least one control protein.
44. The method of any of claims 1-43, further comprising quantifying the signals by normalizing the signals associated with the target probes by the signals associated with the control probes.
45. The method of any of claims 1-44, further comprising extracting a nucleic acid molecule from the same sample for nucleic acid analysis.

46. The method of claim 45, further comprising subjecting the nucleic acid molecule to a nucleic acid analysis.
47. The method of claim 46, wherein the nucleic acid analysis comprises sequencing, quantitative polymerase chain reaction (PCR), multiplexed PCR, DNA sequencing, RNA sequencing, de novo sequencing, next-generation sequencing such as massively parallel signature sequencing (MPSS), polony sequencing, pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing, ion semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, single molecule real time (SMRT) sequencing, nanopore DNA sequencing, sequencing by hybridization, sequencing with mass spectrometry, microfluidic Sanger sequencing, microscopy-based sequencing techniques, RNA polymerase (RNAP) sequencing, or any combinations thereof.
48. The method of any of claims 45-47, wherein the target molecules to be detected in the sample comprise proteins, thereby detecting proteins and nucleic acid molecules from the same sample.
49. A kit for multiplexed detection of a plurality of different target molecules from a sample comprising:
    a. a plurality of target probes, wherein each target probe in the plurality comprises:
        i. a target-binding molecule that specifically binds to a distinct target molecule in the sample;
        ii. an identification nucleotide sequence that identifies the target-binding molecule; and
        iii. a cleavable, non-hybridizable linker between the target-binding molecule and the identification nucleotide sequence;
    b. a plurality of reporter probes, wherein each reporter probe comprises:
        i. a first target probe-specific region that is capable of binding a first portion of the identification nucleotide sequence; and
        ii. a detectable label that identifies the reporter probe.
50. The kit of claim 49, further comprising a plurality of capture probes, wherein each capture probe comprises (i) a second target probe-specific region that is capable of binding a second portion of the identification nucleotide sequence; and (ii) an affinity tag.
51. The kit of claim 49 or 50, wherein the detectable label of the reporter probes comprises one or more labeling molecules that create a unique signal for each reporter probe.
52. The kit of claim 51, wherein the unique signal is an optical signal.
53. The kit of claim 52, wherein the optical signal comprises a sequence of light-emitting signals.

54. The kit of any of claims 51-53, wherein the one or more labeling molecules are selected from the group consisting of a fluorochrome moiety, a fluorescent moiety, a dye moiety, a chemiluminescent moiety, and any combinations thereof.
55. The kit of any of claims 49-54, wherein the target-binding molecule comprises proteins, peptides, metabolites, lipids, carbohydrates, toxins, growth factors, hormones, cytokines, cells, and any combination thereof.
56. The kit of any of claims 49-55, wherein the cleavable, non-hybridizable linker is sensitive to an enzyme, pH, temperature, light, shear stress, sonication, a chemical agent (e.g., dithiothreitol), or any combination thereof.
57. The kit of any of claims 49-56, wherein the cleavable, non-hybridizable linkers are selected from the group consisting of hydrolyzable linkers, redox cleavable linkers, phosphate-based cleavable linkers, acid cleavable linkers, ester-based cleavable linkers, peptide-based cleavable linkers, photocleavable linkers, and any combinations thereof.
58. The kit of any of claims 49-57, wherein the cleavable, non-hybridizable linker comprises a disulfide bond, a tetrazine-trans-cyclooctene group, a sulfhydryl group, a nitrobenzyl group, a nitoindoline group, a bromo hydroxycoumarin group, a bromo hydroxyquinoline group, a hydroxyphenacyl group, a dimethozybenzoin group, or any combinations thereof.
59. The kit of any of claims 49-58, wherein the cleavable, non-hybridizable linker comprises a photocleavable linker.
60. The kit of claim 59, wherein the photocleavable linker is selected from the group consisting of molecules (i)-(xiv) and any combinations thereof, wherein the chemical structures of the molecules (i)-(xiv) are shown as follows:



wherein each of the black dots in each molecule represents a connecting or coupling point that connects, directly or indirectly, to the target-binding molecule or the indentification nucleotide sequence.

61. The kit of claim 59, wherein the photocleavable linker comprises the molecule (xiv).
62. The kit of any of claims 49-61, further comprising a plurality of control probes, wherein each control probe in the plurality comprises:
    i. a control-binding molecule that specifically binds to one control molecule in the sample;
    ii. an identification control sequence that identifies the control-binding molecule; and
    iii. a cleavable linker between the control-binding molecule and the identification control sequence.
63. The kit of claim 62, wherein the control-binding molecule binds to a control protein.
64. The kit of claim 63, wherein the control protein is selected from the group consisting of housekeeping proteins, control IgG isotypes, mutant non-functional or non-binding proteins, and any combinations thereof.
65. The kit of any of claims 49-64, further comprising a reagent selected from the group consisting of a hybridization reagent, a purification reagent, an immobilization reagent, an imaging agent,

a cell permeabilization agent, a blocking agent, a cleaving agent for the cleavable linker, and any combinations thereof.

66. The kit of any of claims 49-65, further comprising a device, wherein the device comprises a surface for immobilization of the capture probes upon coupling to the identification nucleotide sequences.

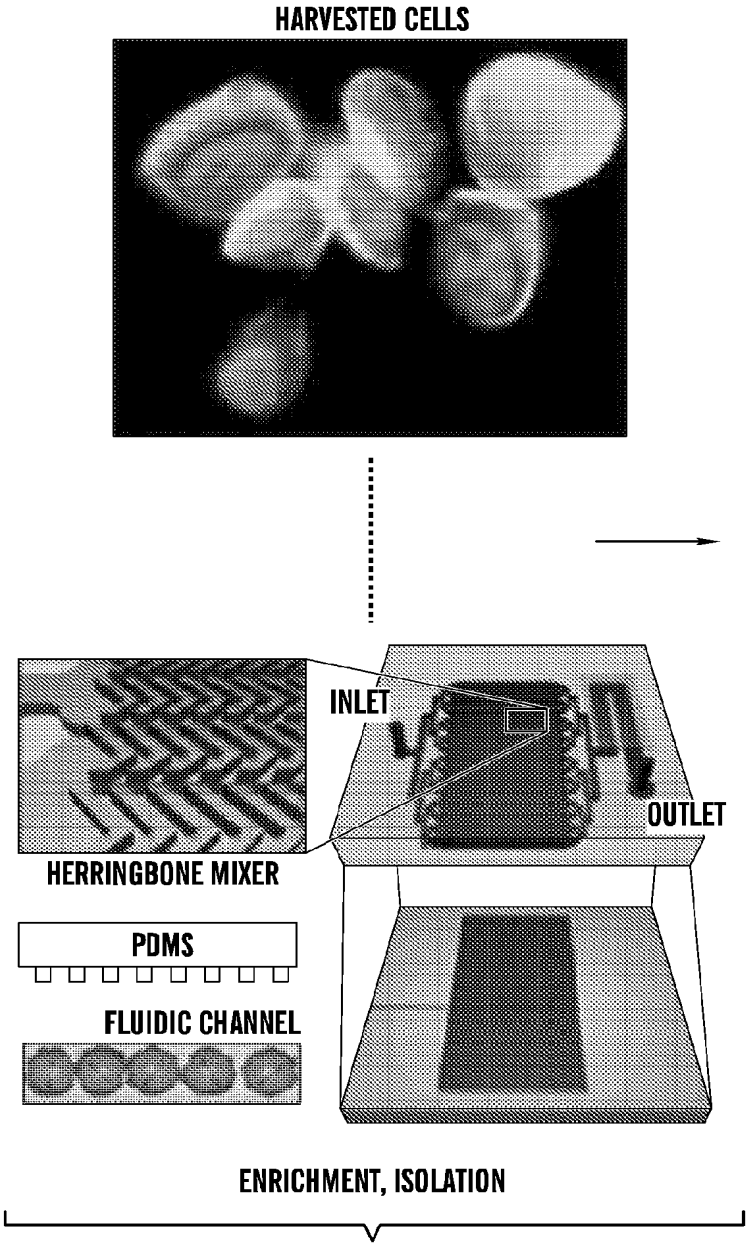

HARVESTED CELLS
INLET
OUTLET
HERRINGBONE MIXER
PDMS
FLUIDIC CHANNEL
ENRICHMENT, ISOLATION

FIG. 1A

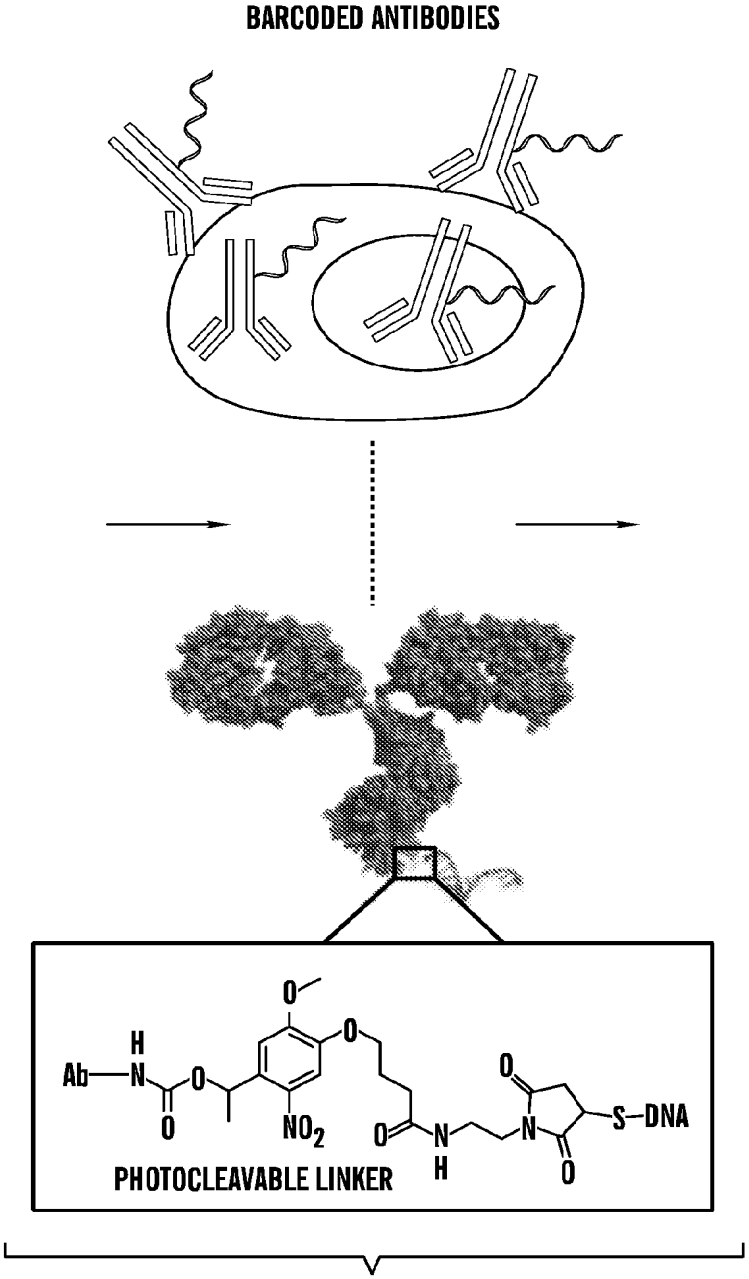


FIG. 1B

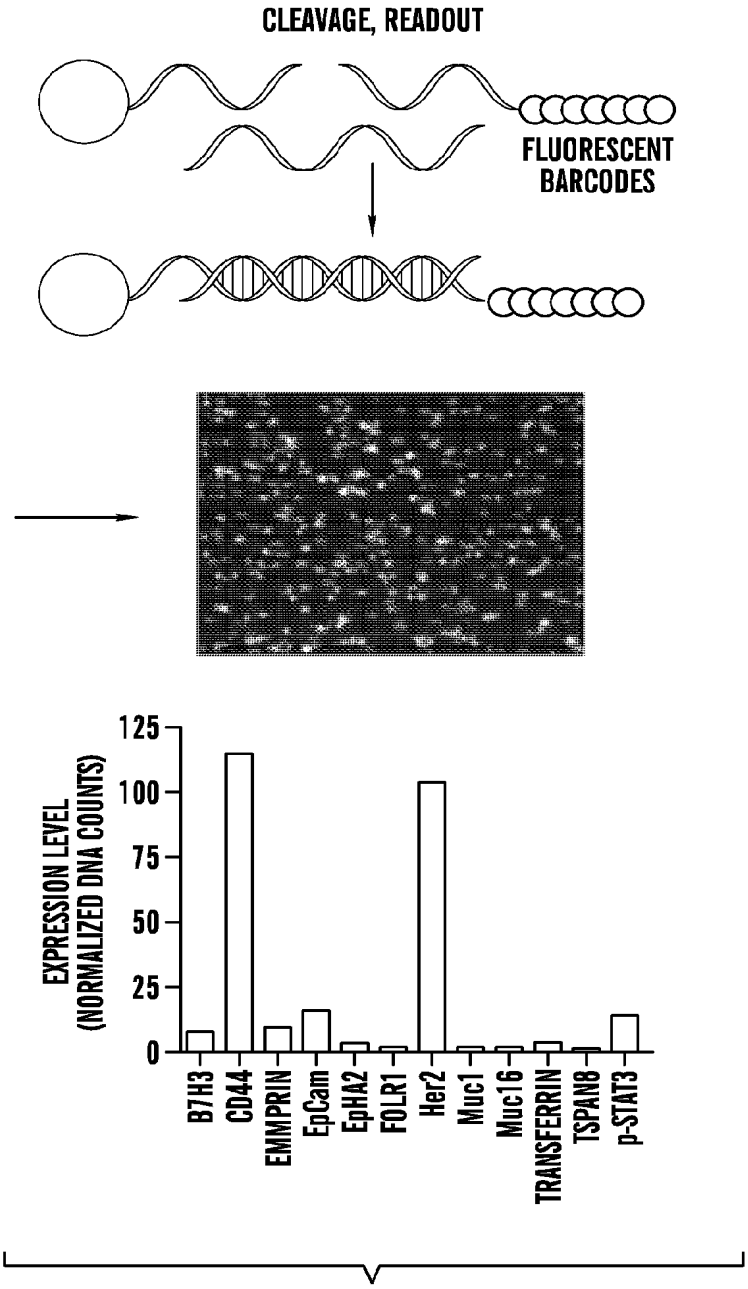

CLEAVAGE, READOUT
FLUORESCENT BARCODES
EXPRESSION LEVEL (NORMALIZED DNA COUNTS)
125
100
75
50
25
0
B7H3
CD44
EMMPRIN
EpCam
EphA2
FOLR1
Her2
Muc1
Muc16
TRANSFERRIN
TSPAN8
p-STAT3

FIG. 1C

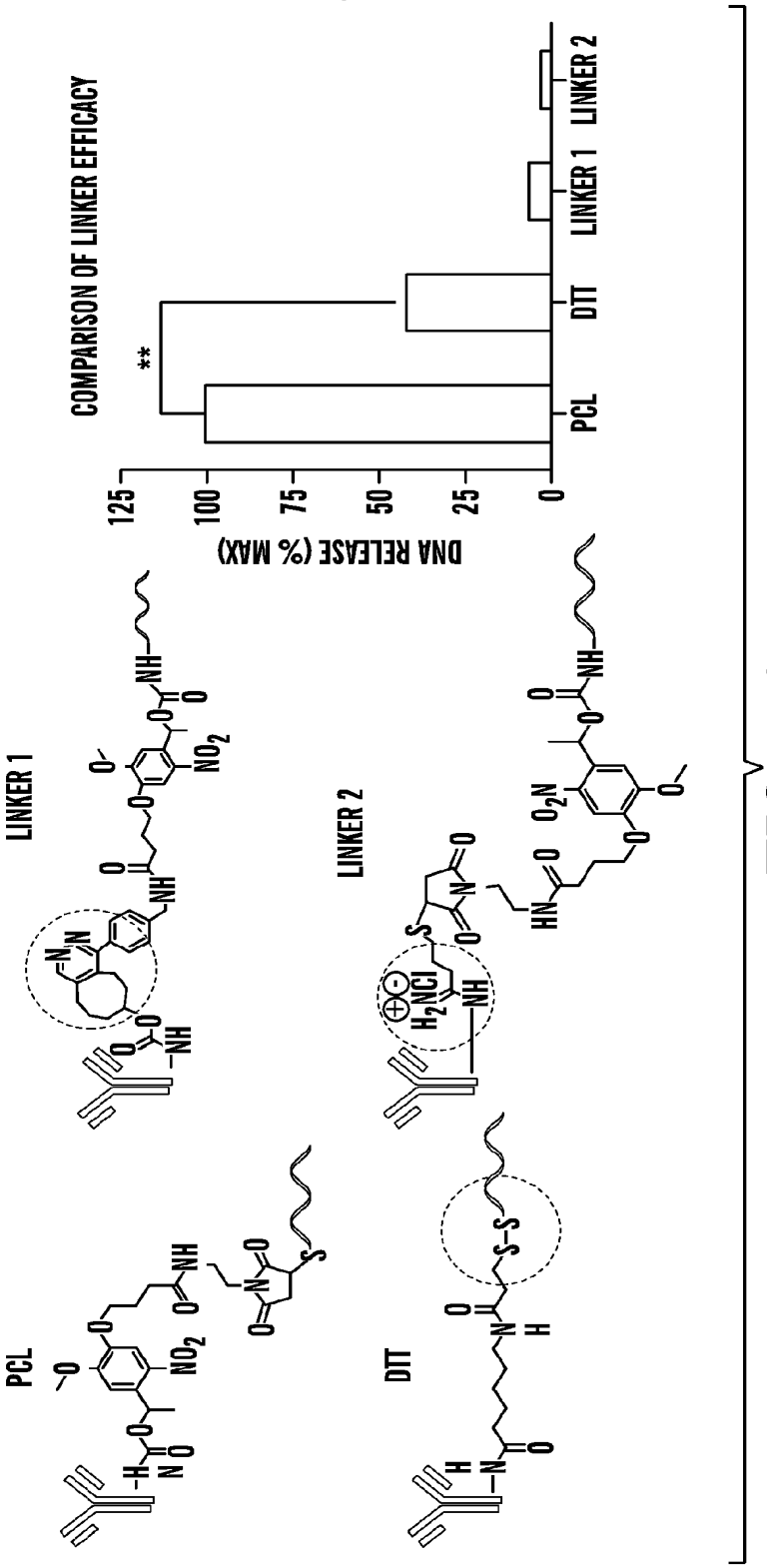

COMPARISON OF LINKER EFFICACY
DNA RELEASE (% MAX)
125
100
75
50
25
0
**
PCL
DTT
LINKER 1
LINKER 2
LINKER 1
LINKER 2
PCL
DTT

FIG. 2A

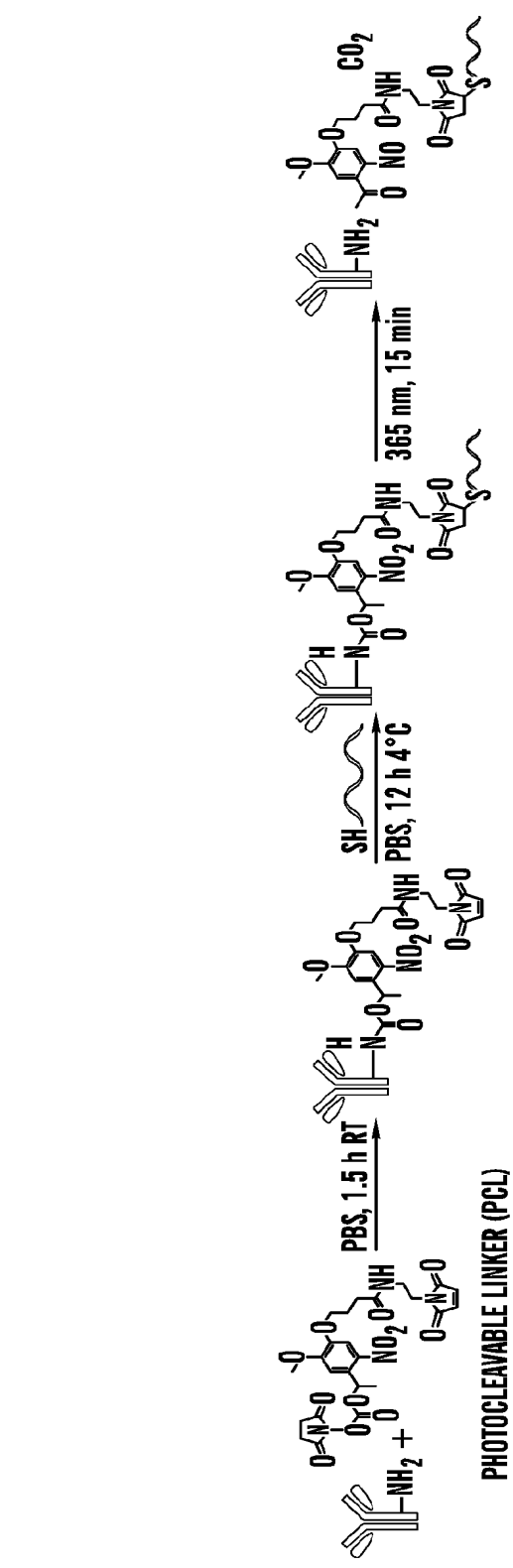


FIG. 2B

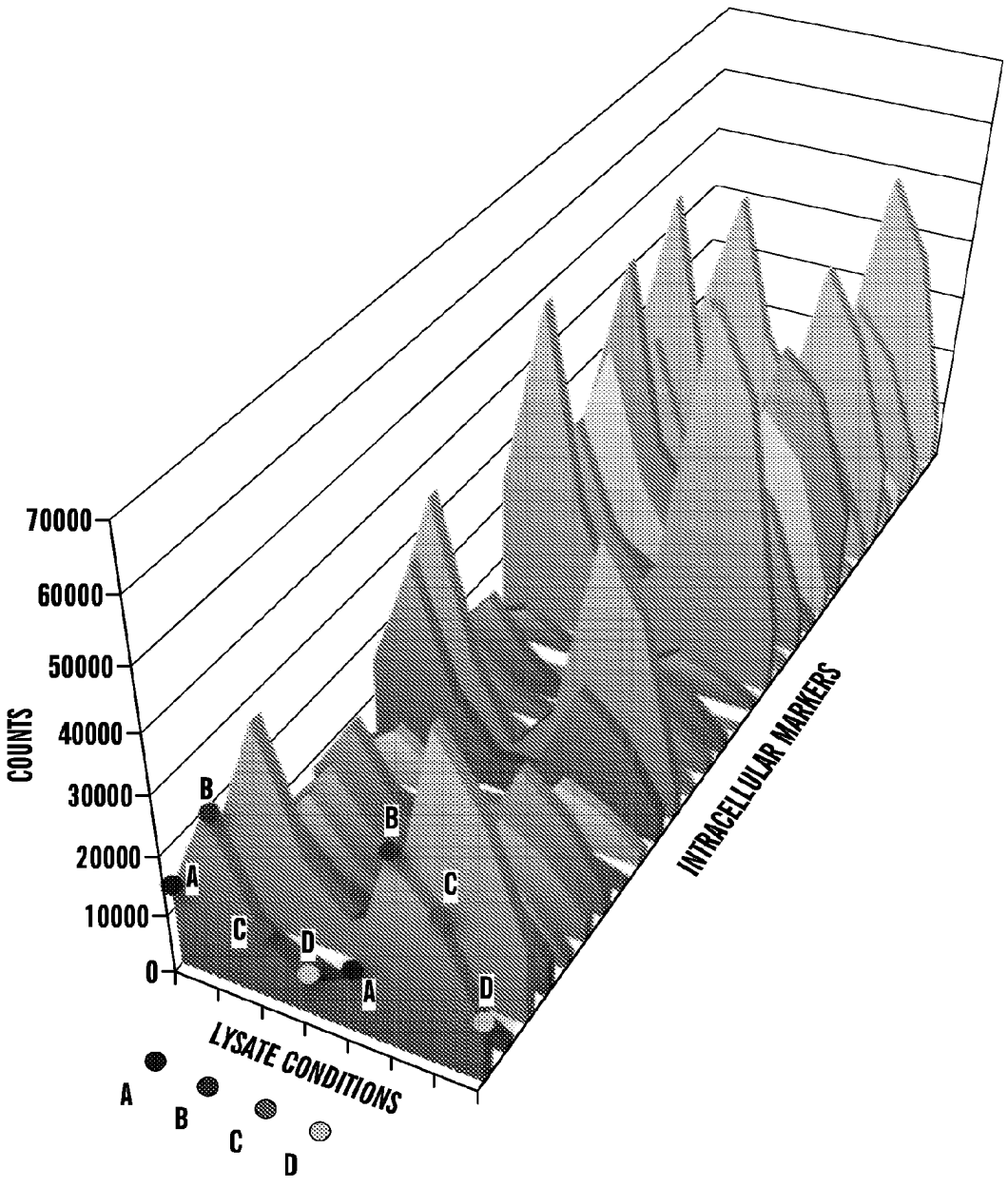

70000
60000
50000
40000
30000
20000
10000
0
COUNTS
LYSATE CONDITIONS
INTRACELLULAR MARKERS
A
B
C
D

FIG. 3

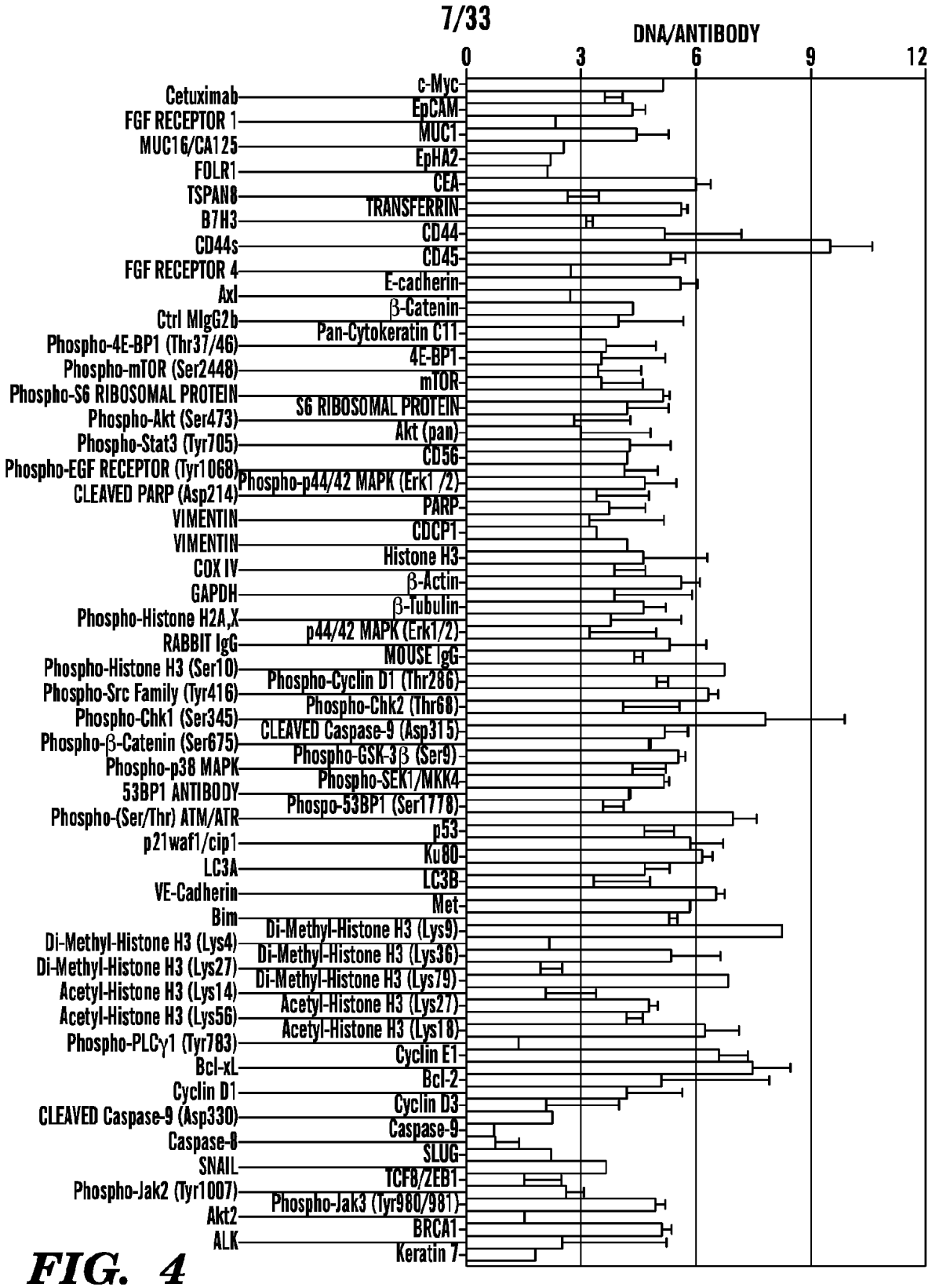


FIG. 4

SUBSTITUTE SHEET (RULE 26)

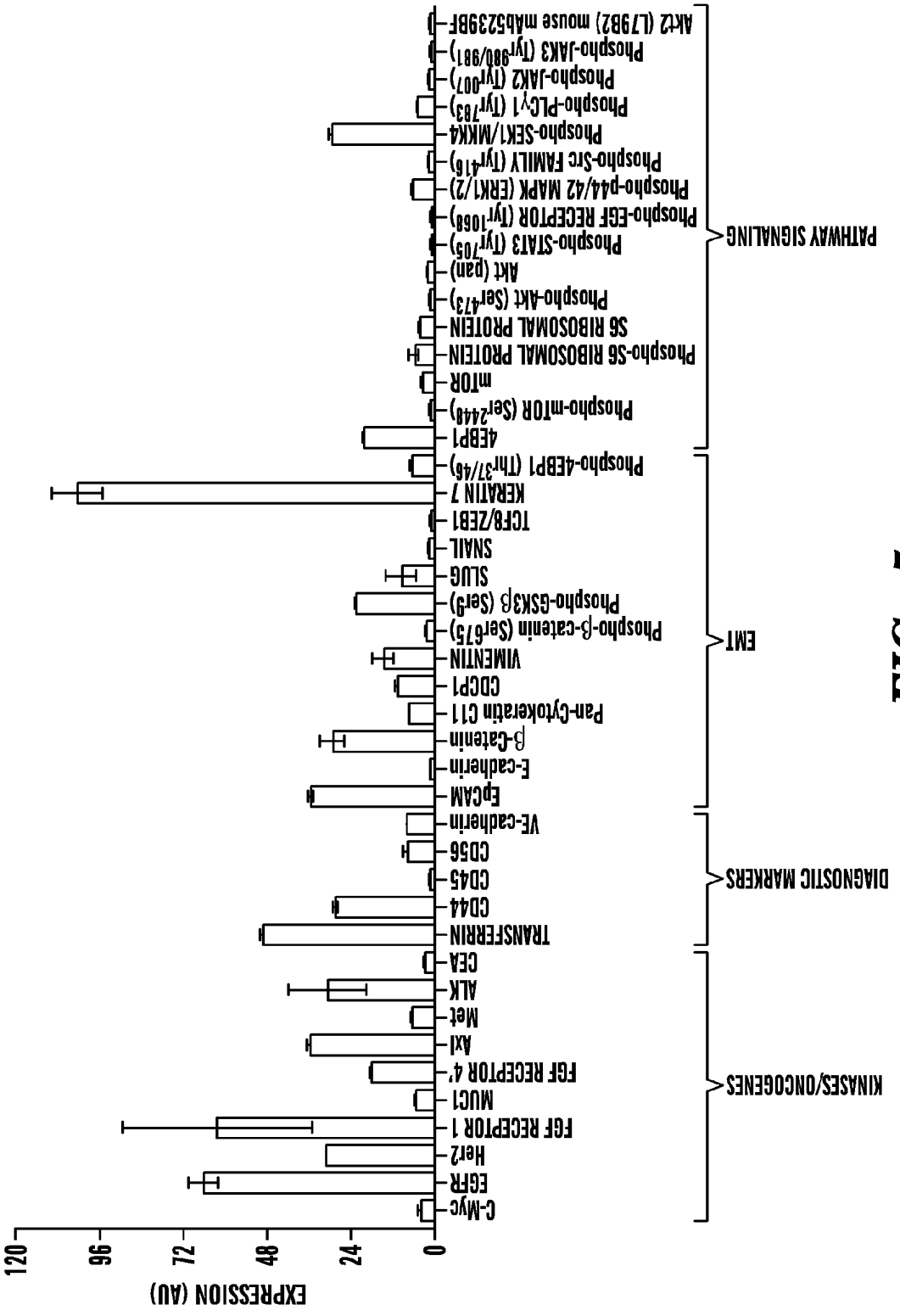

EXPRESSION (AU)
0
24
48
72
96
120
C-Myc
EGFR
Her2
FGF RECEPTOR 1
MUC1
FGF RECEPTOR 4
Axl
Met
ALK
CEA
TRANSFERRIN
CD44
CD45
CD56
VE-cadherin
EpCAM
E-cadherin
β-Catenin
Pan-Cytokeratin C11
CDCP1
VIMENTIN
Phospho-β-catenin (Ser675)
Phospho-GSK3β (Ser9)
SLUG
SNAIL
TCF8/ZEB1
KERATIN 7
Phospho-4EBP1 (Thr37/46)
4EBP1
Phospho-mTOR (Ser2448)
mTOR
Phospho-S6 RIBOSOMAL PROTEIN
S6 RIBOSOMAL PROTEIN
Phospho-Akt (Ser473)
Akt (pan)
Phospho-STAT3 (Tyr705)
Phospho-EGF RECEPTOR (Tyr1068)
Phospho-p44/42 MAPK (ERK1/2)
Phospho-Src FAMILY (Tyr416)
Phospho-SEK1/MKK4
Phospho-PLCγ1 (Tyr783)
Phospho-JAK2 (Tyr1007)
Phospho-JAK3 (Tyr980/981)
Akt2 (L79B2) mouse mAb5239BF
KINASES/ONCOGENES
DIAGNOSTIC MARKERS
EMT
PATHWAY SIGNALING

FIG. 5

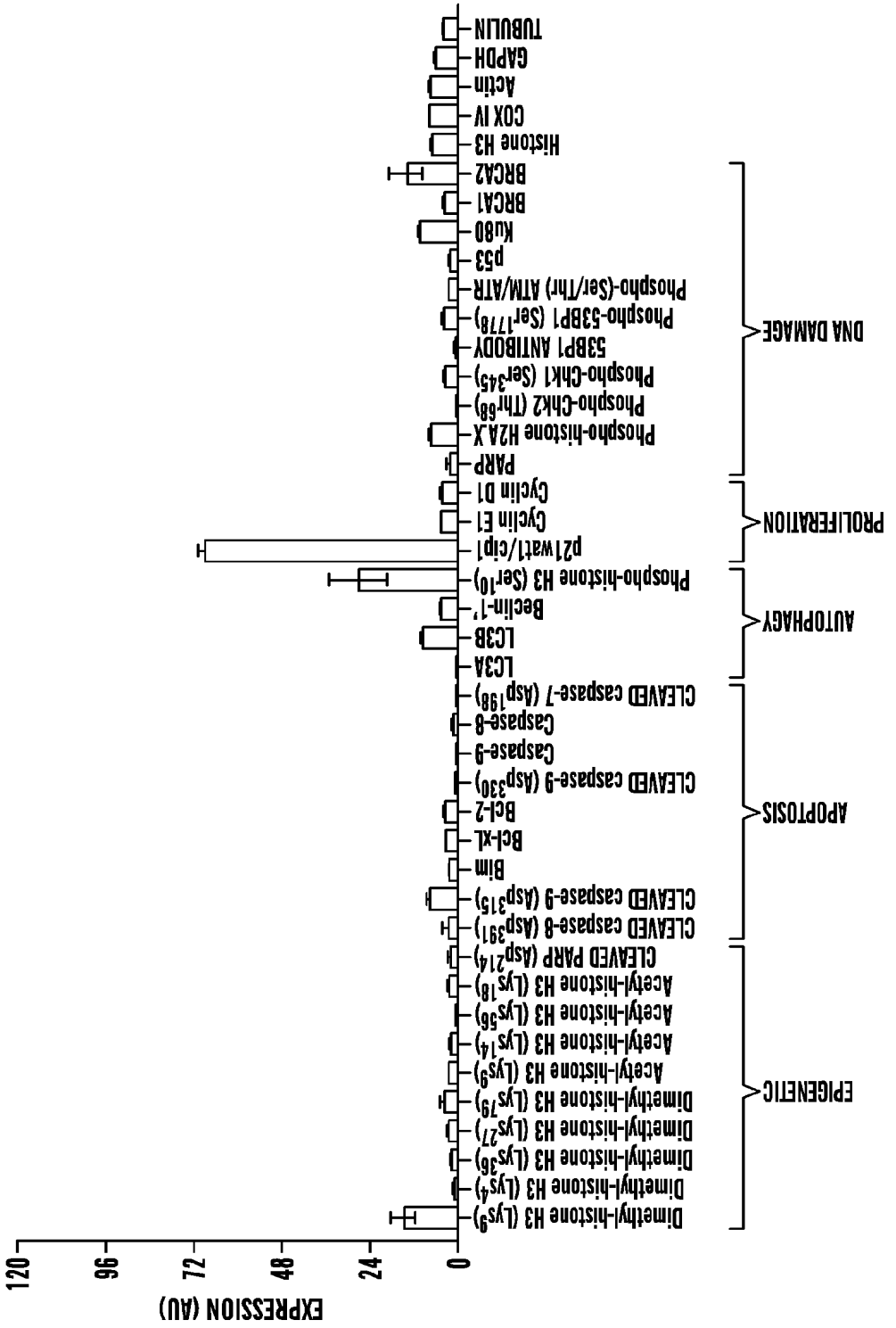

EXPRESSION (AU)
120
96
72
48
24
0
EPIGENETIC
Dimethyl-histone H3 (Lys9)
Dimethyl-histone H3 (Lys4)
Dimethyl-histone H3 (Lys36)
Dimethyl-histone H3 (Lys27)
Dimethyl-histone H3 (Lys79)
Acetyl-histone H3 (Lys9)
Acetyl-histone H3 (Lys14)
Acetyl-histone H3 (Lys56)
Acetyl-histone H3 (Lys18)
CLEAVED PARP (Asp214)
APOPTOSIS
CLEAVED caspase-8 (Asp391)
CLEAVED caspase-9 (Asp315)
Bim
Bcl-xL
Bcl-2
CLEAVED caspase-9 (Asp330)
Caspase-9
Caspase-8
CLEAVED caspase-7 (Asp198)
AUTOPHAGY
LC3A
LC3B
Beclin-1
Phospho-histone H3 (Ser10)
PROLIFERATION
p21wat1/cip1
Cyclin E1
Cyclin D1
PARP
DNA DAMAGE
Phospho-histone H2A.X
Phospho-Chk2 (Thr68)
Phospho-Chk1 (Ser345)
53BP1 ANTIBODY
Phospho-53BP1 (Ser1778)
Phospho-(Ser/Thr) ATM/ATR
p53
Ku80
BRCA1
BRCA2
Histone H3
COX IV
Actin
GAPDH
TUBULIN

FIG. 5 (cont.)

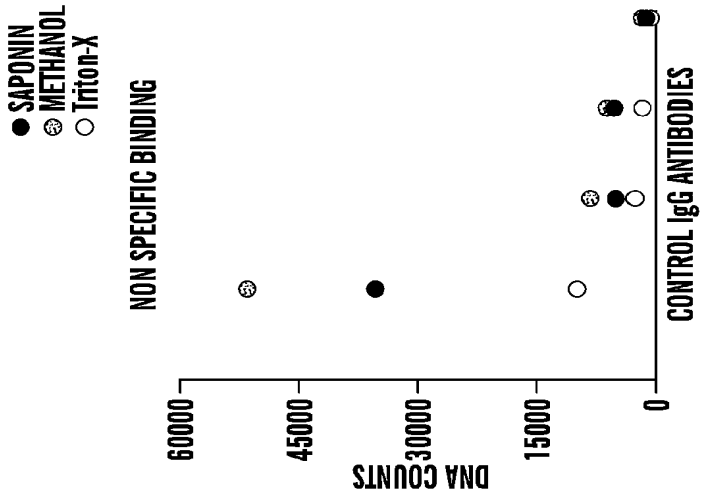

SAPONIN
METHANOL
Triton-X
NON SPECIFIC BINDING
DNA COUNTS
60000
45000
30000
15000
0
CONTROL IgG ANTIBODIES

FIG. 6B

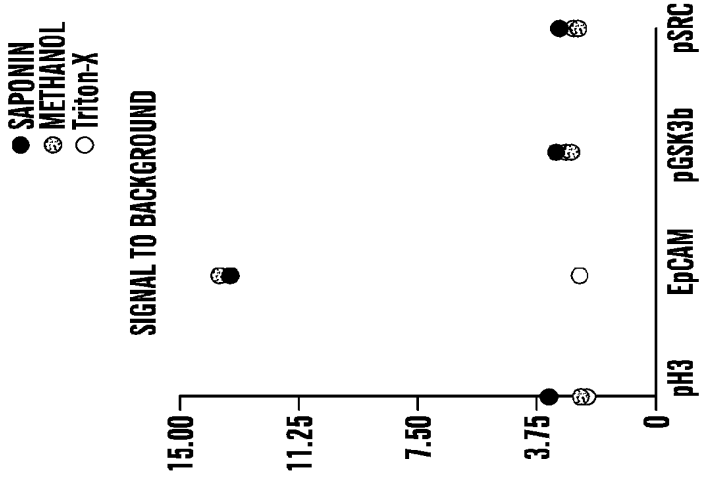

SAPONIN
METHANOL
Triton-X
SIGNAL TO BACKGROUND
15.00
11.25
7.50
3.75
0
pH3
EpCAM
pGSK3b
pSRC

FIG. 6A

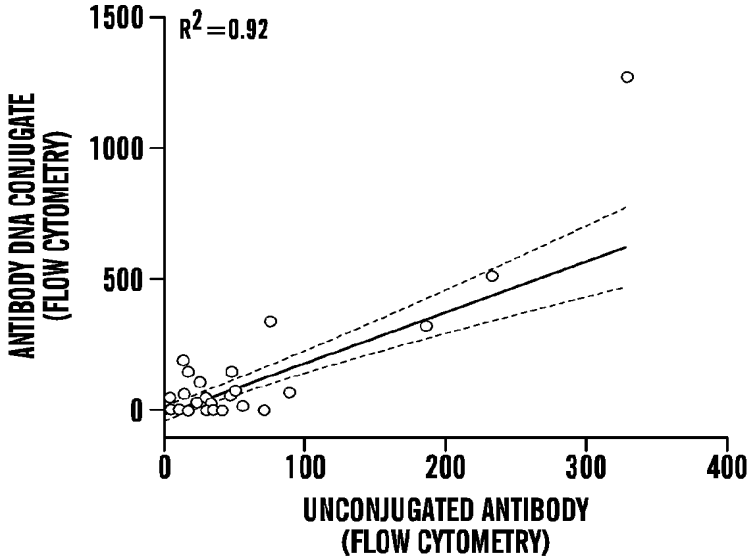

1500
1000
500
0
R^2=0.92
ANTIBODY DNA CONJUGATE (FLOW CYTOMETRY)
0
100
200
300
400
UNCONJUGATED ANTIBODY (FLOW CYTOMETRY)

FIG. 7A

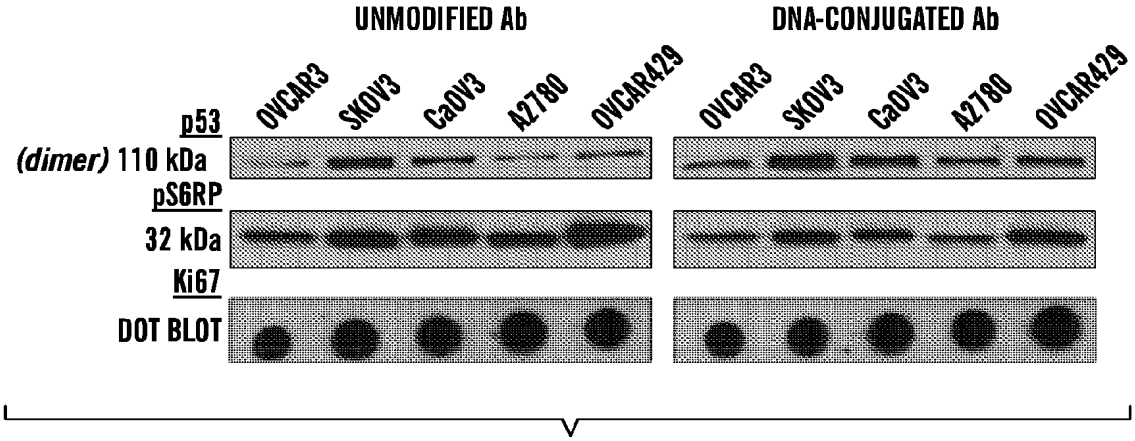

UNMODIFIED Ab
DNA-CONJUGATED Ab
OVCAR3
SKOV3
CaOV3
AZ780
OVCAR429
p53
(dimer) 110 kDa
pS6RP
32 kDa
Ki67
DOT BLOT

FIG. 7B

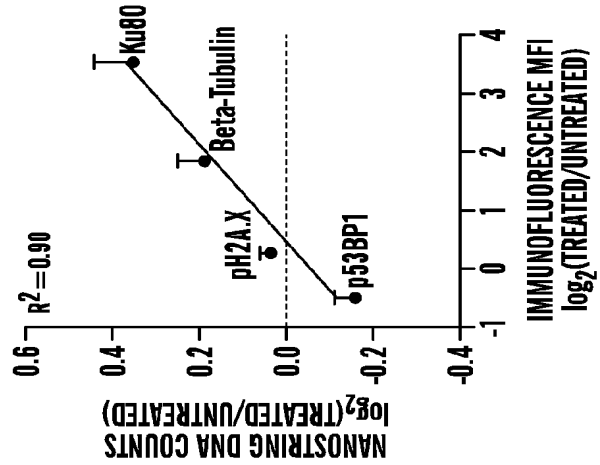


FIG. 8C

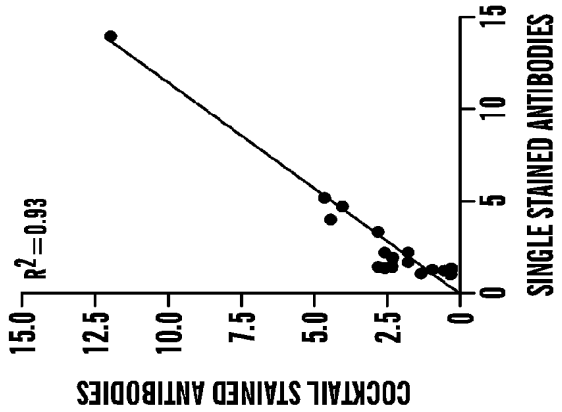


FIG. 8B

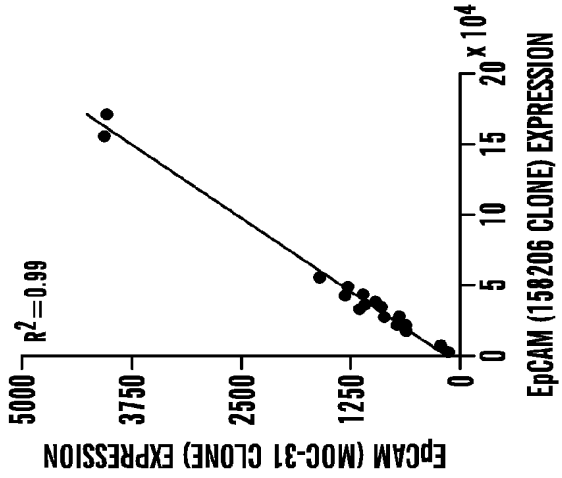


FIG. 8A

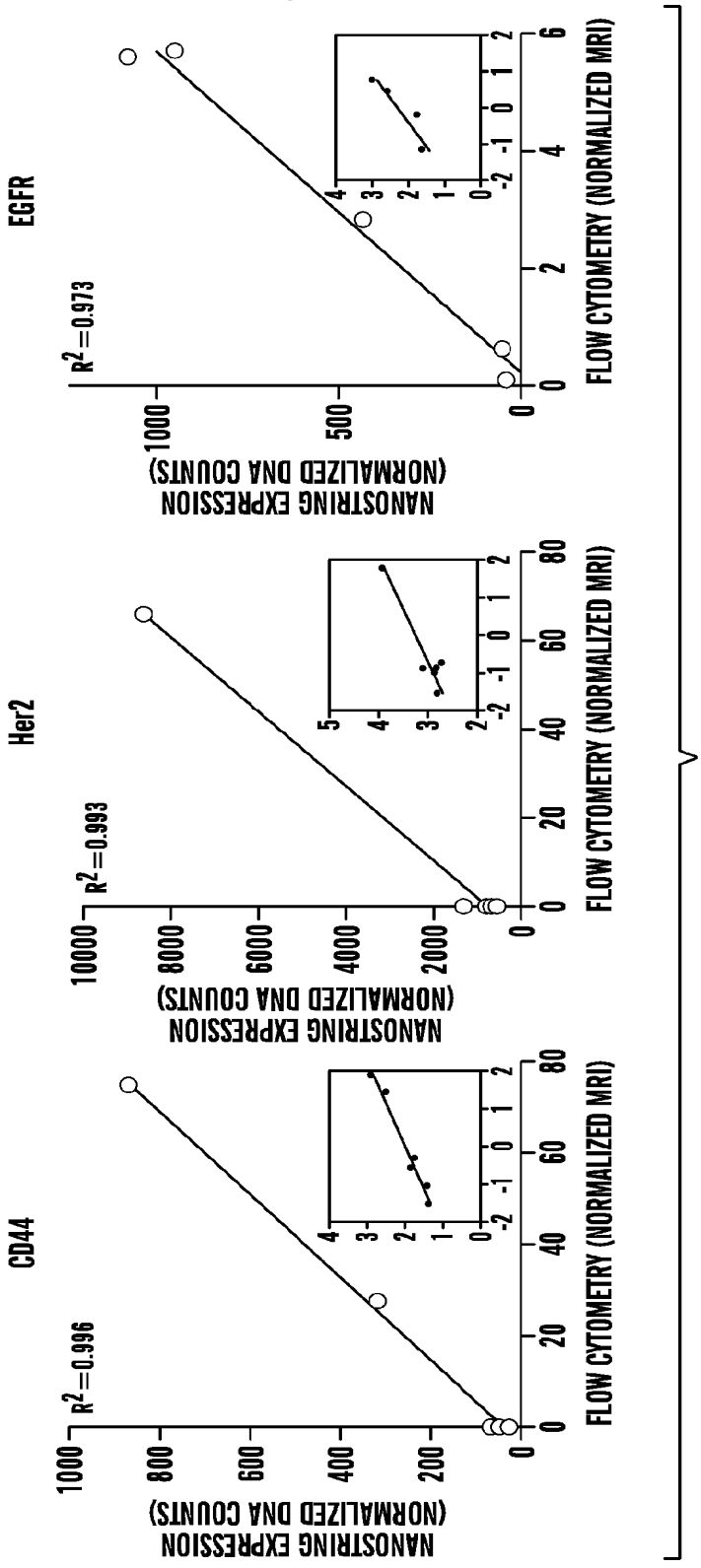

CD44
R^2=0.996
NANOSTRING EXPRESSION (NORMALIZED DNA COUNTS)
FLOW CYTOMETRY (NORMALIZED MRI)
Her2
R^2=0.993
NANOSTRING EXPRESSION (NORMALIZED DNA COUNTS)
FLOW CYTOMETRY (NORMALIZED MRI)
EGFR
R^2=0.973
NANOSTRING EXPRESSION (NORMALIZED DNA COUNTS)
FLOW CYTOMETRY (NORMALIZED MRI)

FIG. 9

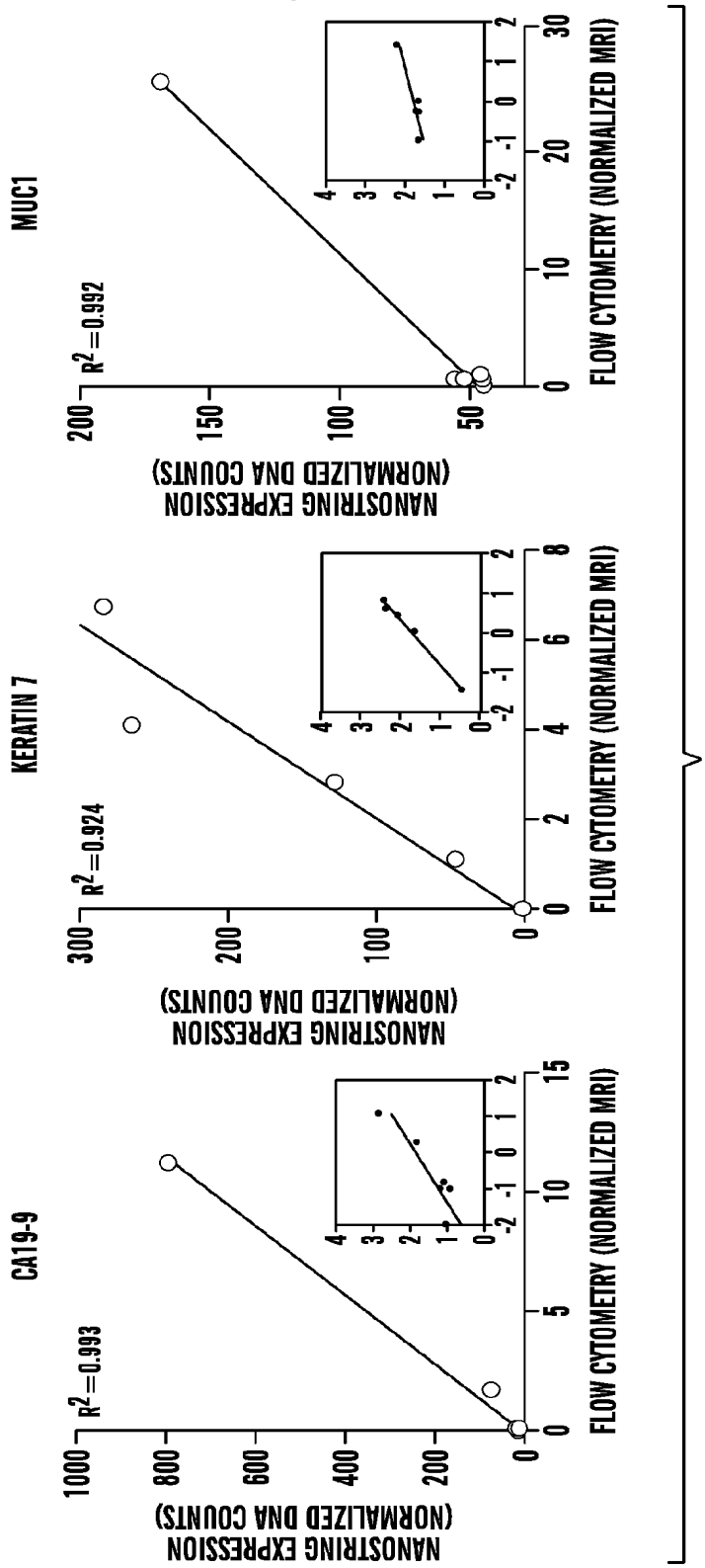

CA19-9
R^2=0.993
NANOSTRING EXPRESSION (NORMALIZED DNA COUNTS)
FLOW CYTOMETRY (NORMALIZED MRI)
KERATIN 7
R^2=0.924
NANOSTRING EXPRESSION (NORMALIZED DNA COUNTS)
FLOW CYTOMETRY (NORMALIZED MRI)
MUC1
R^2=0.992
NANOSTRING EXPRESSION (NORMALIZED DNA COUNTS)
FLOW CYTOMETRY (NORMALIZED MRI)

FIG. 9 (cont.)

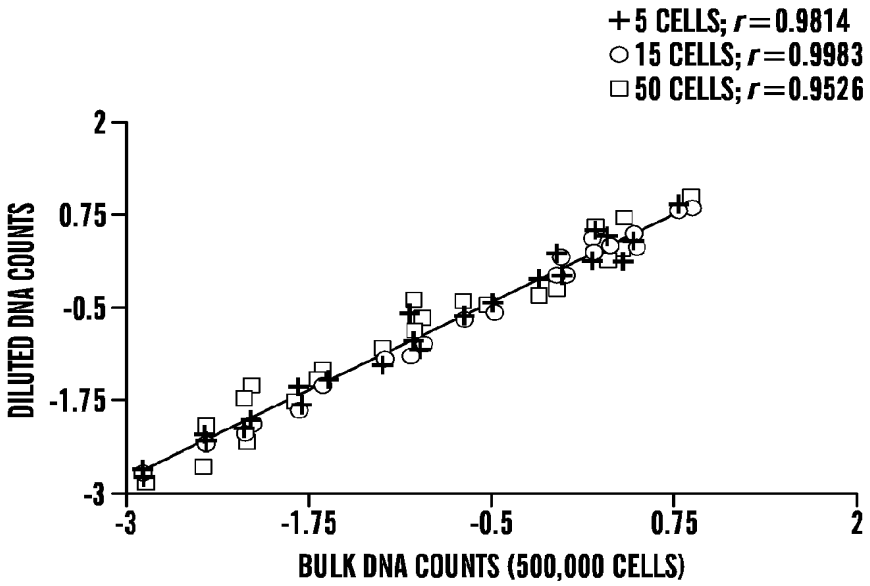


FIG. 10A

CORRELATION TO BULK	SPEARMAN *r*	CONFIDENCE INTERVAL
SINGLE CELL 1	0.938	0.897-0.963
SINGLE CELL 2	0.963	0.939-0.978
SINGLE CELL 3	0.53	0.317-0.686
SINGLE CELL 4	0.61	0.426-0.746

FIG. 10B

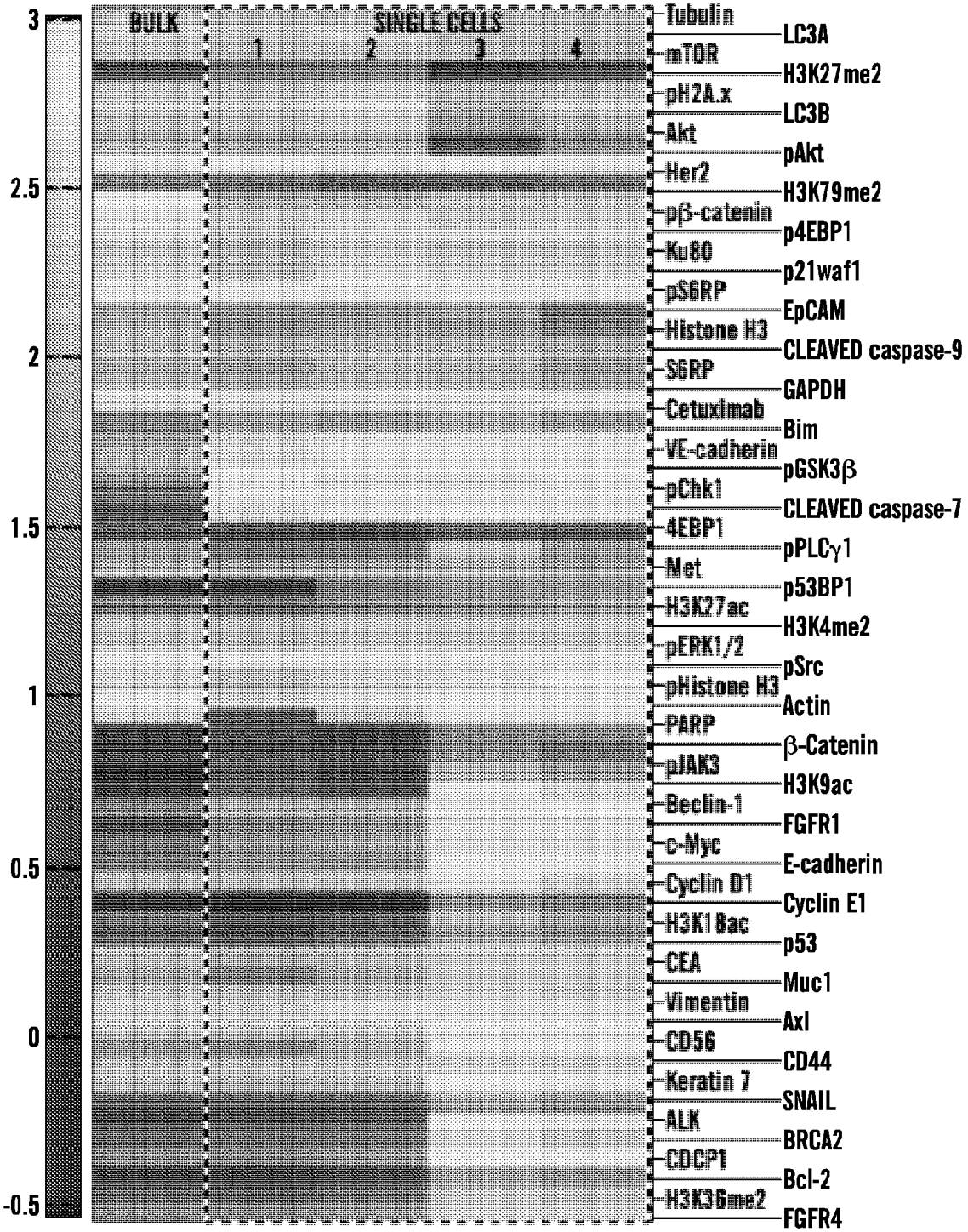

3
2.5
2
1.5
1
0.5
0
-0.5
BULK
SINGLE CELLS
1
2
3
4
Tubulin
LC3A
mTOR
H3K27me2
pH2A.x
LC3B
Akt
pAkt
Her2
H3K79me2
pβ-catenin
p4EBP1
Ku80
p21waf1
pS6RP
EpCAM
Histone H3
CLEAVED caspase-9
S6RP
GAPDH
Cetuximab
Bim
VE-cadherin
pGSK3β
pChk1
CLEAVED caspase-7
4EBP1
pPLCγ1
Met
p53BP1
H3K27ac
H3K4me2
pERK1/2
pSrc
pHistone H3
Actin
PARP
β-Catenin
pJAK3
H3K9ac
Beclin-1
FGFR1
c-Myc
E-cadherin
Cyclin D1
Cyclin E1
H3K18ac
p53
CEA
Muc1
Vimentin
Axl
CD56
CD44
Keratin 7
SNAIL
ALK
BRCA2
CDCP1
Bcl-2
H3K36me2
FGFR4

FIG. 10C

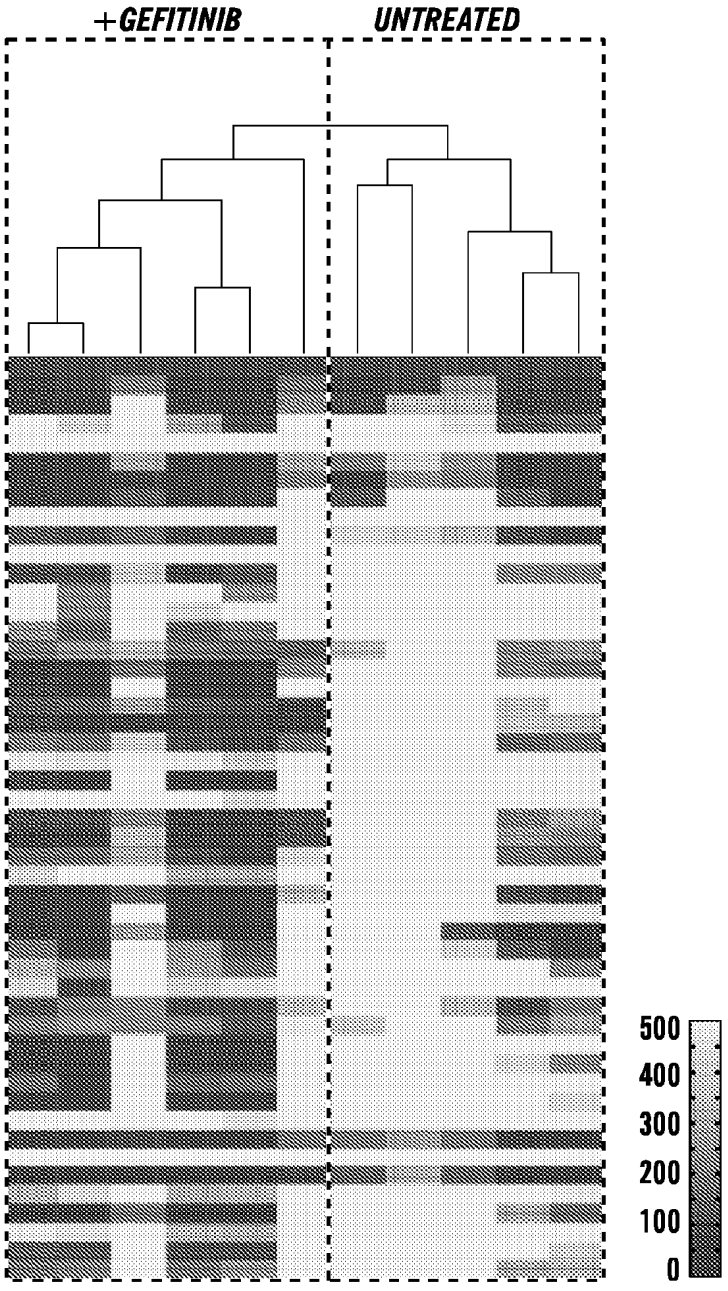


FIG. 11A

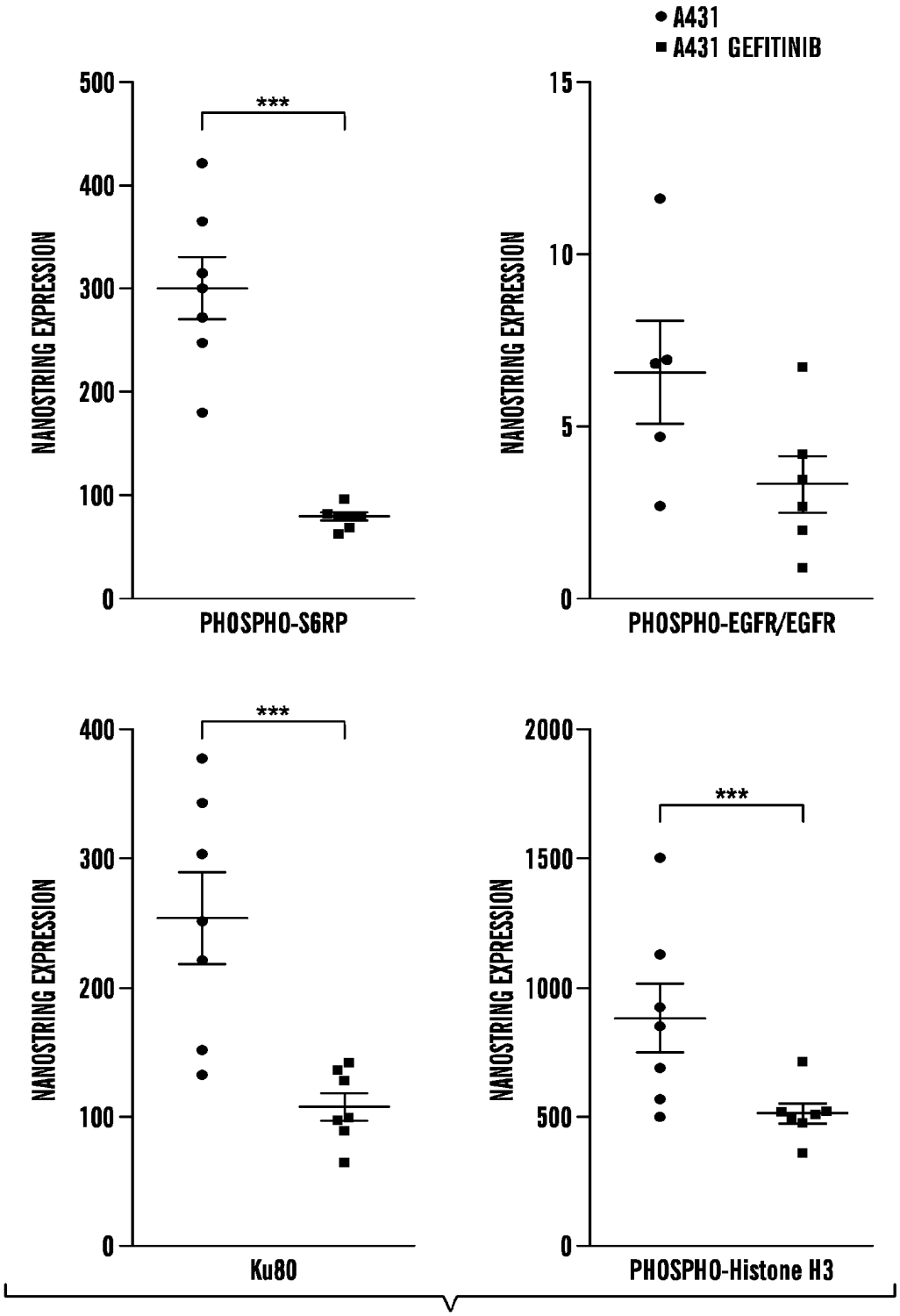

A431
A431 GEFITINIB
NANOSTRING EXPRESSION
500
400
300
200
100
0
***
PHOSPHO-S6RP
15
10
5
0
PHOSPHO-EGFR/EGFR
Ku80
2000
1500
1000
500
PHOSPHO-Histone H3

FIG. 11B

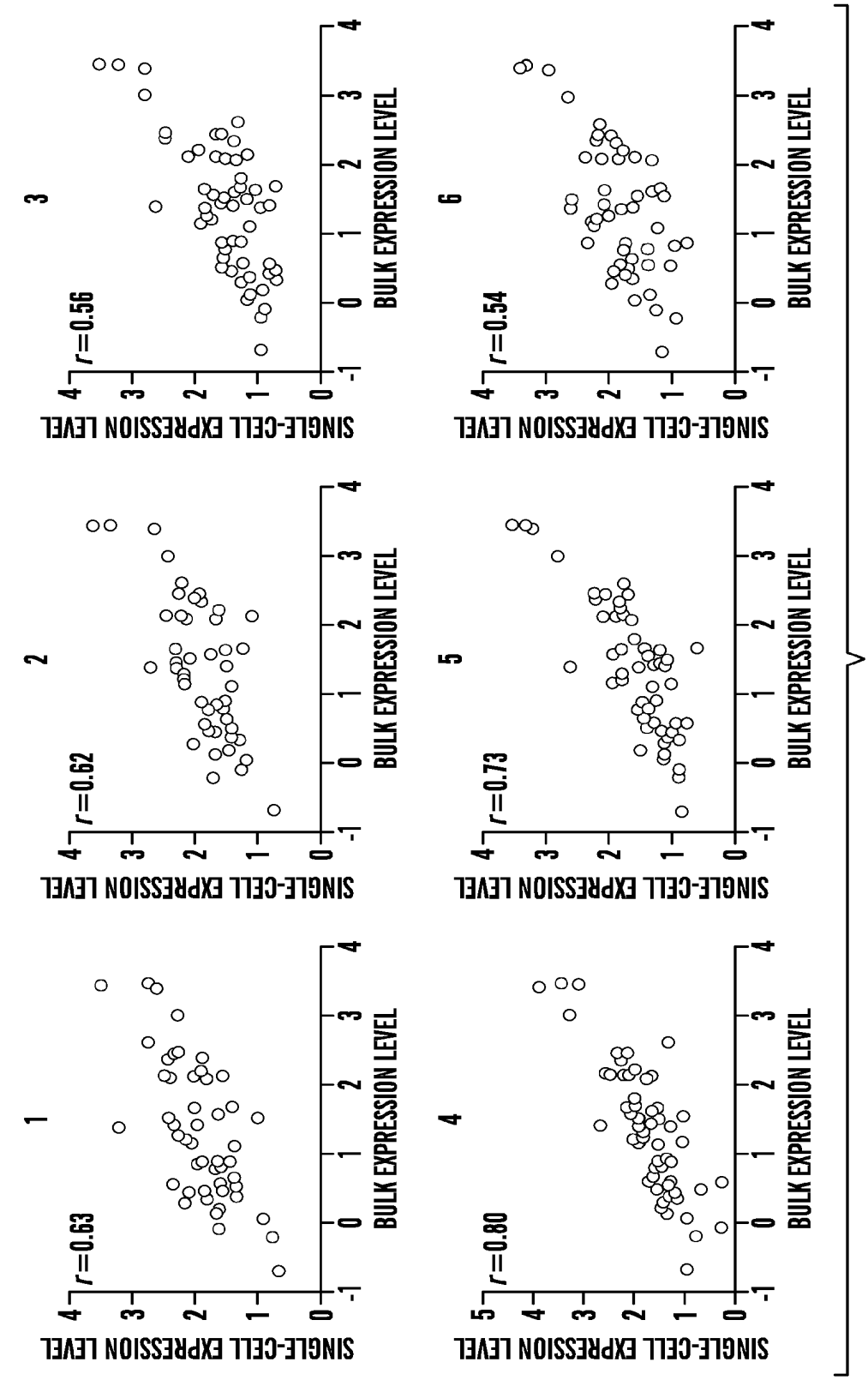

1
r=0.63
2
r=0.62
3
r=0.56
4
r=0.80
5
r=0.73
6
r=0.54
SINGLE-CELL EXPRESSION LEVEL
BULK EXPRESSION LEVEL

FIG. 12A

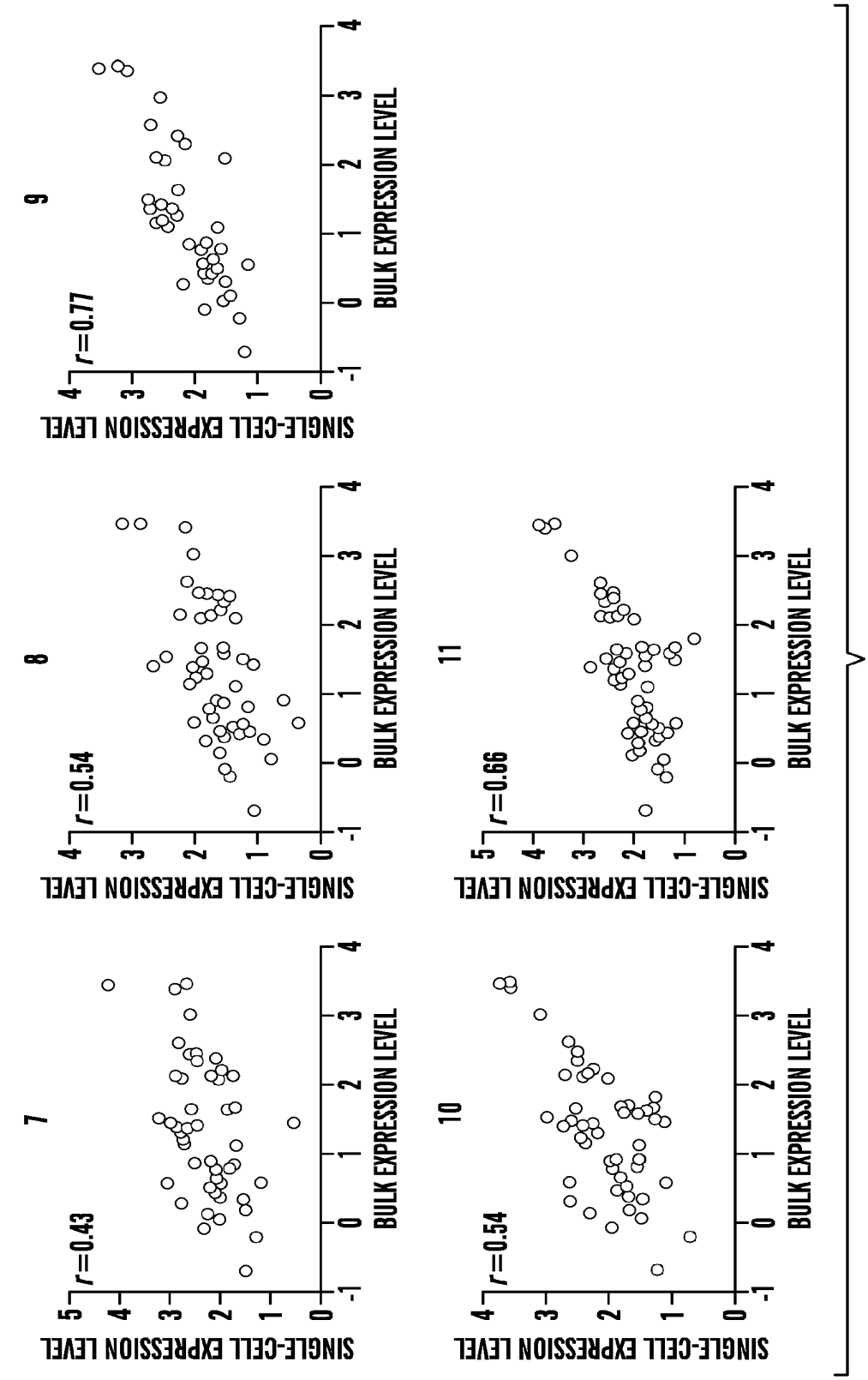

9
r=0.77
8
r=0.54
11
r=0.66
7
r=0.43
10
r=0.54
SINGLE-CELL EXPRESSION LEVEL
BULK EXPRESSION LEVEL

FIG. 12A (cont.)

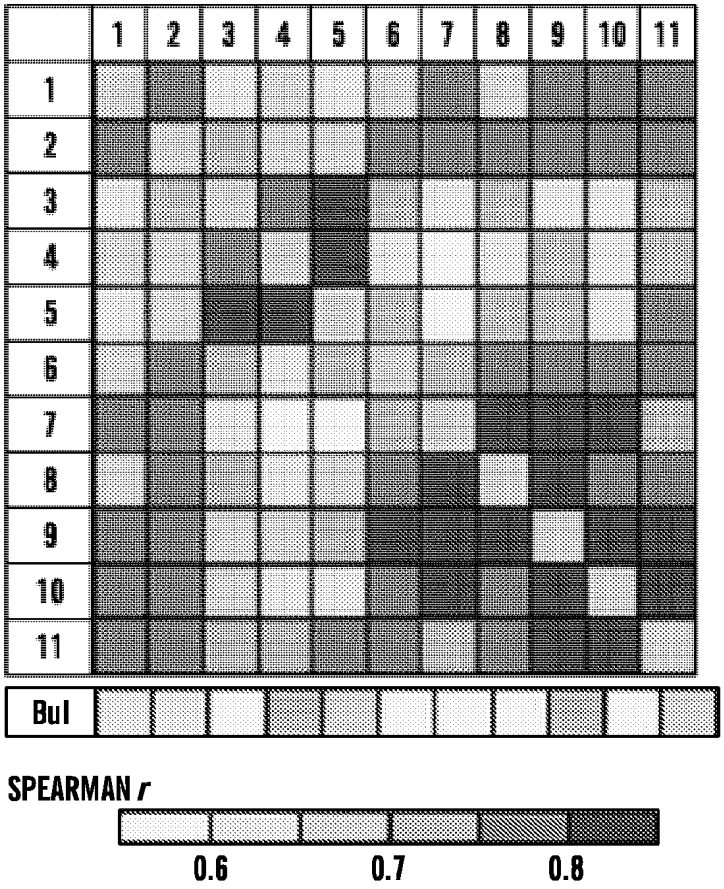


FIG. 12B

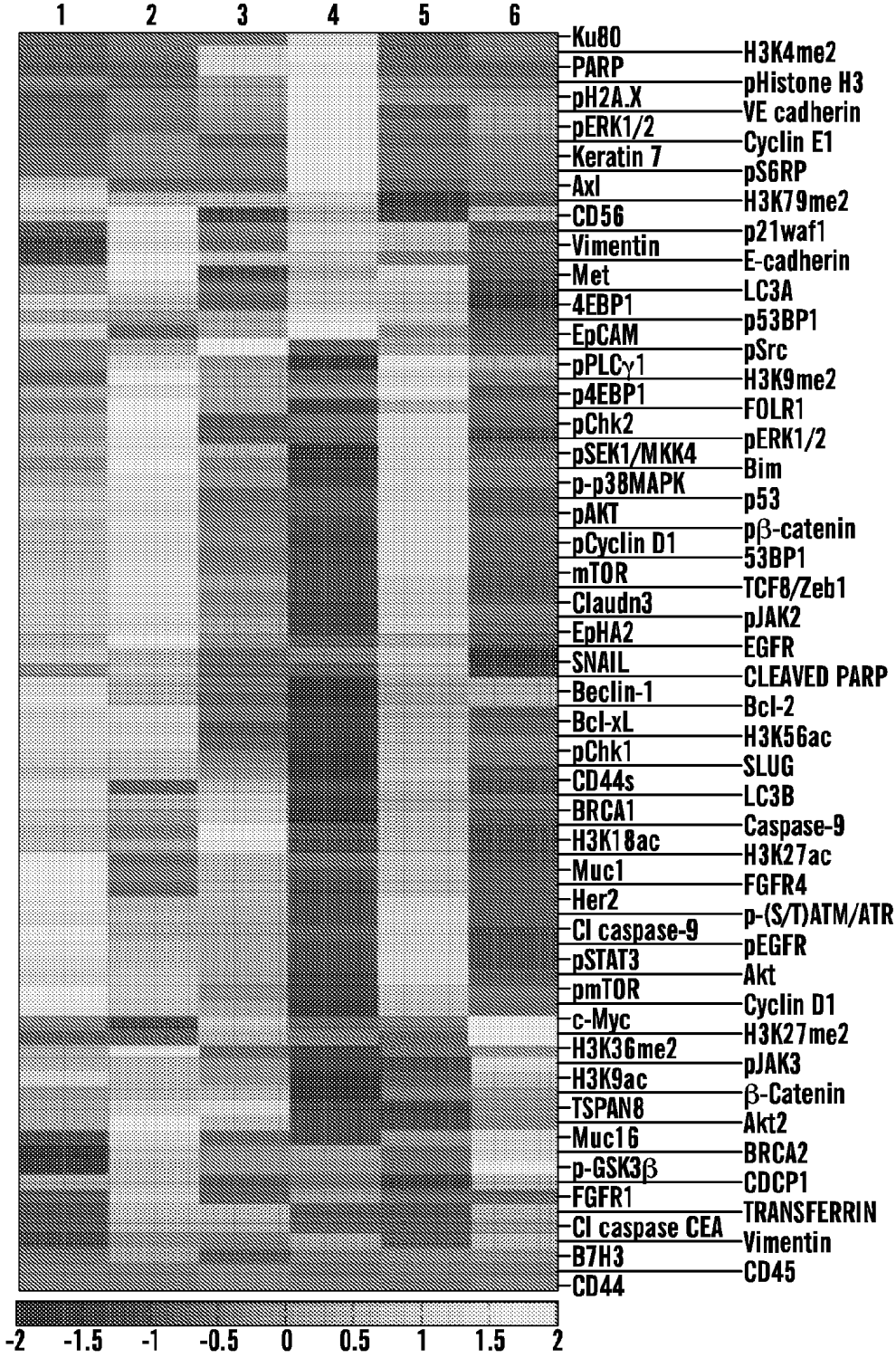


FIG. 13

SUBSTITUTE SHEET (RULE 26)

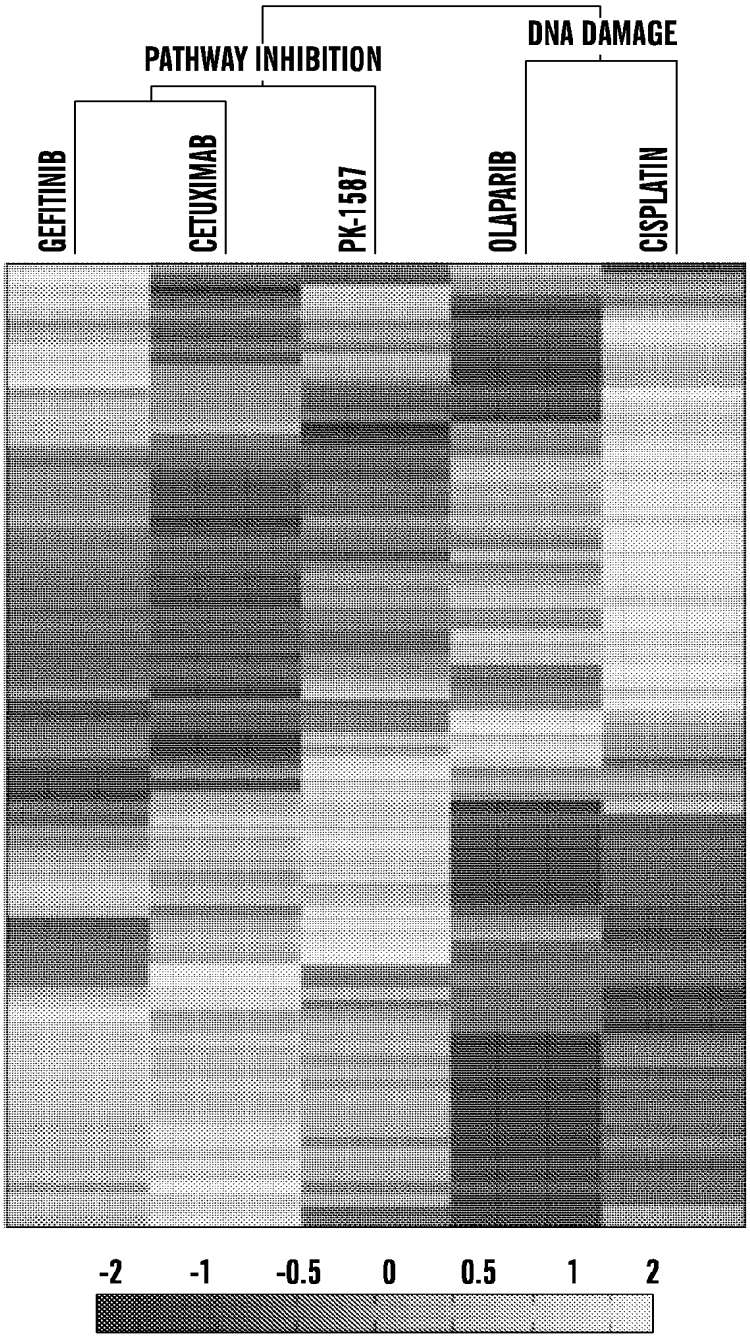

PATHWAY INHIBITION
DNA DAMAGE
GEFITINIB
CETUXIMAB
PK-1587
OLAPARIB
CISPLATIN
-2
-1
-0.5
0
0.5
1
2

FIG. 14A

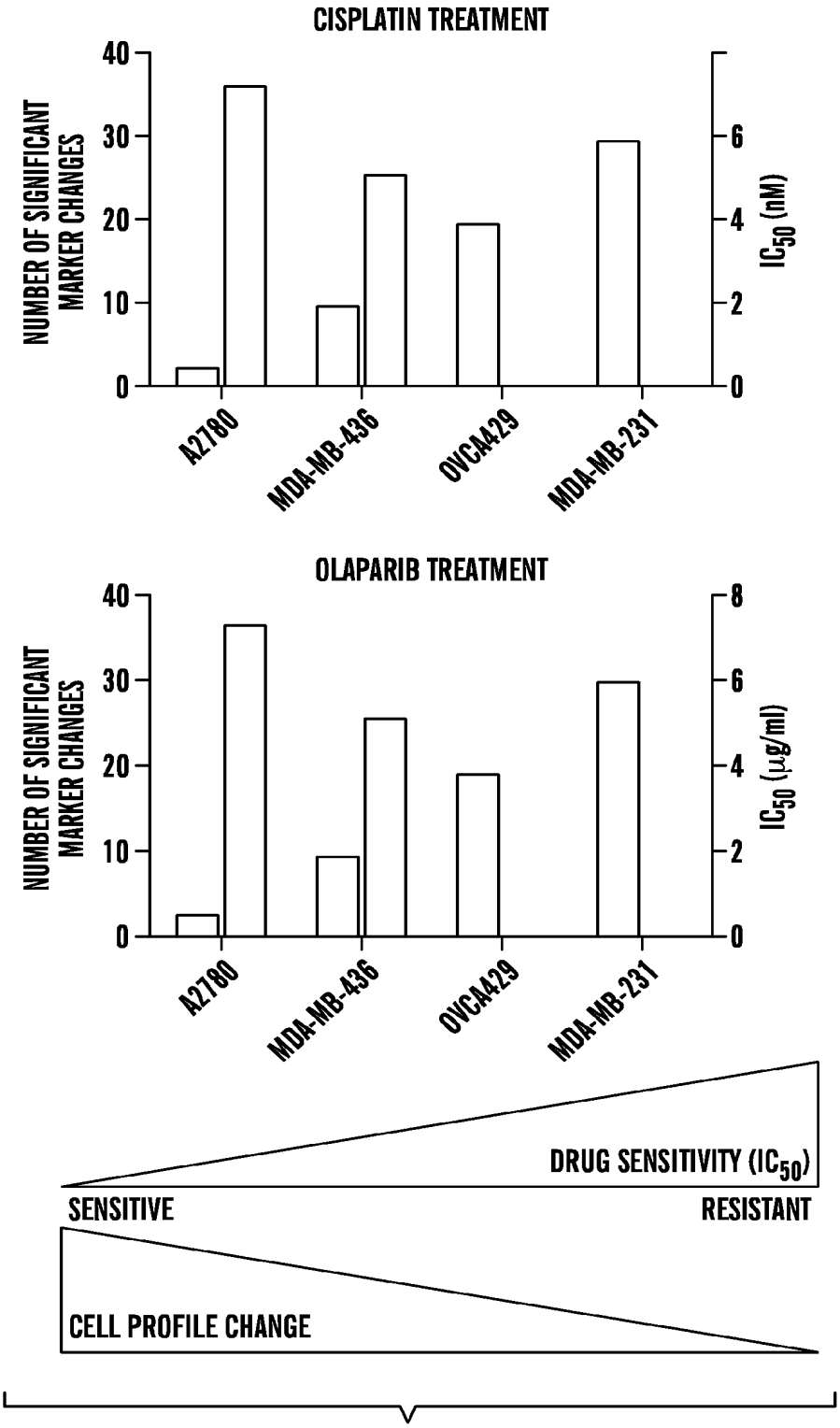

CISPLATIN TREATMENT
NUMBER OF SIGNIFICANT MARKER CHANGES
40
30
20
10
0
6
4
2
0
IC50 (nM)
A2780
MDA-MB-436
OVCA429
MDA-MB-231
OLAPARIB TREATMENT
NUMBER OF SIGNIFICANT MARKER CHANGES
40
30
20
10
0
8
6
4
2
0
IC50 (μg/ml)
A2780
MDA-MB-436
OVCA429
MDA-MB-231
DRUG SENSITIVITY (IC50)
SENSITIVE
RESISTANT
CELL PROFILE CHANGE

FIG. 14B

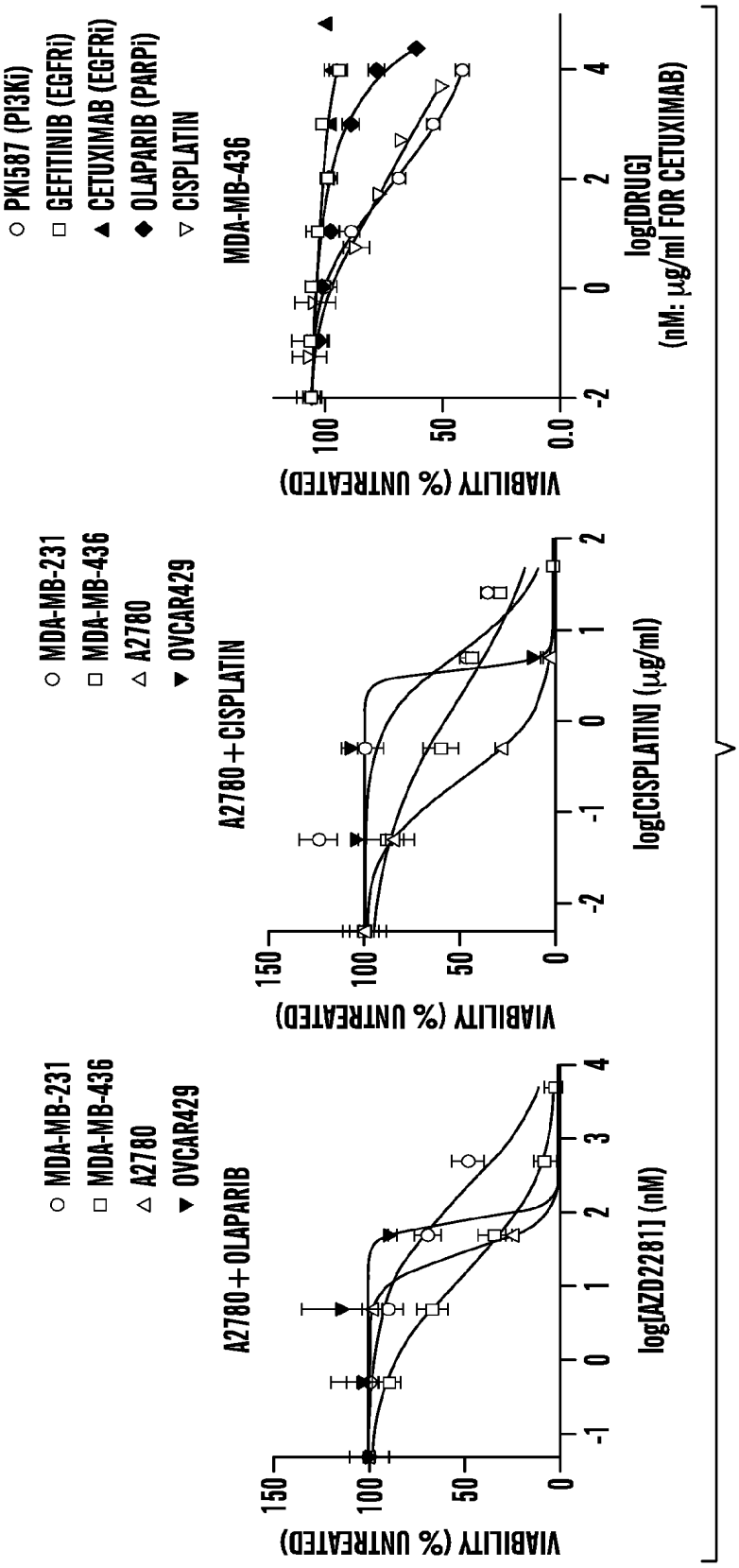

PKI587 (PI3Ki)
GEFITINIB (EGFRi)
CETUXIMAB (EGFRi)
OLAPARIB (PARPi)
CISPLATIN
MDA-MB-436
VIABILITY (% UNTREATED)
log[DRUG]
(nM: μg/ml FOR CETUXIMAB)
MDA-MB-231
MDA-MB-436
A2780
OVCAR429
A2780+CISPLATIN
log[CISPLATIN] (μg/ml)
A2780+OLAPARIB
log[AZD2281] (nM)

FIG. 15A

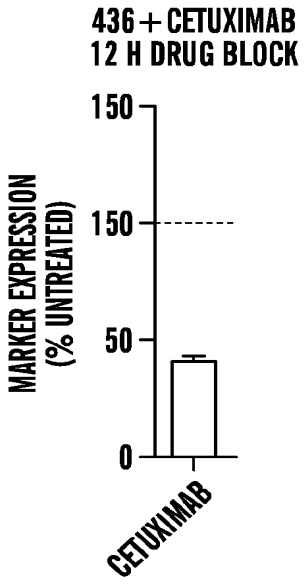


FIG. 15B

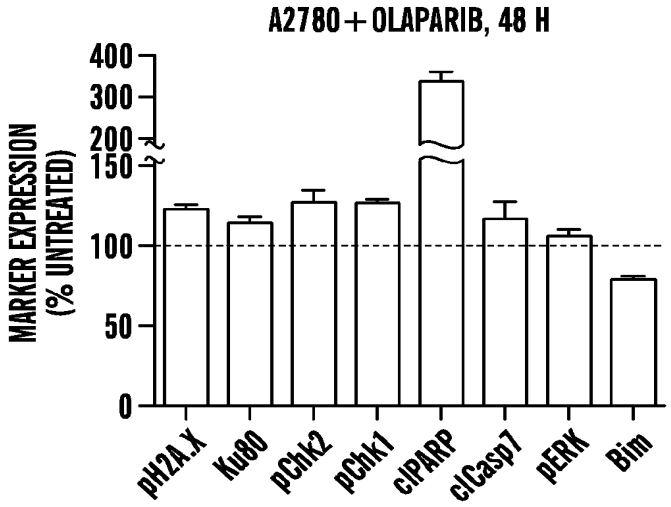


FIG. 15C

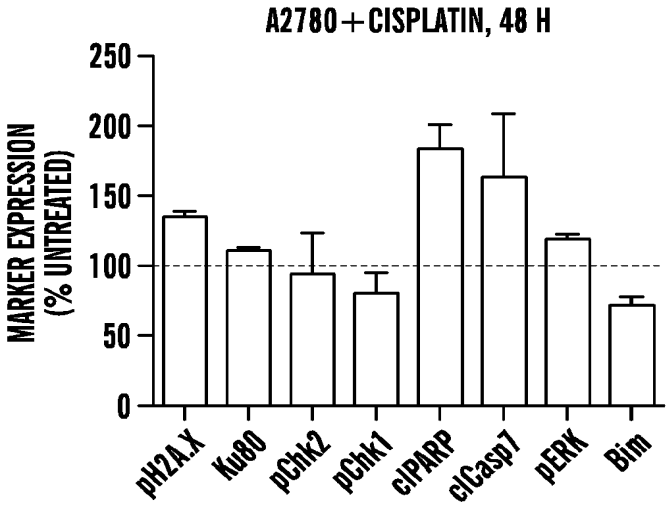


FIG. 15D

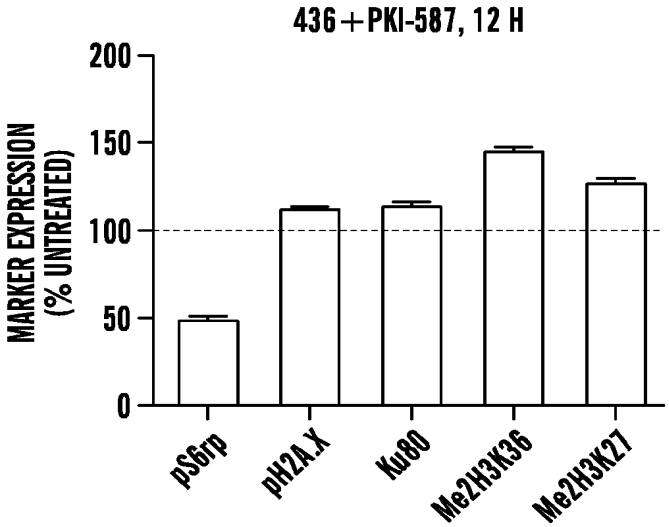


FIG. 15E

MARKER	EC50 (nM)	R^2
Phos-53BP1	7.76	0.76
pCyclinD	34.15	0.88
Ku80	52.92	0.78
FGFR4	26.81	0.95
Cyclin E1	4.971	0.91
BRCA2	17.11	0.84
pERK1/2	12.74	0.78

FIG. 16B

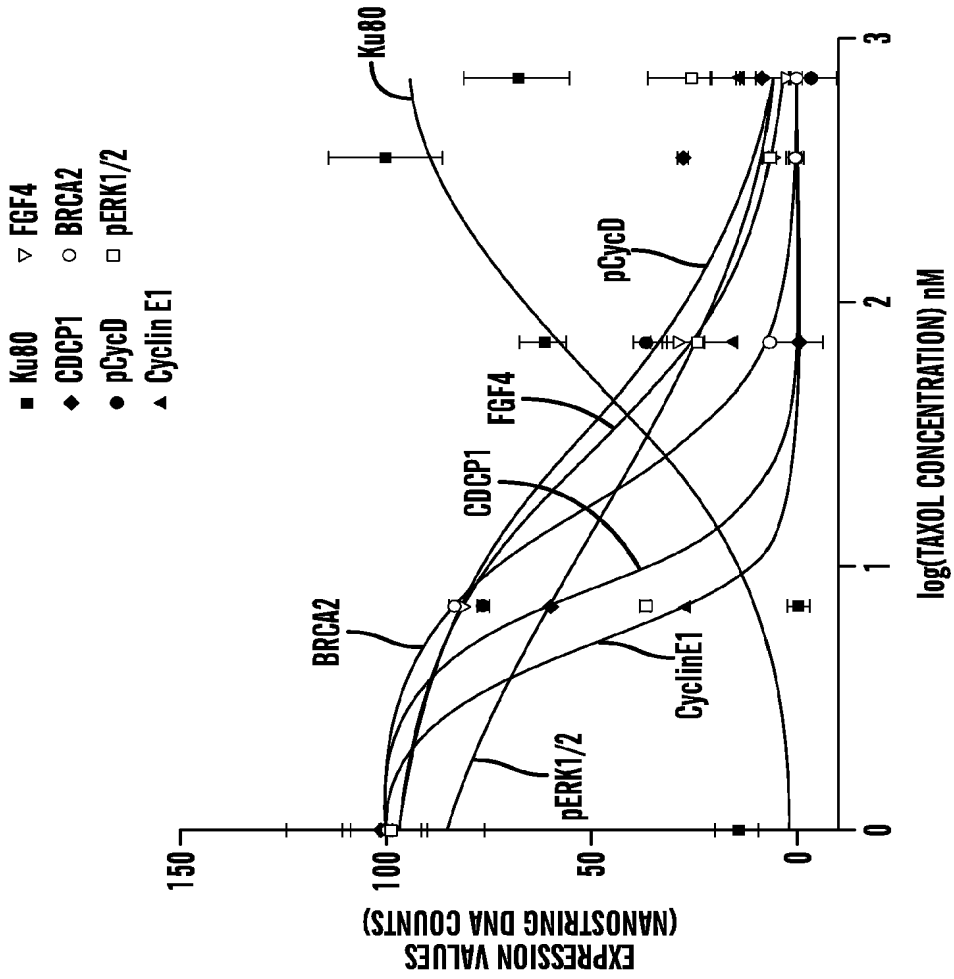


FIG. 16A

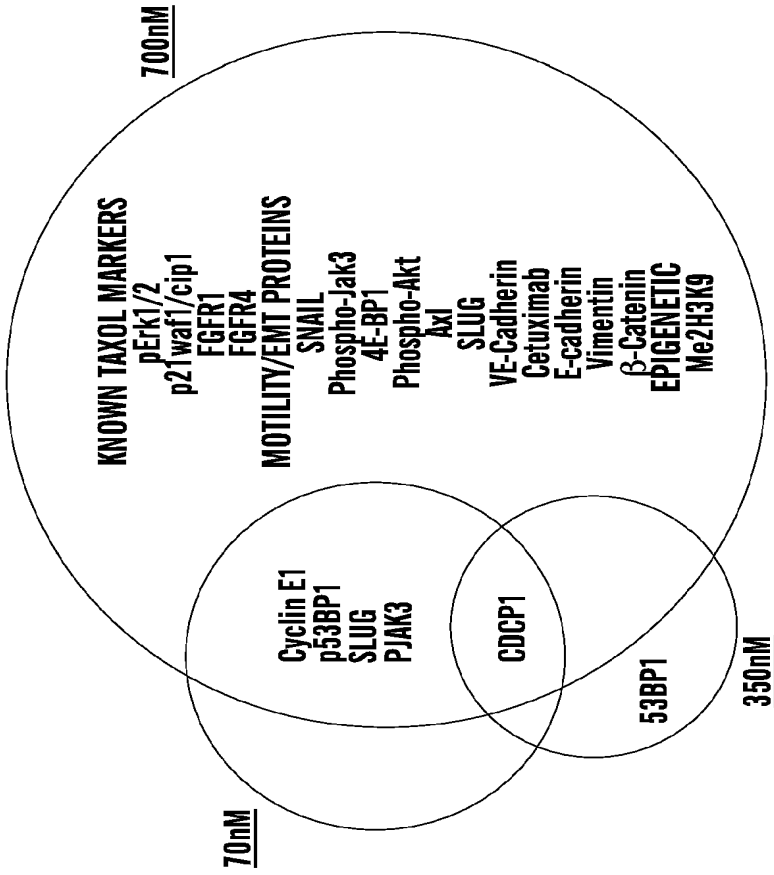


FIG. 16D

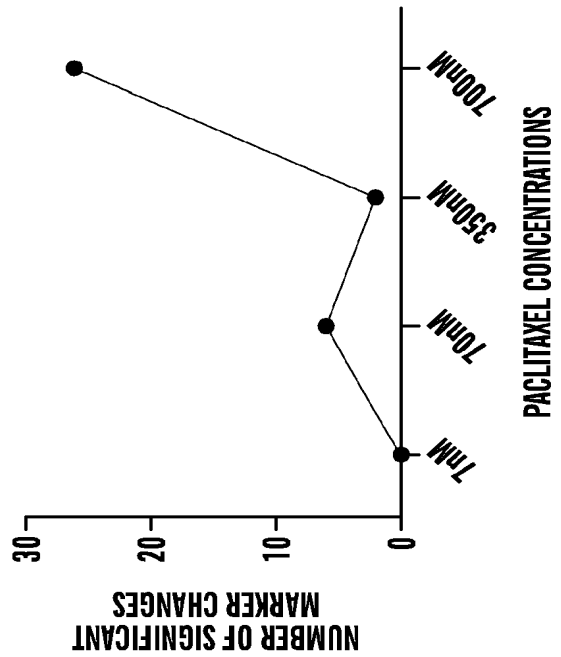


FIG. 16C

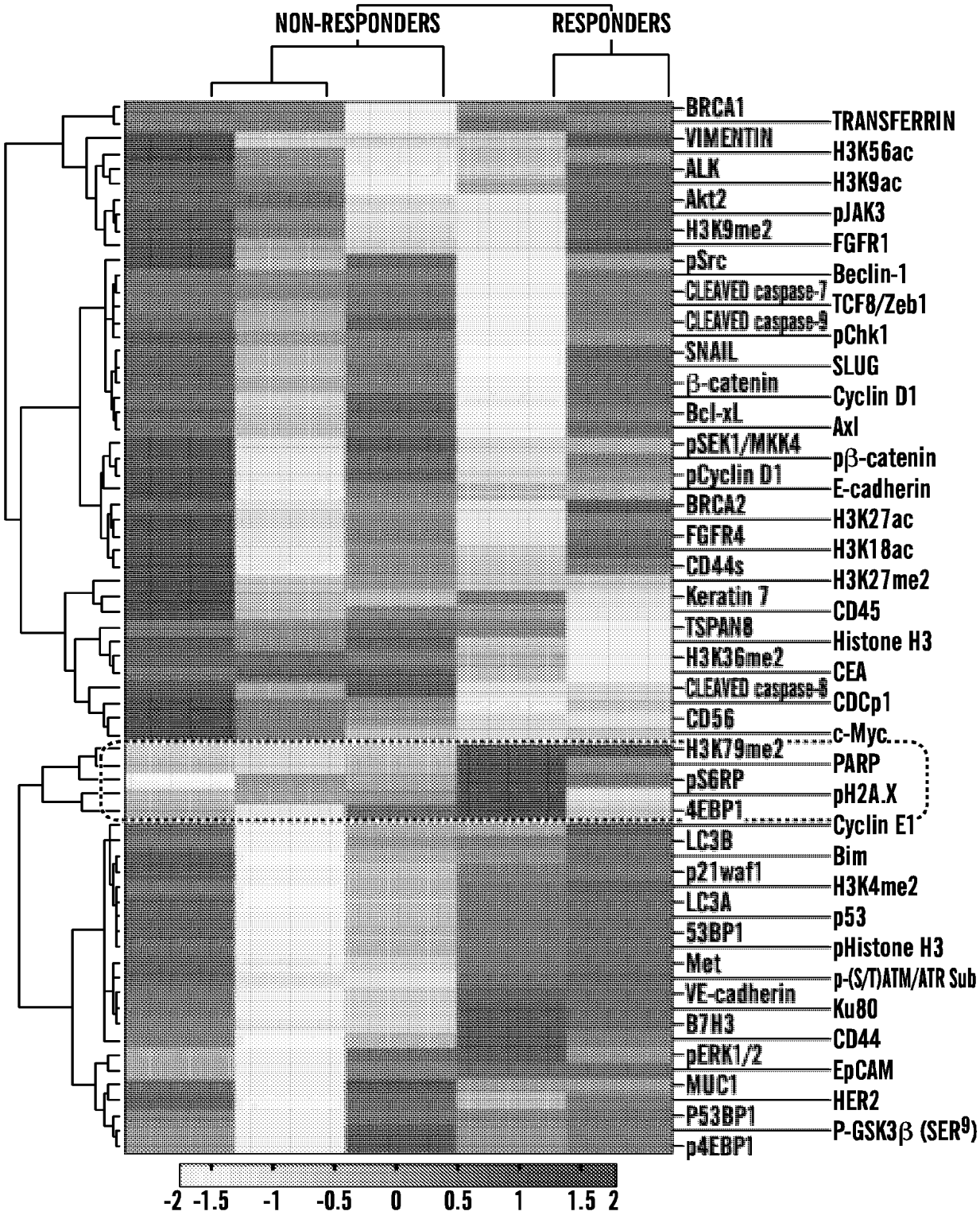


FIG. 17A

SUBSTITUTE SHEET (RULE 26)

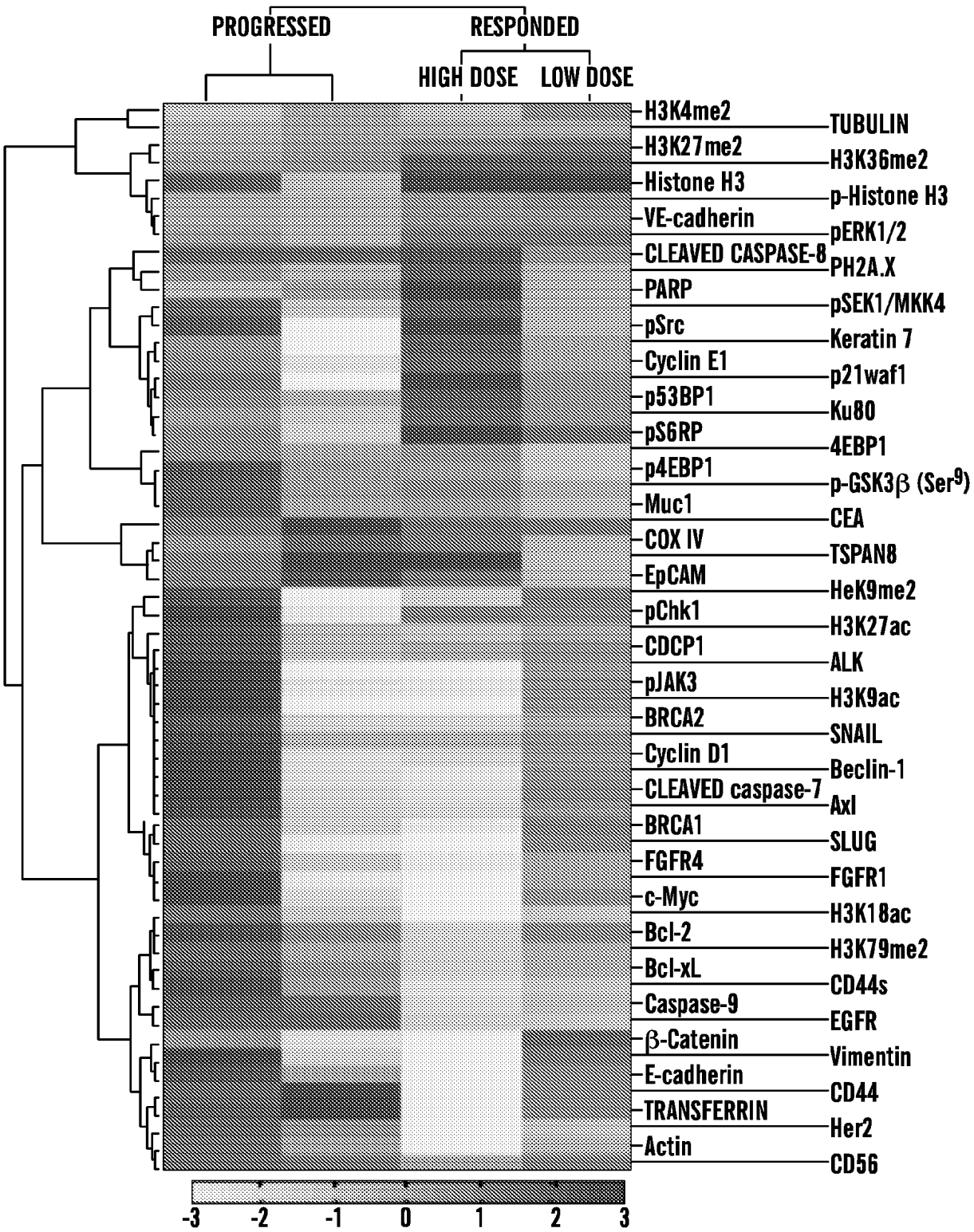


FIG. 17B

SUBSTITUTE SHEET (RULE 26)

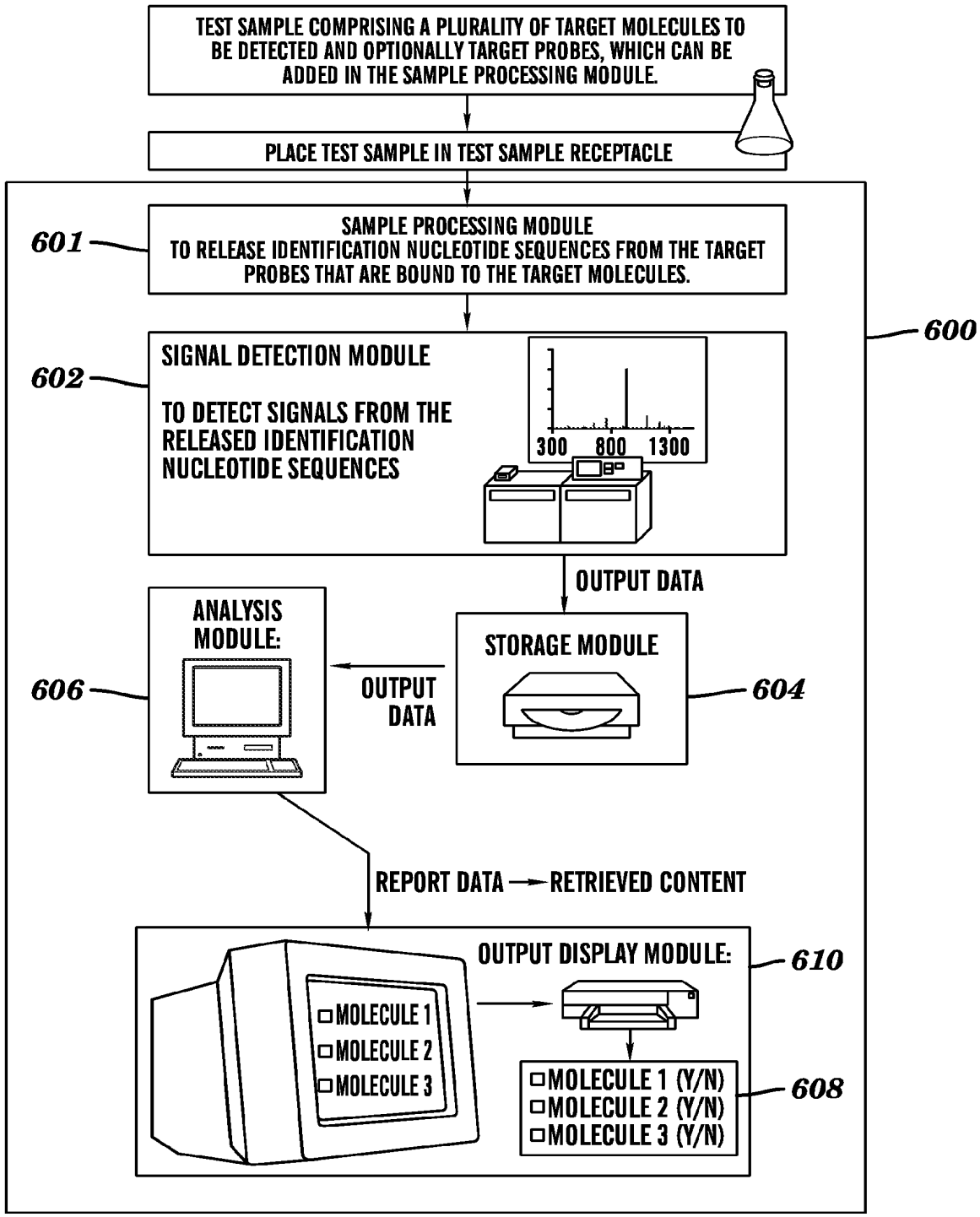

TEST SAMPLE COMPRISING A PLURALITY OF TARGET MOLECULES TO BE DETECTED AND OPTIONALLY TARGET PROBES, WHICH CAN BE ADDED IN THE SAMPLE PROCESSING MODULE.
PLACE TEST SAMPLE IN TEST SAMPLE RECEPTACLE
601
SAMPLE PROCESSING MODULE TO RELEASE IDENTIFICATION NUCLEOTIDE SEQUENCES FROM THE TARGET PROBES THAT ARE BOUND TO THE TARGET MOLECULES.
600
602
SIGNAL DETECTION MODULE
TO DETECT SIGNALS FROM THE RELEASED IDENTIFICATION NUCLEOTIDE SEQUENCES
300 800 1300
OUTPUT DATA
ANALYSIS MODULE:
606
OUTPUT DATA
STORAGE MODULE
604
REPORT DATA
RETRIEVED CONTENT
OUTPUT DISPLAY MODULE:
610
MOLECULE 1
MOLECULE 2
MOLECULE 3
MOLECULE 1 (Y/N)
MOLECULE 2 (Y/N)
MOLECULE 3 (Y/N)
608

FIG. 18A

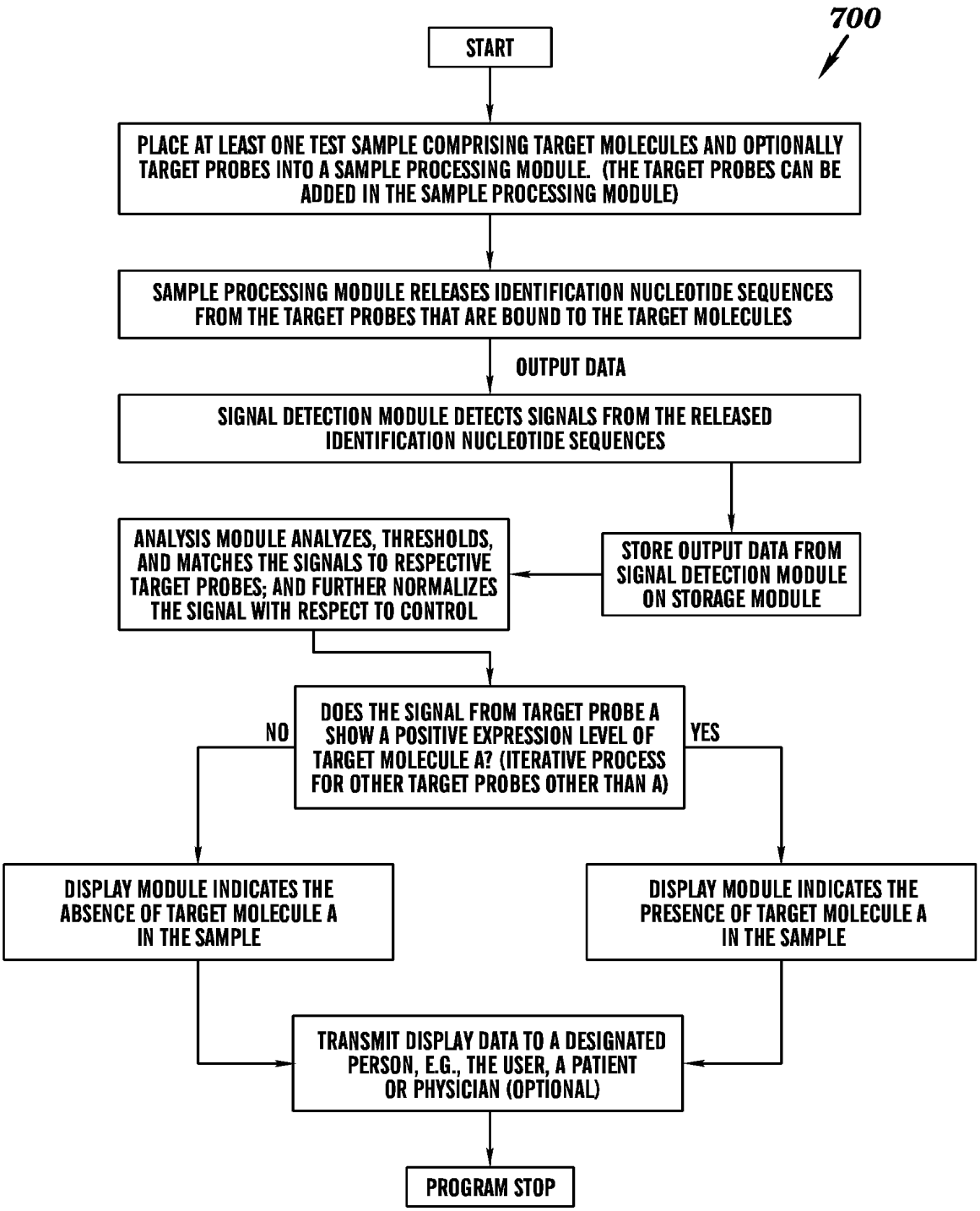

700
START
PLACE AT LEAST ONE TEST SAMPLE COMPRISING TARGET MOLECULES AND OPTIONALLY TARGET PROBES INTO A SAMPLE PROCESSING MODULE. (THE TARGET PROBES CAN BE ADDED IN THE SAMPLE PROCESSING MODULE)
SAMPLE PROCESSING MODULE RELEASES IDENTIFICATION NUCLEOTIDE SEQUENCES FROM THE TARGET PROBES THAT ARE BOUND TO THE TARGET MOLECULES
OUTPUT DATA
SIGNAL DETECTION MODULE DETECTS SIGNALS FROM THE RELEASED IDENTIFICATION NUCLEOTIDE SEQUENCES
ANALYSIS MODULE ANALYZES, THRESHOLDS, AND MATCHES THE SIGNALS TO RESPECTIVE TARGET PROBES; AND FURTHER NORMALIZES THE SIGNAL WITH RESPECT TO CONTROL
STORE OUTPUT DATA FROM SIGNAL DETECTION MODULE ON STORAGE MODULE
DOES THE SIGNAL FROM TARGET PROBE A SHOW A POSITIVE EXPRESSION LEVEL OF TARGET MOLECULE A? (ITERATIVE PROCESS FOR OTHER TARGET PROBES OTHER THAN A)
NO
YES
DISPLAY MODULE INDICATES THE ABSENCE OF TARGET MOLECULE A IN THE SAMPLE
DISPLAY MODULE INDICATES THE PRESENCE OF TARGET MOLECULE A IN THE SAMPLE
TRANSMIT DISPLAY DATA TO A DESIGNATED PERSON, E.G., THE USER, A PATIENT OR PHYSICIAN (OPTIONAL)
PROGRAM STOP

FIG. 18B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/040731

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

   a. (means)

   ☐ on paper

   ☒ in electronic form

   b. (time)

   ☐ in the international application as filed

   ☐ together with the international application in electronic form

   ☒ subsequently to this Authority for the purposes of search

2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs:1 and 2 were searched.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/040731

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 5-48, 54-66
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/040731

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12P 19/34 (2014.01)
CPC - C12Q 1/6804 (2014.09)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - C12N 15/09, 15/10; C12P 19/34; C12Q 1/68; G01N 1/00, 33/563 (2014.01)
USPC - 424/1.57; 435/6, 6.1, 6.11, 6.12, 6.18, 91.2; 536/24.31

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
CPC - C12Q 1/6804, 1/6834; 1/686; G01N 33/53, 33/533 (2014.09)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase, PubMed, Google Patents, Google

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	AGASTI et al. 'Photocleavable DNA Barcode–Antibody Conjugates Allow Sensitive and Multiplexed Protein Analysis in Single Cells,' Journal of the American Chemical Society, 23 October 2012 (23.10.2012), Vol. 134, Pgs. 18499-18502. entire document	1, 2 ------ 3, 4
X -- Y	US 2005/0079520 A1 (WU) 14 April 2005 (14.04.2005) entire document	49-53 --------- 3, 4
P, X	ULLAL et al. 'Cancer cell profiling by barcoding allows multiplexed protein analysis in fine needle aspirates,' Science Translational Medicine, 15 January 2014 (15.01.2014), Vol. 6, Iss. 219, Pgs. 1-22. entire document	1-4, 49-53
A	NAM et al. 'Nanoparticle-Based Bio–Bar Codes for the Ultrasensitive Detection of Proteins,' Science, 26 September 2003 (26.09.2003), Vol. 301, Pgs. 1884-1886. entire document	1-4, 49-53

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
03 October 2014	29 OCT 2014
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 September 2015 (03.09.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/128272 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/EP2015/053681

(22) International Filing Date:
23 February 2015 (23.02.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/944,996 26 February 2014 (26.02.2014) US

(71) Applicant: **VENTANA MEDICAL SYSTEMS, INC.** [US/US]; 1910 Innovation Park Dr., Tucson, Arizona 85755 (US).

(71) Applicant *(for CL only)*: **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacher Strasse 124, CH-4070 Basel (CH).

(72) Inventors: **ALEXANDER, Nelson**; 11440 N. Vista Ranch Place, Marana, Arizona 85658 (US). **MORRISON, Larry**; 1120 W. Moonlit Place, Oro Valley, Arizona 85737 (US). **OZTURK, Sedide**; 1500 E. Pusch Wilderness Drive, #17107, Oro Valley, Arizona 85837 (US).

(74) Agents: **GOEHRING, Frank** et al.; P.O. Box 11 52, Nonnenwald 2, 82372 Penzberg (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: PHOTO-SELECTIVE METHOD FOR BIOLOGICAL SAMPLE ANALYSIS FIELD

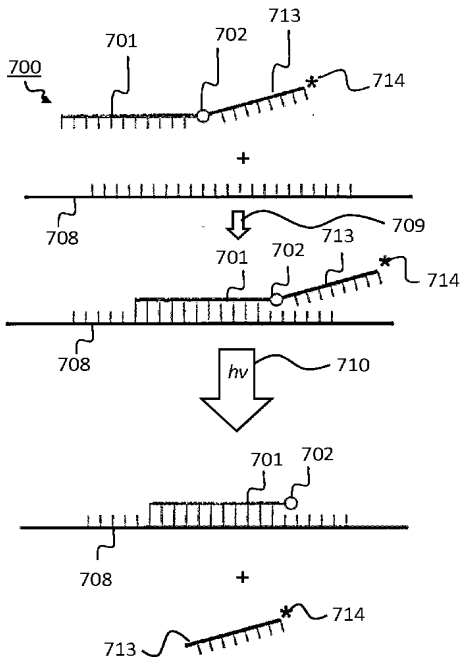


FIG. 7

(57) Abstract: The present disclosure concerns novel methods for analyzing biological samples using photo-cleavable moieties that facilitate target detection and/or biomarker isolation. In some embodiments, the method concerns using a probe comprising a photo-cleavable moiety and a selective irradiation technique for activating/cleaving the photo-cleavable moiety thereby providing facile isolation of the probe or probe components. Also disclosed herein is a system and kit for implementing the methods disclosed herein.

WO 2015/128272 A2

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

PHOTO-SELECTIVE METHOD FOR BIOLOGICAL SAMPLE ANALYSIS

FIELD

[0001] The present disclosure concerns embodiments of a method for photo-selective detection and/or isolation of a target, as well as probes and assays suitable for practicing these embodiments for biological analysis.

BACKGROUND

[0002] Detecting proteins from formalin-fixed, paraffin-embedded (FFPE) tissue in a multiplexed and quantifiable fashion is an important aspect of tissue diagnostics. Current methods, such as immunohistochemistry, have low sensitivities, narrow dynamic ranges, and limited multiplexing capabilities. Unlike nucleic acids, proteins cannot be amplified; therefore, multiplexed detection techniques and quantifying low expressing protein targets from FFPE tissue can be difficult. Also, proteins of interest are typically expressed in a sample over a wide range of locations, making simultaneous detection of multiple targets from the same tissue sample even more difficult. Highly multiplexed tests, which are useful for detecting multiple targets in a single sample (*e.g.*, a tissue sample), require sample isolation. Lysing is a typical method used to remove testable analytes from the tissue for analysis. Typical lysing methods, such as macroscopic dissection and laser capture microdissection (LCM), can be problematic as these methods disrupt overall tissue context, can cause tissue destruction, and/or require physically removing sections of the tissue from a sample. Multiplexing assays in tissue samples also are limited to analyzing a limited number of targets due to the probes and labels used in these types of assays.

[0003] Chromosomal alterations within individual cells also can be analyzed using *in situ* hybridization techniques (ISH). Currently, these analyses are limited to aberrations larger than whole genes. Particular targets or modifications (*e.g.*, SNPs, promoter methylation, miRNA, and structural alterations substantially smaller than 100kb) cannot be detected using ISH. Reference is made to “Cancer Cell Profiling by Barcoding Allows Multiplexed

Protein Analysis in Fine-Needle Aspirates" published in Science Translational Medicine, by A.V. Ullal et al. (2014), which is hereby incorporated by reference, in its entirety, for disclosure related to multiplexed oncology diagnostics. In particular, the reference describes the importance of multiplexed diagnostic methods to expand diagnostic capability in oncology. While that disclosure provides robust analysis of the medical value provided by highly multiplexed assays, the samples were limited to fine needle aspirates which are divorced from contextual information as described herein. A need exists in the art for biological assays that produce quantifiable data and that can combine sensitive and multiplexed target analysis using next generation sequencing, PCR, and/or array technology in a variety of biological samples.

SUMMARY

[0004] Disclosed herein are probes for analyzing a biological sample, and methods of using such probes. The methods disclosed herein can be used to perform contextual molecular diagnostics of various different types of samples. In one disclosed method, contextual molecular diagnostics is performed by contacting a sample with a probe comprising a photo-cleavable moiety, a specific binding moiety, and a detection tag using conditions sufficient to facilitate binding of the specific binding moiety to a target in the sample; removing probe that does not bind to the sample; selecting a region of the sample for irradiation based on contextual information; irradiating the selected region of the sample with light of a wavelength and an intensity sufficient to cleave the photo-cleavable moiety, thereby freeing the detection tag from the selected region of the sample; and detecting the detection tag. This method may further include contacting that includes multiple probes having specific binding moieties to multiple targets, the multiple different probes comprising unique detection tags and unique specific binding moieties and detecting the detection tag includes detecting the multiple unique detection tags. Illustratively, the method may further include selecting additional regions so that the method further includes iteratively and separately irradiating the

additional regions. The method may further include iteratively depositing the detection tags from the additional regions in additional separate reaction vials. The additional separate reaction vials may contain PCR primers with distinct ID sequences. As such, the method may include appending the PCR primers to the detection tags. Illustratively, the method may include iteratively and separately detecting the detection tag from the additional selected regions. The method may further comprise relating the detecting of the detection tag from the selected regions to the sample so as to create a molecular profile of the sample, wherein the molecular profile relates each selected area with the detection tags detected from the selected area across the sample.

[0005] In another disclosed method, contextual molecular diagnostics can be performed by contacting a sample with a probe comprising a photo-reactive moiety, a specific binding moiety, and a detection tag using conditions sufficient to facilitate binding of the specific binding moiety to a target in the sample; removing probe that does not bind to the sample; selecting a region of the sample for irradiation based on contextual information; irradiating the selected region of the sample with light of a wavelength and an intensity sufficient to cause the photo-reactive moiety to react, thereby freeing the detection sequence for further reaction; exposing the tissue sample to an enzyme that acts on the detection oligonucleotide to effect a change; and detecting the change. In one embodiment, the detection tag is a polymerase chain reaction primer, the photo-reactive moiety is a photo-cleavable blocking group, the enzyme is a polymerase, and the change is an extension of the primer sequence. In another embodiment, the method further includes ligating adjacent detection tags after irradiating the selected region.

[0006] In a disclosed system for performing contextual diagnostics using a photo-activated probe, the systemincludes a slide configured to accept a sample; a vessel containing the photo-activated probe; a slide imaging device configured to receive the slide, and a light source configured to selectively irradiate the sample. The photo-activated probe comprises a specific binding moiety, a photo-activated moiety, and a detection tag, wherein the photo-

activated probe is configured to be detectable upon selective irradiation. In a disclosed kit for performing contextual diagnostics, the kit includes a probe reagent solution comprising a specific binding moiety, a photo-cleavable moiety, and a detection tag.

[0007] The foregoing and other objects, features, and advantages of the invention will become more apparent from the following detailed description, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. 1(A) – (E) are a schematic drawings of a probe showing (A) a specific binding moiety and a detection tag coupled through a photo-reactive linker, (B) a probe in which the specific binding moiety is an anti-body, and (C) a probe in which the specific binding moiety is an oligonucleotide/polynucleotide, (D) a specific binding moiety, a detection tag, and a photo-reactive moiety coupled thereto, and (E) a probe in which the specific binding moiety is an oligonucleotide/polynucleotide.

[0009] FIG. 2(A) – (B) show (A) a schematic representation of a detection tag with various identification components coupled to a photo-cleavable moiety and (B) a sequence (SEQ ID NO: 1) representation of a detection tag showing exemplary identification parts.

[0010] FIG. 3 is a schematic drawing showing a method of performing contextual molecular diagnostics.

[0011] FIG. 4 is a photomicrograph showing an approach to selecting two regions of a sample.

[0012] FIG. 5(A) – (C) are schematic drawings showing the use of (A) a microscope objective, (B) a liquid crystal screen, and (C) a digital micromirror device to selectively irradiate a region of a sample.

[0013] FIGS. 6(A) – 6(F) are schematic drawings illustrating embodiments of a disclosed method, wherein (A) illustrates a probe bound to a sample prior to exposure to a light exposure step illustrated in (B). FIG. 6(C) illustrates subsequent probe extension and (D) shows the extended probe being

subjected to next generation sequencing. FIG. 6(E)-(F) show an alternative light-activation approach analogous to FIG. 6(A)-(B).

[0014] FIG. 7 is a schematic drawing illustrating an embodiment of a disclosed method whereby a probe comprising an oligo/polynucleotide, a photo-cleavable moiety, and a tag sequence modified with a detectable label is added to a sample and irradiated with light.

[0015] FIG. 8 is a schematic drawing illustrating an embodiment of a disclosed method whereby a caged ATP moiety is used so that irradiating with light liberates the caged ATP and enables detection.

[0016] FIG. 9 is a schematic drawing illustrating an embodiment of a disclosed method whereby a ligase joins two different adjacent probes bound to targets within a sample.

[0017] FIG. 10 is a schematic drawing illustrating a method for performing contextual tissue diagnostics which shows iterative irradiation of a sample and separate collection of detection tags therefrom into distinct vessels, which may or may not include location tags.

[0018] FIG. 11 is a photographic image illustrating results obtained from a control sample used in an embodiment of a disclosed method.

[0019] FIG. 12 is a photographic image illustrating selective irradiation of a sample.

[0020] FIG. 13 is a photographic image illustrating selective irradiation of a sample.

[0021] FIG. 14 is an exploded view of the photographic image of FIG. 13, further illustrating the difference in signal obtained from a sample that has been selectively irradiated.

[0022] FIG. 15 is a photographic image of a patterned sample obtained by selectively irradiating a slide with a digital micromirror device and a patterned mask.

[0023] FIG. 16 is an exploded view of the image illustrated in FIG. 15.

[0024] FIG. 17 is an image of an electrophoresis gel illustrating the detection of unique sequence identifiers after photoselective cleavage from a specific

binding moiety.

[0025] FIG. 18 is a schematic drawing illustrating an *in situ* polymerase extension embodiment whereby reverse primers are used to detect regions of the extended probe that are further away from the initial probe sequence.

[0026] FIG. 19 is an image of an electrophoresis gel analyzing isolated sequences obtained using *in situ* polymerase extension.

[0027] FIG. 20 is an image of an electrophoresis gel analyzing isolated CHR17 sequences obtained using *in situ* polymerase extension.

[0028] FIG. 21 is an image of an electrophoresis gel analyzing isolated sequences obtained from extension of two different probes within the same sample using *in situ* polymerase extension.

[0029] FIG. 22(A) – (B) is a (A) a schematic drawing illustrating an *in situ* polymerase extension embodiment and (B) is an image of an electrophoresis gel analyzing isolated sequences therefrom.

[0030] FIG. 23(A) – (B) are images of an electrophoresis gel analyzing isolated sequences obtained for BRAF Exon 13 (A) and for HER2 Exon 10 (B) probes extended with and without caged thymidines.

DETAILED DESCRIPTION

I. Terms

[0031] Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes VII*, published by Oxford University Press, 2000 (ISBN 019879276X); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Publishers, 1994 (ISBN 0632021829); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by Wiley, John & Sons, Inc., 1995 (ISBN 0471186341); and other similar references.

[0032] As used herein, the singular terms “a,” “an,” and “the” include plural referents unless context clearly indicates otherwise. Similarly, the word “or” is intended to include “and” unless the context clearly indicates otherwise. Also,

as used herein, the term "comprises" means "includes." Hence "comprising A or B" means including A, B, or A and B. Nucleotide sizes or amino acid sizes, and molecular weight or molecular mass values, stated for nucleic acids or polypeptides or other compounds are generally approximate, and are provided for description.

[0033] Concentrations, amounts, and other numerical data may be expressed or presented herein as a range. Such ranges are used merely for convenience and brevity and should be interpreted to include not only the numerical values explicitly recited as the range limits, but also to include all individual numerical values or sub-ranges encompassed within that range as if each numerical value and sub-range is explicitly recited. As an illustration, a numerical range of "about 1 to about 5" includes not only the explicitly recited values of about 1 to about 5, but also values and sub-ranges within the indicated range, such as 2, 3, and 4 and sub-ranges such as from 1-3, from 2-4, and from 3-5, etc., as well as 1, 2, 3, 4, and 5 individually. This same principle applies to ranges that recite only one numerical value as a minimum or a maximum. Furthermore, such an interpretation should apply regardless of the breadth of the range or the characteristics being described.

[0034] Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

[0035] The following explanations of specific terms are provided to facilitate review of the various disclosed embodiments:

[0036] **"Antibody"** occasionally abbreviated "Ab," refers to immunoglobulins or immunoglobulin-like molecules (including by way of example and without limitation, IgA, IgD, IgE, IgG and IgM, combinations thereof), similar molecules produced during an immune response, (*e.g.*, in

mammals such as humans, goats, rabbits and mice) as well as antibody fragments, that specifically bind to a molecule of interest (or a group of highly similar molecules of interest) to the substantial exclusion of binding to other molecules. For example, antibodies and antibody fragments typically have a binding constant for the molecule of interest that is at least 10^3 M^{-1} greater, at least 10^4 M^{-1} greater, or at least 10^5 M^{-1} greater than a binding constant for other molecules in a biological sample. "Antibody" further refers to a polypeptide ligand comprising at least a light chain or heavy chain immunoglobulin variable region that specifically recognizes and binds an epitope of an antigen. Antibodies may comprise of a heavy and a light chain, each of which has a variable region, termed the variable heavy (VH) region and the variable light (VL) region. Together, the VH region and the VL region are responsible for binding an antigen recognized by the antibody. "Antibody" also includes intact immunoglobulins and the variants and portions of them well known in the art. Antibody fragments include proteolytic antibody fragments, such as F(ab')2 fragments, Fab' fragments, Fab'-SH fragments and Fab fragments as are known in the art; recombinant antibody fragments, such as sFv fragments, dsFv fragments, bispecific sFv fragments, bispecific dsFv fragments, F(ab)'2 fragments; single chain Fv proteins ("scFv"); disulfide stabilized Fv proteins ("dsFv"); diabodies; triabodies (as are known in the art); and camelid antibodies (see, for example, U.S. Patent Nos. 6,015,695; 6,005,079, 5,874,541; 5,840,526; 5,800,988; and 5,759,808). "Antibody" also includes monoclonal antibodies, which are characterized by being produced by a single clone of B lymphocytes or by a cell into which the light and heavy chain genes of a single antibody have been transfected. Monoclonal antibodies are produced by methods known to those of skill in the art. Monoclonal antibodies include humanized monoclonal antibodies.

[0037] **"Conjugate"** refers to two or more moieties directly or indirectly coupled together. For example, a first moiety may be covalently or noncovalently (*e.g.*, electrostatically) coupled to a second moiety. Indirect

attachment is possible, such as by using a linker, which may be a molecule or group of atoms positioned between two moieties.

[0038] **"Contacting"** refers to placement that allows association between two or more moieties, particularly direct physical association, for example both in solid form and/or in liquid form (for example, the placement of a biological sample, such as a biological sample affixed to a slide, in contact with a composition, such as a solution containing the probes disclosed herein).

[0039] **"Context, -ual"** refers to the environment associated with a sample, wherein the integrity of, or information obtained from, a certain portion of the sample can be determined by comparing it with the entire sample environment. Contextual information is that information that can be understood by considering a sample's spatial and relational characteristics. As an example, contextual information for a tissue sample would include the type of cells and the spatial relationship between the various cells in the tissue sample. According to one aspect, pathologists gather and interpret contextual information as they examine a tissue sample microscopically or through an image using the cellular structures, morphological features of various cells and tissues, or through the interpretation of a stain (*i.e.*, immunohistochemistry (IHC) antibody stain) within the sample. The contextual information can include classification of a cell or region as tumor or normal tissue. Contextual information can also include an observation of spatial tumor heterogeneity or the recognition of an invasive margin or other biological features evident through visual inspection. Contextual information can include identification of tissue features such as the vasculature, or of immune cell infiltration. As used herein, it is understood that contextual information is destroyed in some way when a sample is lysed or ground so as to create a solution from the sample. The creation of this solution takes the various components from their biological locations and places them in a solution where at least some of the spatial relationships are lost. Contextual information can be preserved when the chemical or molecular diagnostic information can be related back to an image

or sample location. In another example, chemical or molecular imaging of a sample can be considered as preserving the contextual information of a sample.

[0040] **"Detect"** refers to the ability to determine if an agent (such as a signal or particular antigen, protein or nucleic acid) is present or absent, for example, in a sample. In some examples, this can further include quantification, and/or localization, for example localization within a cell or particular cellular compartment. "Detecting" refers to any method of determining if something exists, or does not exist, such as determining if a target molecule is present in a biological sample. For example, "detecting" can include using a visual or a mechanical device to determine if a sample displays a specific characteristic. In certain examples, light microscopy and other microscopic means are used to detect a detectable label bound to or proximally to a target. Detecting can also encompass performing sequencing, array analysis and/or PCR amplification of a detection tag.

[0041] **"Detectable Label"** refers to a molecule or material that can be detected or that can produce a detectable (such as visually, electronically or otherwise) signal that indicates the presence, number, location, and/or concentration of a target, such as a target molecule, in a sample, such as a tissue sample. When conjugated to a molecule capable of binding directly or proximally to a target, the detectable label can be used to locate and/or quantify the target. A detectable label can be detected directly or indirectly, and several different detectable labels can be used in combination to detect one or more targets.

[0042] **"Hybridization"** refers to forming base pairs between complementary regions of two strands of DNA, RNA, or between DNA and RNA, thereby forming a duplex molecule. Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method and the composition and length of the hybridizing nucleic acid sequences. Generally, the temperature of hybridization and the ionic strength (such as the Na^+ concentration) of the hybridization buffer will

determine the stringency of hybridization. The presence of a chemical which decreases hybridization (such as formamide) in the hybridization buffer will also determine the stringency (Sadhu *et al., J. Biosci.* 6:817-821, 1984). Calculations regarding hybridization conditions for attaining particular degrees of stringency are discussed in Sambrook *et al.*, (1989) Molecular Cloning, second edition, Cold Spring Harbor Laboratory, Plainview, NY (chapters 9 and 11). Hybridization conditions for ISH are also discussed in Landegent *et al., Hum. Genet.* 77:366-370, 1987; Lichter *et al., Hum. Genet.* 80:224-234, 1988; and Pinkel *et al., Proc. Natl. Acad. Sci. USA* 85:9138-9142, 1988.

[0043] **"Multiplex**, -**ed**, -**ing**" refers to detecting multiple targets in a sample, or multiple targets in multiple samples, concurrently, substantially simultaneously, or sequentially. Multiplexing can include identifying and/or quantifying multiple distinct nucleic acids (*e.g.*, DNA, RNA, mRNA, miRNA) and polypeptides (*e.g.*, proteins), both individually and in any and all combinations.

[0044] **"Oligonucleotide"** refers to a plurality of joined nucleotides joined by phosphodiester bonds. While the number of nucleotides can vary, oligonucleotides typically include from about 6 to about 500 nucleotides. "Oligonucleotide" refers to DNA oligonucleotides, RNA oligonucleotides, synthetic oligonucleotides (*e.g.*, sequences that are, or contain, non-naturally occurring DNA or RNA sequences), and oligonucleotide analogs, which are moieties that function similarly to oligonucleotides but which have non-naturally occurring portions. For example, oligonucleotide analogs can contain non-naturally occurring portions, such as altered sugar moieties or inter-sugar linkages, such as a phosphorothioate oligodeoxynucleotide. Functional analogs of naturally occurring polynucleotides can bind to RNA or DNA, and include peptide nucleic acid molecules.

[0045] "**Proximal**" refers to being situated at or near a reference point. As used herein, proximal means within about 5000 nm, within about 2500 nm, within about 1000 nm, within about 500 nm, within about 250 nm, within about 100 nm, within about 50 nm, within about 10 nm, or within about 5 nm of the

reference point.

[0046] "**Sample**" refers to a biological specimen containing biological molecules, such as genomic DNA, RNA (including mRNA), genes, amino acids, peptides, proteins, or combinations thereof, obtained from a subject. Examples include, but are not limited to, blood (*e.g.*, peripheral blood), urine, saliva, tissue biopsy, surgical specimen, amniocentesis samples and autopsy material.

[0047] "**Specific binding moiety**" refers to a moiety/compound/sequence that specifically binds to a second member of a specific binding pair or a target to the substantial exclusion of binding to other potential binding pairs or non-targeted areas of a sample. Specific binding pairs are pairs of molecules that bind each other to the substantial exclusion of binding to other molecules (for example, specific binding pairs can have a binding constant that is at least 10^3 M^{-1} greater, 10^4 M^{-1} greater or 10^5 M^{-1} greater than a binding constant for either of the two members of the binding pair with other molecules in a biological sample). Exemplary specific binding moieties include, but are not limited to, proteins (*e.g.*, antibodies or a binding variable region thereof), polynucleotides (*e.g.*, DNA, RNA, etc.), nucleotides, oligonucleotides, polypeptides, peptides, amino acids, aptamers, a member of a specific binding pair, a primer, a plasmid, or combinations thereof.

[0048] **"Target nucleic acid sequence or molecule"** refers to a defined region or particular portion of a nucleic acid molecule, for example a portion of a genome (such as a gene or a region of mammalian genomic DNA containing a gene of interest). In an example where the target nucleic acid sequence is a target genomic sequence, such a target can be defined by its position on a chromosome (*e.g.,* in a normal cell), for example, according to cytogenetic nomenclature by reference to a particular location on a chromosome; by reference to its location on a genetic map; by reference to a hypothetical or assembled contig; by its specific sequence or function; by its gene or protein name; or by any other means that uniquely identifies it from among other genetic sequences of a genome. In some examples, the target nucleic acid

sequence is mammalian genomic sequence (for example human genomic sequence).

[0049] In some examples, alterations of a target nucleic acid sequence (*e.g.*, genomic nucleic acid sequence) are "associated with" a disease or condition. In some examples, detection of the target nucleic acid sequence can be used to infer the status of a sample with respect to the disease or condition. For example, the target nucleic acid sequence can exist in two (or more) distinguishable forms, such that a first form correlates with absence of a disease or condition and a second (or different) form correlates with the presence of the disease or condition. The two different forms can be qualitatively distinguishable, such as by polynucleotide polymorphisms, and/or the two different forms can be quantitatively distinguishable, such as by the number of copies of the target nucleic acid sequence that are present in a cell.

[0050] **"Unique sequence identifier"** refers to a nucleic acid sequence that is configured to have a sequence that is specifically correlated to a particular target and/or specific binding moiety of a probe. Each unique sequence identifier can have a sequence that differs from other unique sequence identifiers present in a sample and can therefore be used to identify a particular target and or specific binding moiety of a probe.

II. Overview

[0051] The present disclosure concerns various analytical techniques for biological samples, such as tissue and cell samples. The various analytical techniques contemplated by the present disclosure can include, but are not limited to, target detection, genetic sequencing (*e.g.*, next generation sequencing), multiplexing, amplification (*e.g.*, polymerase chain reaction, ligase chain reaction, etc.), mass spectrometry-based protein analysis, and the like. The analytical techniques disclosed herein can be facilitated by using photochemical-based technology.

[0052] The present disclosure also concerns analytical techniques and isolation techniques that address a current need in the art to provide efficient

methods for analyzing multiple targets, and in some situations, a large number of targets. Analytical and isolation techniques also are disclosed that do not destroy the sample undergoing analysis, which thereby maintains sample context integrity. For example, methods for detecting targets in a tissue sample that are currently known in the art are limited to detecting only a few targets, given the types of probes available for tissue analysis. These methods also typically require various cumbersome and tissue-destructive isolation techniques (*e.g.*, tissue lysis or laser capture microdissection) to provide isolated targets that undergo subsequent analysis. The present disclosure addresses these problems and limitations by providing probes, methods of using the probes, and selective isolation techniques exemplified by the disclosure provided herein.

III. Probes

[0053] Disclosed herein are probe embodiments for detecting one or more targets in a sample and/or providing genetic information useful for determining the presence of genetic aberrations. In some embodiments, a probe is used to facilitate analyzing formalin-fixed paraffin embedded (FFPE) samples, fresh samples, or frozen samples. The samples may be tissue samples, cell samples, or soluble protein samples. Additionally, disclosed probes can provide substantially unlimited numbers of combinations of detection tags. This unique feature is particularly suitable for detecting multiple targets in multiplexing assays.

[0054] Particular disclosed probe embodiments comprise a photo-cleavable moiety which confers particular benefits for analytical methods and isolation. For example, using disclosed photo-cleavable moieties may reduce background signals during analysis by decreasing non-specific binding event detection. For example, duplexes formed between a template sequence and a tag sequence attached to a specifically bound probe comprising a photo-cleavable moiety can be cleaved by an irradiation event and can be subsequently isolated and analyzed. In contrast, non-specifically bound template sequences (which

contribute to background noise) that hybridize with sample through a non-photo-labile interaction (*e.g.*, electrostatic and/or hydrophobic interactions) may not be cleaved by the irradiation step and therefore can potentially be sequestered on the sample rather than isolated with the desired probes that were cleaved during the irradiation event. Under this scenario, only the desired probes (or desired portions of the probe) undergo subsequent analysis (*e.g.*, amplification, genetic sequencing, array analysis, etc.). As such, non-photo-labile probes can be added to the sample to block sample sequences that are creating background issues. While certain probes described herein include a covalently bound photo-cleavable or photo-reactive moiety, the term probe is not so restricted. In other embodiments, the probe may include a photo-reactive moiety or a photo-cleavable moiety present on a separate molecule that interacts with the specific binding moiety and/or labeling components of the probe subsequent to irradiation.

A. Probe Components

[0055] Certain embodiments of the probe comprise a specific binding moiety (SBM) suitable for binding to a particular binding pair, such as a biological target. The probe also may comprise a photo-reactive moiety (PRM). In illustrative embodiments, the photo-reactive moiety is a photo-cleavable moiety. In particular, the photo-cleavable moiety may be capable of being cleaved from other probe components, such as the specific binding moiety, upon exposure to light having a suitable wavelength. Embodiments of the probe may further comprise a detection tag (TAG) suitable for downstream analysis (*e.g.*, array-based analysis, sequencing, and/or PCR amplification). Each of these components, as well as other probe components, is discussed in more detail below.

[0056] Exemplary probe embodiments are illustrated in FIGS. 1(A) – (E). FIG. 1(A) provides a generic illustration of a probe embodiment 100 comprising specific binding moiety 101, photo-cleavable moiety 102, and detection tag 103. FIG. 1(B) illustrates exemplary probe 110, wherein the specific binding moiety 101 is an antibody 111; the probe further comprising photo-cleavable moiety

102 and unique sequence identifier 113, which acts as the detection tag. Unique sequence identifier 113 may include a particular sequence, or can be configured to have a particular size (*e.g.*, certain number of nucleotides). These variable and selectable characteristics provide the ability to uniquely and discriminately detect each different unique sequence identifier effectively coupled to a sample target by a probe. FIG. 1(C) illustrates exemplary probe 120, wherein the specific binding moiety is an oligonucleotide 121; the probe further comprising photo-cleavable moiety 122 and unique sequence identifier 123.

[0057] FIG. 1(D) provides a generic illustration of a probe embodiment 130 comprising specific binding moiety 131, detection tag 132, and photo-cleavable moiety 133. FIG. 1(E) illustrates exemplary probe 140, wherein the specific binding moiety is an oligonucleotide 141; the probe further comprising unique sequence identifier 143 and photo-cleavable moiety 142. While not shown, a probe analogous to exemplary probe 140 and 110 is understood to be within the scope of the present disclosure. In particular, another exemplary probe would include an antibody specific binding moiety with a coupling pattern analogous to 130.

1. Photo-Reactive Moiety (PRM)

[0058] The photo-reactive moiety is a moiety that reacts or is cleaved by light. In some embodiments, the photo-reactive moiety is cleaved from one or more probe components. In some embodiments, a photo-cleavable moiety is capable of being chemically activated by light and being separated, typically chemically separated (such as by bond cleavage), from the component (or components) to which it is coupled. In some embodiments, the cleavage event separates portions of the probe (*e.g.,* separates the specific binding moiety from the detection tag). In other embodiments, the cleavage event cleaves a protecting group. In some embodiments, the photo-cleavable moiety is capable of absorbing light of a certain wavelength. In particular disclosed embodiments, the photo-cleavable moiety may be cleaved by ultraviolet light and/or visible light. Suitable ultraviolet light typically has a wavelength ranging

from about 200 nm to about 400 nm, such as from about 250 nm to about 375 nm, or from about 260 nm to about 365 nm, or from about 280 nm to about 350 nm. Suitable visible light typically has a wavelength ranging from about 400 nm to about 790 nm, such as from about 450 nm to about 750 nm, or from about 500 nm to about 725 nm, or from about 550 nm to about 700 nm.

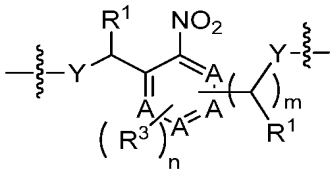
[0059] The photo-cleavable moiety used in certain probe embodiments may be selected based on considering certain physical properties of the probe and conditions associated with using the probe in a biological assay. For example, the photo-cleavable moiety may be selected based on the kinetics of the photo-based cleavage reaction. Thus, certain exemplary criteria for selection include the rate at which the photo-cleavable moiety is photochemically cleaved, the efficiency of photochemical cleavage, its size, and its stability under the analytical/storage conditions selected. In some embodiments, the photo-cleavable moiety should be photochemically cleaved substantially immediately upon exposure to light of the appropriate wavelength, such as at within 1 minute of being exposed to light of the appropriate wavelength, and preferably no more than about to about 5 minutes after being exposed to light of the appropriate wavelength. Preferred embodiments concern using a photo-cleavable moiety that can be cleaved under 1 minute after being exposed to light of the appropriate wavelength. In some embodiments, substantially all photo-cleavable moieties exposed to light should be cleaved within this time frame. For example, from about 50% to about 100% of the photo-cleavable moieties exposed to light of the appropriate wavelength should be cleaved, with some embodiments having from about 60% to about 99% cleavage, or from about 70% to about 95% cleavage, or from about 80% to about 90% cleavage.

[0060] The probe comprising the photo-cleavable moiety should be sufficiently stable to conditions associated with a particular analytical technique to substantially preclude degradation. For example, the probe comprising the photo-cleavable moiety should be substantially non-degraded by process solutions used, pH conditions used, and/or temperatures associated with a particular technique. Furthermore, the probe and the photo-cleavable moiety

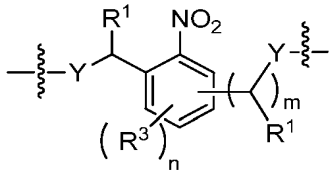
should be sufficiently robust before and after cleavage to prevent any destructive by-products from interfering with sample analysis.

[0061] In some disclosed embodiments, the photo-cleavable moiety can comprise a linker suitable for linking it with probe components disclosed herein. Suitable linkers include, but are not limited to aliphatic linkers, alkylene oxide linkers, disulfide linkers, commercially available linkers, polymeric linkers, and the like. The linker may be selected to impart increased hydrophilicity or hydrophobicity to the probe. Solely by way of example, alkylene oxide linkers disclosed herein can be used to link the photo-cleavable moiety to one or more of the probe components disclosed herein and to increase the hydrophilic nature of the entire probe.

[0062] Some embodiments of the disclosed photo-cleavable moiety have a Formula I or Formula II, provided below.



Formula I

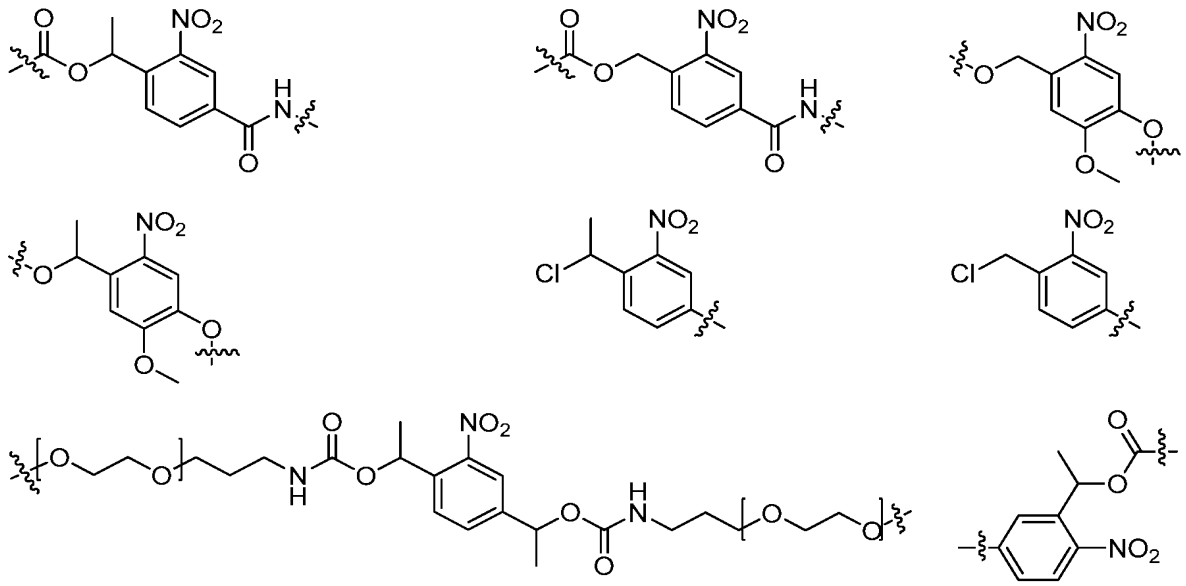


Formula II

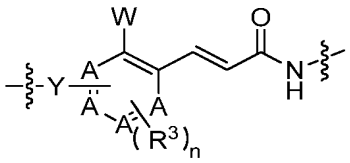
[0063] With reference to Formula I or Formula II, each R^1 independently may be halogen, hydrogen or alkyl (*e.g.*, methyl, ethyl, propyl, and the like); each R^3 independently may be selected from alkyl, aryl, hydrogen, hydroxyl, alkoxy, ether, ester, amine, amide, or halogen; each Y independently can be selected from oxygen, sulfur, NH, NR^4 (where R^4 is aliphatic, particularly alkyl), –C(O)–, –OC(O)–, –NHC(O)–, or halogen (*e.g.*, Cl, Br, Fl, or I); each A independently may be carbon or nitrogen; n can be 0, 1, 2, or 3; and m can be 0 or 1.

[0064] Exemplary species of the photo-cleavable moiety are illustrated below.

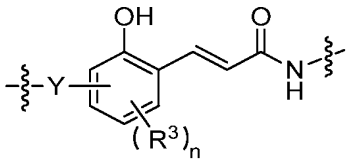
With reference to these species and the general formulas provided herein, a " ~~~ " symbol is used to indicate a truncated bond to other components to which the photo-cleavable moiety may be bound (*e.g.*, a specific binding moiety, a linker, a detection tag, etc.).



[0065] In some embodiments, the photo-cleavable moiety may have Formula III or Formula IV, provided below.



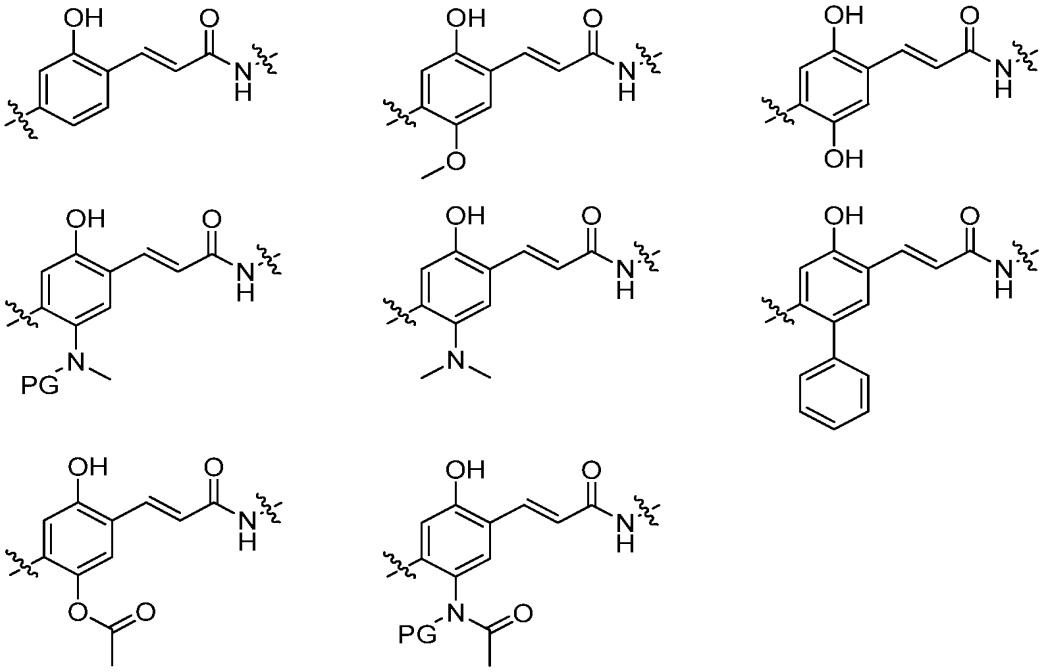
Formula III



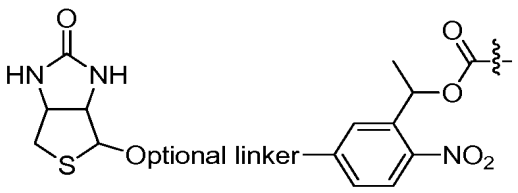
Formula IV

[0066] With reference to both Formula III and Formula IV, each of R^3, Y, and n may be as indicated above for Formula I. The "W" variable of Formula III can be selected from OH or $N(R^4)_2$ (where at least one of R^4 is hydrogen and the

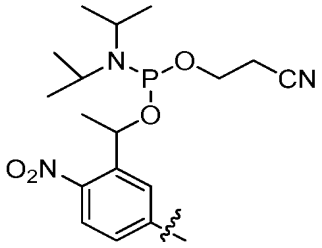
other R^4 moiety may be hydrogen or aliphatic, particularly alkyl). Exemplary embodiments of a photo-cleavable linker having a Formula II are provided below. With reference to these examples, PG refers to an amine protecting group, which would be readily recognized by persons of ordinary skill in the art.



[0067] The photo-cleavable moiety may include or be bound to, either directly or indirectly, additional functional groups or moieties suitable for incorporating and/or facilitating the use of the photo-cleavable moiety in the methods disclosed herein. For example, the photocleavable moiety may be bound to a phosphoramidite to facilitate incorporation of the photo-cleavable moiety into a nucleic acid specific binding moiety. In other embodiments, the photocleavable moiety may be bound to a detectable label, such as a biotin molecule. In particular disclosed embodiments, the photo-cleavable moiety is selected from a phosphoramidite-containing photo-cleavable moiety, a biotin-containing photo-cleavable moiety, and/or an amine-containing photo-cleavable moiety. Exemplary embodiments comprising a photo-cleavable moiety are provided below.



Biotin-containing photo-cleavable moiety

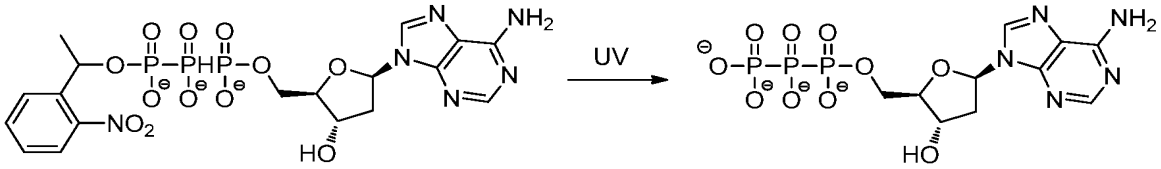


Phosphoramidite-containing photo-cleavable moiety

[0068] In some embodiments, the photo-cleavable moiety may act as a linker to couple two different components together. For certain exemplary probe embodiments, the photo-cleavable moiety couples a specific binding moiety to another probe component, such as a detection tag. In other disclosed embodiments, such as when the specific binding moiety is a primer sequence, the photo-cleavable moiety may serve as a polymerase blocking moiety on the 3' end of the primer sequence, the 5' end of the primer sequence, or both; but more typically the 3' end of the primer sequence. The photo-cleavable moiety can thereby block the 3' end of the primer sequence from being extended by a polymerase enzyme on those particular portions of the sample that are not exposed to light of a suitable wavelength. Only those primers present on the irradiated portion of the sample are extended by the polymerase enzyme as the blocking portion is removed and the primer is therefore capable of being chemically modified by a polymerase enzyme. In another embodiment, caged thymidines can be used analogously. It was observed that caged thymidines can be used to inhibit *in situ* polymerase extension by interfering with base-pairing.

[0069] Photo-cleavable moieties also may be present on a component other than the probe. For example, particular disclosed embodiments of the method use photo-cleavable moieties to block, or cage, ATP molecules – that is, the

photo-cleavable moiety prevents ATP from being used in an ATP-dependent enzymatic reaction until a sample is irradiated with light of a suitable wavelength to cleave the photo-cleavable moiety and release the ATP molecule. An exemplary blocked ATP compound (**100**) is provided below in Scheme 1. Upon exposure to light of a suitable wavelength, ATP compound **102** is obtained.



100 **102**

Scheme 1

[0070] Reference is made to U.S. Patent No. 7,432,368 and U.S. Patent Publication No. 2014/0051605, which is incorporated by reference herein in its entirety for disclosure related to methods and compositions for photo-cleavable moieties.

[0071] Similarly, in another embodiment caged deoxynucleoside triphosphates are employed to prevent extension of a primer hybridized to a target nucleic acid sequence by a polymerase enzyme (*e.g.*, DNA polymerase and its variants, when the target is DNA, or reverse transcriptase when the target is RNA). All four deoxynucleoside triphposphates may be caged or as few as a single nucleoside trip may be caged with the remaining deoxynucleoside triphosphases being uncaged to prevent significant primer extension in the absence of illumination. An example of a caged deoxynucleoside triphosphate is caged dATP, an example of which is pictured in scheme 1 except that a deoxyribose sugar would replace the ribose sugar moiety. An example of synthesizing caged dATP is described in *The Development of Methods for the Time-Resolved Imaging of the Replication of Single DNA Molecules*, Richard David Perrins Ph.D. thesis, the University of Birmingham, August 2005. The preparation of caged dideoxyribosylthymine

triphosphate, dideoxyadenosine triphosphate and arabinosylcytosine triphosphate is described by Meldrum *et al.* in "Kinetics and mechanism of DNA repair, *Biochem. J.* (1990) 266, 885-890. Walker *et al.* describe "Photolabile 1-(2-Nitropheny1)ethyl Phosphate Esters of Adenine Nucleotide Analogues. Synthesis and Mechanism of Photolysis" in *J. Am. Chem. Soc.* 1988, 110, 7170-7177. In another reference, Meldrum *et al.* also describe the "Use of Caged Compounds in Studies of the Kinetics of DNA Repair" in *METHODS IN ENZYMOLOGY*, VOL. 291, 1998.

2. Specific Binding Moieties

[0072] Specific binding moieties can be selected from an amino acid, a peptide, a protein (*e.g.*, an antibody or a fragment thereof), a nucleic acid, a nucleotide, an oligonucleotide, a polynucleotide (*e.g.*, DNA, RNA, miRNA, etc.), a member of a specific binding pair (*e.g.*, biotin, avidin, streptavidin, etc.), a primer, an aptamer, a plasmid, or combinations thereof. Suitable antibodies for use as specific binding moieties include, but are not limited to, primary antibodies and/or secondary antibodies. For example, the antibody may be a primary antibody that binds directly to a target in a sample, or it may be a secondary antibody (*e.g.*, anti-species antibody) that binds indirectly to a target in a sample through another antibody. The antibody also can be a monoclonal antibody. Suitable exemplary specific antibodies include, but are not limited to, a HER-2/neu rabbit monoclonal antibody, a PR rabbit monoclonal antibody, a topoisomerase IIa rabbit monoclonal antibody, an ER rabbit monoclonal antibody, a Ki-67 rabbit monoclonal antibody, a p53 antibody, and any other antibody that may be suitable for use in analyzing a sample having known or hereafter determined targets, including any of the particular target sequences disclosed herein. The probes disclosed herein are not limited to whole antibodies as specific binding moieties and may comprise an antibody fragment, such as a binding variable region.

[0073] Suitable oligo/polynucleotides for use as specific binding moieties include, but are not limited to, those suitable for hybridizing with a particular

target sequence (or target sequence fragment). In particular disclosed embodiments, the specific binding moiety may be an oligo/polynucleotide suitable for complimentary binding with DNA or RNA. In other disclosed embodiments, the specific binding moiety may be a primer, such as a DNA primer or RNA primer that comprises at least one end suitable for extension via polymerase-mediated extension. For example, the primer may comprise a free hydroxyl group which allows extension of the primer using a polymerase enzyme. Other suitable specific binding moieties include aptamers and/or plasmids.

[0074] Members of a specific binding pair suitable for practicing the disclosed embodiments include, but are not limited to, those members that are typically employed in immunoassays and/or hybridization procedures. Particular disclosed embodiments of the specific binding pair may include avidin, biotin, streptavidin, and the like.

[0075] Certain probe embodiments concern specific binding moieties that are oligo/polynucleotides capable of binding specifically to a target sequence, or a portion of a target sequence, or antibodies capable of binding to an epitope in a target protein. Exemplary target sequences are provided herein, with certain targets of interest being selected from BCL2 (18q21.3; *e.g.*, GENBANK™ Accession No. NC_000018, complement, nucleotides 58941559 59137593), TOP2A (17q21-q22; *e.g.*, GENBANK™ Accession No. NC_000017, complement, nucleotides 35798321 35827695), ERG (21q22.3; *e.g.*, GENBANK™ Accession No. NC_000021, complement, nucleotides 38675671 38955488), PTEN (10q23.3; *e.g.*, GENBANK™ Accession No. NC_000010, nucleotides 89613175 89716382), p53 (17p13.1; e.g., GENBANK™ Accession No. NC_000017, complement, nucleotides 7512464 7531642), EGFR gene (7p12; *e.g.*, GENBANK™ Accession No. NC_000007, nucleotides 55054219 55242525), or any one of SEQ ID NOs: 8-19.

[0076] The oligo/polynucleotide specific binding moiety may have any sequence of nucleotides capable of complementarily binding to the sequences, or portion of the sequences, of these targets, and may comprise at least 5

nucleotides to about 5,000 nucleotides (or such as about 5 to about 4,500 nucleotides, about 10 to about 4,000 nucleotides, about 20 to about 3,500 nucleotides) that bind to the target sequence (or portion of the target sequence). The oligo/polynucleotide specific binding moiety can also comprise a photo-cleavable moiety (such as a photo-cleavable phosphoramidite as disclosed herein). These embodiments of the specific binding moiety can further comprise a unique sequence identifier comprising any one of (or more) of the following: an adapter sequence, an alignment sequence, a reagent barcode sequence, with the reagent barcode sequence comprising at least 6 to about 200 nucleotides selected to have a random sequence unique to the particular oligo/polynucleotide specific binding moiety present in the probe, and/or a subject index sequence.

[0077] Embodiments of a probe comprising an antibody specific binding moiety may have the following exemplary conjugated structures: HER (such as HER1, HER2, HER3, or HER4) rabbit monoclonal antibody – photo-cleavable moiety – unique sequence identifier; PR rabbit monoclonal antibody– photo-cleavable moiety – unique sequence identifier; TOP2a rabbit monoclonal antibody– photo-cleavable moiety – unique sequence identifier; ER rabbit monoclonal antibody– photo-cleavable moiety – unique sequence identifier; Ki-67 rabbit monoclonal antibody– photo-cleavable moiety – unique sequence identifier; p53 antibody– photo-cleavable moiety – unique sequence identifier. The unique sequence identifier may have any suitable sequence, such as one of the exemplary sequences of SEQ ID NOs: 1-3 or 21-23. In other embodiments, each unique sequence identifier provided in these conjugates may be replaced with a tag sequence.

[0078] Exemplary embodiments of an oligo/polynucleotide sequence that may be used as a specific binding moiety in the disclosed probe include a chromosome 17 alpha satellite repeat sequence, such as SEQ ID NO: 4; an ALU repetitive sequence, such as SEQ ID NO: 5, or a chromosome 7 alpha satellite repeat sequence, such as SEQ ID NO: 20.

[0079] Other suitable probes include DNP-tagged probes, such as DNP-

TATTTT-DNP-TATTTT, and probes comprising primer sites, such as SEQ ID NO: 6 and/or SEQ ID NO: 7. Other probes include, but are not limited to, those comprising olio/polynucleotide sequences selected from SEQ ID NOs: 34-41. In still other embodiments, the oligonucleotide specific binding moiety can be of the type disclosed in U.S. Patent Publication No. 2011/0160076, which is incorporated by reference herein in its entirety for disclosure related to methods for producing uniquely specific nucleic acid probes. In further exemplary embodiments, probes disclosed in U.S. Ser. No. 61/645,247 or in International Publication No. WO 2013/167387, both of which are incorporated herein by reference in their entirety for disclosure related to uniquely specific probes. In still other embodiments, the oligonucleotide specific binding moiety can be of the type disclosed in U.S. Patent No. 8,420,798 or U.S. Patent No 7,869,959, which are incorporated by reference herein in their entirety for disclosure related to methods for selecting and/or producing nucleic acid probes.

[0080] As a specific example, the probe may comprise a primary antibody, 4B5 (available from Ventana Medical Systems, Inc.), which is capable of recognizing an epitope in the Her2 intracellular domain, such as amino acids 1231-1250 of SEQ ID NO: 10 (GAPPSTFKGTPTAENPEYLG), in the intracellular domain. The probe further comprises a photo-cleavable moiety bound to the antibody. A unique sequence identifier having any one of the sequences described in SEQ ID NOs: 1-3 or 21-23 (or a tag sequence) also is bound to the photo-cleavable moiety.

3. Detection Tag

[0081] Certain disclosed embodiments of the probe comprise a detection tag. In some embodiments, the detection tag can be a unique sequence identifier. In other embodiments, the detection tag can be a tag sequence.

[0082] Each unique sequence identifier has a unique sequence of nucleotides. The sequence can be designed to provide desired information. In some embodiments, the number of nucleotides in a unique sequence identifier can range from greater than one nucleotide to at least about 100 nucleotides, more typically from about 3 to about 150 nucleotides, even more typically from

about 50 to about 90 nucleotides or from about 18 to about 25 nucleotides. In one embodiment, the barcode sequences are at least 18 nucleotides, while tag sequences for NGS (index tags) are typically 4-6 nucleotides. The individual sequence of each unique sequence identifier provides the ability to readily discriminate between a potentially infinite number of unique sequence identifiers present in a sample based on the sequence and/or size of the unique sequence identifier. For example, in a sample comprising mixed cell types (*e.g.*, human and mouse cells), the unique sequence identifiers can be used to distinguish the cell types, as human cells can be detected with a probe comprising a first specific unique sequence identifier, and the mouse cells detected with a second, completely different unique sequence identifier. The unique sequence identifier may be amplified to improve detection using techniques known to those of ordinary skill in the art, such as by using the polymerase chain reaction (PCR). The unique sequence identifiers also may be analyzed for encoded information after a desired target is detected. In some embodiments, the unique sequence identifier may be sequenced, such as by using next generation sequencing. In other embodiments, the unique sequence identifier may be identified using array hybridization, such as by using a microarray (*e.g.*, DNA microarray).

[0083] FIG. 2(A) – (B) provides (A) a schematic representation of a detection tag and (B) an illustrative unique sequence identifier contemplated by the present disclosure. Unique sequence identifier 200 can be coupled to photo-reactive moiety (PRM) 202 and comprise one or more adapter sequences (204 and 212), an alignment sequence 206, an optional patient index sequence 208, and a reagent barcode 210.

[0084] Adapter sequences 204 and/or 212 typically comprise a short nucleotide sequence that may be manipulated to facilitate analysis of the unique sequence identifier. For example, the adapter sequence may comprise a nucleotide sequence that hybridizes with a corresponding nucleic acid sequence present on a substrate used in a next-generation sequencing technique. The adapter sequence, in some embodiments, is configured for use

in next generation sequencing that utilizes solid phase amplification sequencing techniques and thus uses a solid phase amplification substrate (*e.g.*, flow cell). In other embodiments, the adapter sequence may be configured for use in emulsion PCR and thereby suitable for binding a microbead substrate. Certain disclosed adapter sequences have a sufficient number of nucleotides to hybridize with the particular next generation sequencing sample (*e.g.*, sequences included on a solid phase amplification flow cell or sequences included on a microbead) being used, such as 1 nucleotide to at least 50 nucleotides, typically about 5 nucleotides to about 40 nucleotides, and more typically from about 15 nucleotides to about 30 nucleotides.

[0085] Unique sequence identifier 200 also may comprise one or more alignment sequences, such as sequence 206. Alignment sequences facilitate accurate sequencing by providing a starting point for reading the particular unique sequence identifier. In particular disclosed embodiments, the alignment sequence is located near the 5' end of the unique sequence identifier. An alignment sequence comprises a sufficient number of nucleotides to allow accurate alignment for reading such as from about 3 nucleotides to at least about 20 nucleotides, more typically from about 5 to about 15 nucleotides.

[0086] Subject index sequences, such as sequence 208, have a particular unique nucleotide sequence that is designed to correlate to and thereby uniquely identify, the particular subject being analyzed. For example, a patient's identity may be equated with a subject index sequence and thus this sequence can be used to identify the unique sequence identifier isolated from a particular sample with a particular patient. In some embodiments, including a subject index sequence can be useful when multiple samples are analyzed at one time because a subject can be easily correlated with a particular sample. Subject index sequences comprise a unique nucleotide sequence comprising from about 3 to at least about 10 nucleotides; for example, 5 nucleotides, 6 nucleotides, 7 nucleotides, 8 nucleotides, 9 nucleotides, or 10 nucleotides.

[0087] Reagent barcodes, such as reagent bar code 210, provide a sequence

of nucleotides that is designed to correlate to each specific binding moiety used to position the probe on targets within the sample (*e.g.*, through hybridization, covalent bonding, antigen/antibody recognition, and the like). For example, a probe comprising an antibody as a specific binding moiety may be bound to a unique sequence identifier comprising a reagent barcode that comprises a sequence of nucleotides unique to (or that correlates to) that particular antibody. In other embodiments, the reagent barcode can be configured to encode other information, such as lot data, manufacturing data, date information, source information, or combinations thereof. In some embodiments, the reagent barcode sequence may comprise from about 6 nucleotides to about 200 nucleotides, such as about 6 nucleotides to about 150 nucleotides, or 6 nucleotides to about 100 nucleotides, or about 6 nucleotides to about 50 nucleotides.

[0088] As illustrated in FIG. 2(A), unique sequence identifier 200 may comprise an adapter sequence 204 and 212 at each end of the component. Typically, at least one adapter sequence (204 or 212) is bound to photo-reactive moiety 202. Photo-reactive moiety 202 may couple unique sequence identifier 200 to a specific binding moiety (not shown)) or photo-reactive moiety 202 may be a protecting group appending unique sequence identifier 200. Alignment sequence 206 typically is located next to adapter sequence 204 that is bound to photo-cleavable moiety 202, and reagent barcode 210 may be located directly next to, or adjacent to, alignment sequence 206. Subject index sequence 208 may be placed in between alignment sequence 206 and reagent barcode 210. Finally, second adapter sequence 212 typically is located next to the reagent barcode. Second adapter sequence 212 may facilitate linking unique sequence identifier 200 to the specific binding moiety.

[0089] A person of ordinary skill in the art will recognize that a suitable unique sequence identifier may comprise all components discussed above, or it may omit certain components. Each component can be positioned in the order illustrated by FIG. 2(A), or different component ordering may be used. For example, a person of ordinary skill in the art would recognize that the particular

method of next generation sequencing used may or may not require adapter sequences. Additionally, the subject index sequence is an optional component and therefore may be present or omitted. Particular embodiments of the disclosed unique sequence identifier may be generated by a computer algorithm, which is capable of generating a substantially infinite number of unique sequence identifier sequences each having a reagent barcode comprising a different sequence of nucleotides and/or combination of adapter sequences, alignment sequences, subject index sequences, and/or reagent barcode sequences.

[0090] An exemplary unique sequence identifier 215 is illustrated in FIG. 2(B). Unique sequence identifier may be coupled to a photo-cleavable linker (not shown) at one end, which is coupled, either directly or indirectly, to an adapter sequence 218. Exemplary unique sequence identifier 215 further comprises an alignment sequence 220, a subject index sequence 222, a reagent barcode sequence 224, and a second adapter sequence 226. Other exemplary unique sequence identifiers include SEQ ID NOs: 1-3 and 21-23.

[0091] In other disclosed embodiments, the detection tag may be a tag sequence. A tag sequence, which is separate and distinct from the unique sequence identifier disclosed herein, may be used to encode information concerning a specific target on a tissue sample (*e.g.*, probe identity, lot data, manufacturing data, date information and source information, and the like). The tag sequence also facilitates binding of a template strand to a sample that is suitable for PCR amplification. In some embodiments, the tag sequence may be bound to a protein target of the tissue sample through a specific binding moiety. Solely by way of example, a tag sequence may be bound to an antibody or an oligo/polynucleotide by a photo-cleavable linker, with the antibody or the oligo/polynucleotide serving as the specific binding moiety that binds to a particular target in the sample. The tag sequence has a sufficient number of nucleotides to facilitate hybridization with a template strand, such as from at least one nucleotide to about 25 nucleotides (such as at least one nucleotide to about 22 nucleotides, or at least one nucleotide to about 20 nucleotides, or at

least one nucleotide to about 18 nucleotides).

[0092] In some embodiments, the tag sequence may comprise, or be modified to further comprise, one or more flanking primer sequences. For example, in some embodiments the tag sequence comprises a flanking primer sequence before it is bound to the sample. In other embodiments, the tag sequence may be bound to the sample and then modified with a primer sequence. In yet other embodiments, the specific binding moiety itself may be a tag sequence and can comprise, or be modified to comprise, a flanking primer sequence bound to the 5' end of the tag sequence, the 3' end of the tag sequence, or both the 5' and 3' ends of the tag sequence. Each tag sequence may include the same flanking primer sequence or different flanking primer sequences. For example, in some embodiments, the same flanking primer sequence may be used for all tag sequences within the sample. This allows simultaneous amplification of all tag sequences present within a sample.

[0093] The tag sequence may be combined with a template sequence prior to analysis in some embodiments. Template sequences, which generally comprise longer strands of nucleotides, are used to facilitate amplification techniques, such as PCR. The template sequence may be, for example, an oligonucleotide having a length suitable for quantitative PCR (*e.g.*, an oligonucleotide having from about 50 nucleotides to about 100 nucleotides, or from about 60 nucleotides to about 90 nucleotides, or from about 70 nucleotides to about 80 nucleotides). In some embodiments, the template sequence can be configured to specifically hybridize (*e.g.*, through sequence complementarity) with the tag sequence to form a stable duplex. A stable duplex, as disclosed herein, typically has a melting temperature (T_m) of greater than 60 ˚C. In some embodiments, this duplex can be transferred into solution for further analysis via cleavage of a photo-cleavable moiety that binds the tag sequence to the specific binding moiety. Using this tag-template duplex probe avoids the need to attach one or more additional recognition biomolecules/peptide sequences directly to the template sequence, which typically is used in the art. For example, immuno-PCR methods used in the art

directly conjugate a template strand to an antibody that will be bound to a target in a sample. These methods however can result in reduced antibody-antigen binding as the large template strand can often interfere with this specific binding. As disclosed herein, the template sequence may be added after the specific binding moiety is bound, which thereby reduces and preferably substantially eliminates specific binding interferences that are induced by large template sequences.

[0094] In other disclosed embodiments, the tag sequence may be modified to include one or more peptide nucleic acids and/or locked nucleic acids. A peptide nucleic acid (PNA) is an artificially synthesized polymer comprising a backbone of N-(2-aminoethyl)-glycine units linked by peptide bonds to which purine and pyrimidine bases are linked. PNAs may be used to improve hybridization with a template sequence (or a tag sequence) as PNAs do not include charged phosphate groups typically found in DNA and therefore exhibit stronger hybridization with the template sequence than would a natural DNA strand. Locked nucleic acids (LNA) comprise one or more modified nucleotides that comprise a bridge in the ribose moiety of the nucleotide that locks the ribose in the 3'-endo conformation. This conformation enhances base stacking and thereby increases the ability of the tag sequence (or a template sequence) to hybridize with the template sequence (or tag sequence). Accordingly, using either PNA or LNA in the tag sequence can reduce background signals produced by non-specific binding of template sequences and further strengthen the duplex formed between the tag sequence and the template sequence.

[0095] The disclosed detection tags are not limited in the types of nucleotides that each detection tag may comprise. For example, detection tags like unique sequence identifiers and tag sequences can comprise one or more naturally occurring nucleotides, synthesized nucleotides, chemically modified nucleotides, and combinations thereof. In some embodiments, the unique sequence identifier and/or the tag sequence can comprise one or more peptide nucleic acid or locked nucleic acid components.

[0096] The base ratio of the unique sequence identifier or tag sequence may be configured to impart chemical and/or physical modifications to the structure of the unique sequence identifier and/or tag sequence. For example, the base composition present in the tag sequence may be configured to influence the thermal melting temperature of this component upon hybridization with a template sequence. That is, the base ratio of the tag sequence may be modified to promote hybridization with a template sequence, or it may be modified to reduce tag/template hybridization. The base ratio of the template also may be modified in a similar manner. Such manipulations are not limited to the detection tag and also may be applied to specific binding moieties such as an oligo/polynucleotide specific binding moiety. These specific binding moieties can be configured to have certain base ratios that promote specific or non-specific binding with a target sequence.

[0097] The base ratio of certain probe components also can be configured to convey information about the probe. For example, the base ratio of a unique sequence identifier and/or tag sequence can be configured to have a certain base ratio that can be used to represent that particular unique sequence identifier and/or tag sequence. Once a probe comprising this particular base ratio is deposited and then isolated, the particular base ratio can be used to track the origin of that particular isolated unique sequence identifier and/or tag on the original sample.

[0098] The base ratio also may be configured to facilitate isolation of the detection tag. For example, some embodiments concern isolating a detection tag for array analysis and/or next generation sequencing. In these embodiments, the base ratio of the substrate to which the detection tag binds during the array and/or next generation sequencing analysis can be configured to specifically hybridize with the complementary base ratio present in the detection tag.

[0099] The disclosed probes also may have one or more linking facilitators that facilitate linking two or more probe components. Suitable linking facilitators include, but are not limited to, functional groups typically used in the

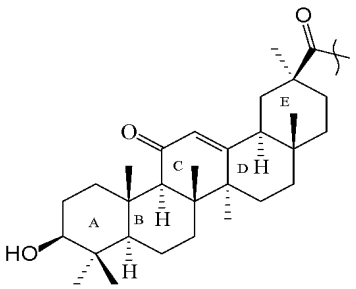
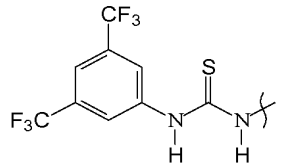
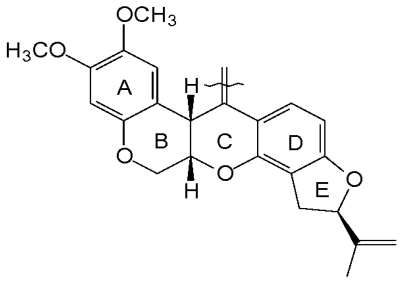
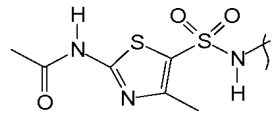
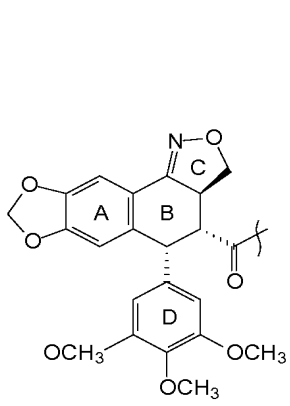
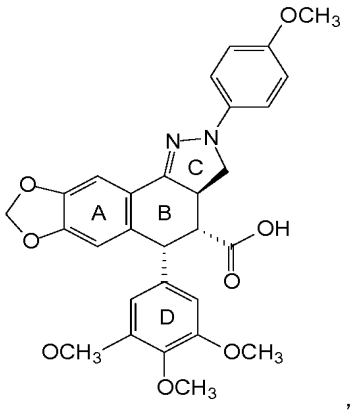
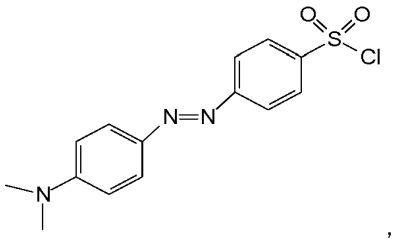
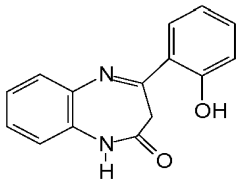
art to promote forming one or more amide bonds, ester bonds, and the like. The linking facilitator may be a chemical compound that promotes linkage between one or more probe components, but is not incorporated into the probe. Exemplary linking facilitators include, but are not limited to, carbodiimide, NHS esters, imidoesters, maleimide, haloacetyls, pyridyldisulfides, hydrazides, alkoxyamines, diazirines, aryl azides, isocyanates, and the like.

[0100] Illustrative detection tags may be selected to be uniquely distinct tags as disclosed by U.S. Patent Publ. No. 2012/0070862, which is hereby incorporated herein by reference in its entirety for disclosure related to methods for producing uniquely distinct nucleic acid tags.

4. Other Components

[0101] In certain embodiments, the probe may comprise one or more additional components other than the specific binding moiety, the photo-cleavable moiety, and the detection tag.

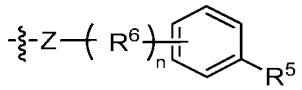
[0102] Certain probe embodiments also may comprise a detectable label, typically coupled to a selected specific binding moiety (or moieties). Particular disclosed embodiments of the detectable label may be selected from a hapten, an enzyme, a chromophore, a fluorophore (*e.g.*, a quantum dot), a member of a specific binding pair, or combinations thereof. These detectable labels can be used to facilitate active purification of the detection tag. Haptens useful in the disclosed method include, but are not limited to, oxazoles, pyrazoles, thiazoles, nitroaryls, benzofurans, triterpenes, ureas, thioureas, rotenoids, coumarins, cyclolignans, di-nitrophenyls, biotin, digoxigenin, fluoresceins, or rhodamines. Particular embodiments of the haptens that may be used in the presently disclosed embodiments are disclosed in U.S. Patent No. 7,695,929, which is incorporated herein by reference in its entirety. Exemplary hapten species include, but are not limited to, 5-nitro-3-pyrazole carbamide, 2-(3,4-dimethoxyphenyl)quinoline-4-carboxylic acid), 3-hydroxy-2-quinoxalinecarbamide, 2,1,3-benzoxadiazole-5-carbamide, and 2-acetamido-4-methyl-5-thiazolesulfonamide. Other exemplary haptens include



,

,


,

,


, and .

[0103] In particular disclosed embodiments, the detectable label may be coupled with a primer to facilitate isolation and/or purification in subsequent isolation steps. For example, the detectable label could be coupled to a nucleotide of the primer (such as through a functionalized base of the nucleotide, or via phosphoramidite-based oligonucleotide synthesis). Also, probes may comprise a detectable label bound to a tag sequence or unique sequence identifier to facilitate detection. Similar base-functionalization and/or

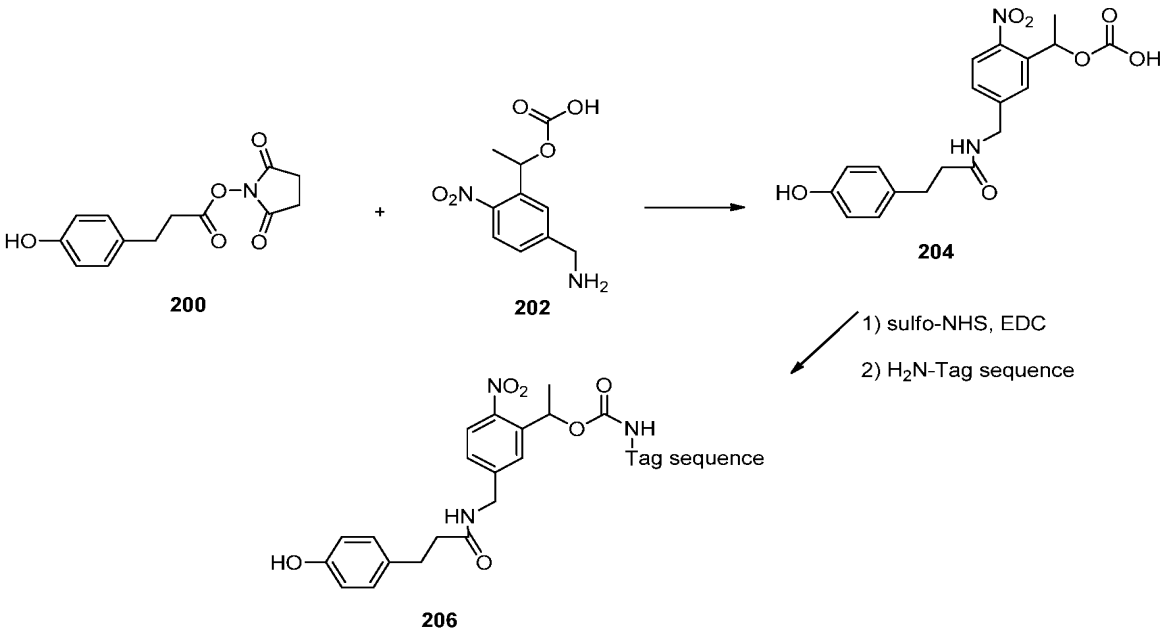
phosphoramidite-based oligonucleotide synthetic methods may be used to produce these particular probes. In some embodiments the detectable label may be attached to a nucleoside triphosphate that becomes incorporated by extension of the primer sequence by a polymerase.

[0104] In some embodiments, the probe may further comprise a signal-amplification moiety, such as a tyramine or tyramine derivative, which can be used to increase the intensity of the signal generated by a detectable label. Accordingly, tyramide signal amplification (TSA) may be used with embodiments of the disclosed probe. Particular examples of suitable tyramine derivatives, as well as methods of attaching the tyramine derivatives, are disclosed in U.S. Patent Publication No. 2013/0109019, which is incorporated herein by reference. A suitable tyramine derivative may have a formula



[0105] where R^5 is selected from hydroxyl, ether, amine, and substituted amine; R^6 is selected from alkyl, alkenyl, alkynyl, aryl, heteroaryl, $-OR_m$, $-NR_m$, and $-SR_m$, where m is 1-20; n is 1-20; Z is selected from oxygen, sulfur, or NR^a wherein R^a is selected from hydrogen, aliphatic, aryl, or alkyl aryl. The tyramine derivative may be attached to a tag sequence and a photo-cleavable moiety through the Z moiety of the tyramine derivative. In some embodiments, the Z moiety of the tyramine derivative may be coupled with a photo-cleavable moiety, which is also coupled with a tag sequence.

[0106] An exemplary synthetic scheme is provided below in Scheme 2. Tyramide compound **200** may be combined with photo-cleavable moiety **202** to form conjugate **204**. Conjugate **204** may be combined with a tag sequence having an exposed amine moiety capable of coupling with conjugate **204** to form tyramide-containing conjugate **206**. A person of ordinary skill in the art will recognize that the particular order of coupling described in Scheme 2 may be modified. For example, photo-cleavable moiety **202** could be first coupled with a tag sequence and then the signal amplification compound **200**.



Scheme 2

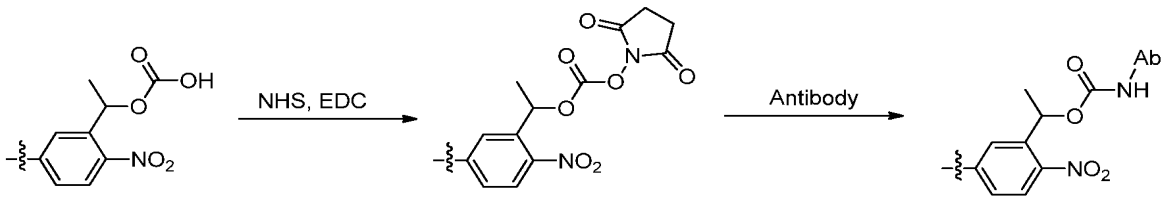
[0107] The tyramine or tyramine derivative may be used to deposit multiple tag sequences proximal to a desire target area of the sample. Adding a peroxidase, such as horseradish peroxidase, to the sample sequentially with, or simultaneously with, the tyramine-containing probe facilitates deposition of the tag sequences.

B. Making the Probe

[0108] The probes described herein may be made by any method now known or hereafter developed. The method for making the probe may be selected based, at least in part, on the particular components being included in the probe. The methods disclosed herein may be altered and modified to accommodate different components than those expressly disclosed herein as will be understood by a person of ordinary skill in the art.

[0109] Probes comprising specific binding moieties, a photo-cleavable moiety, and a detection tag may be made using coupling conditions known in the art. For example, if the specific binding moiety is an antibody, the method may comprise coupling an antibody to the photo-cleavable moiety through a functional group of the antibody (*e.g.*, amine, thiol, carbonyl, etc.). The photo-

cleavable moiety may include (or be modified to include) a reactive functional group, such as an NHS ester, maleimide, etc. These two components can be coupled using coupling reagents typically used in the art, such as EDCI, DCC, EDC, etc. For example, Scheme 3 illustrates an embodiment wherein photo-cleavable moiety **300** is coupled with an NHS group to provide activated NHS ester **302.** Ester 302 is then reacted with an antibody to provide an antibody labeled with the photo-cleavable moiety (conjugate **304**). A person of ordinary skill in the art will recognize that this method can be suitably modified for any of the particular photo-cleavable moieties disclosed herein.

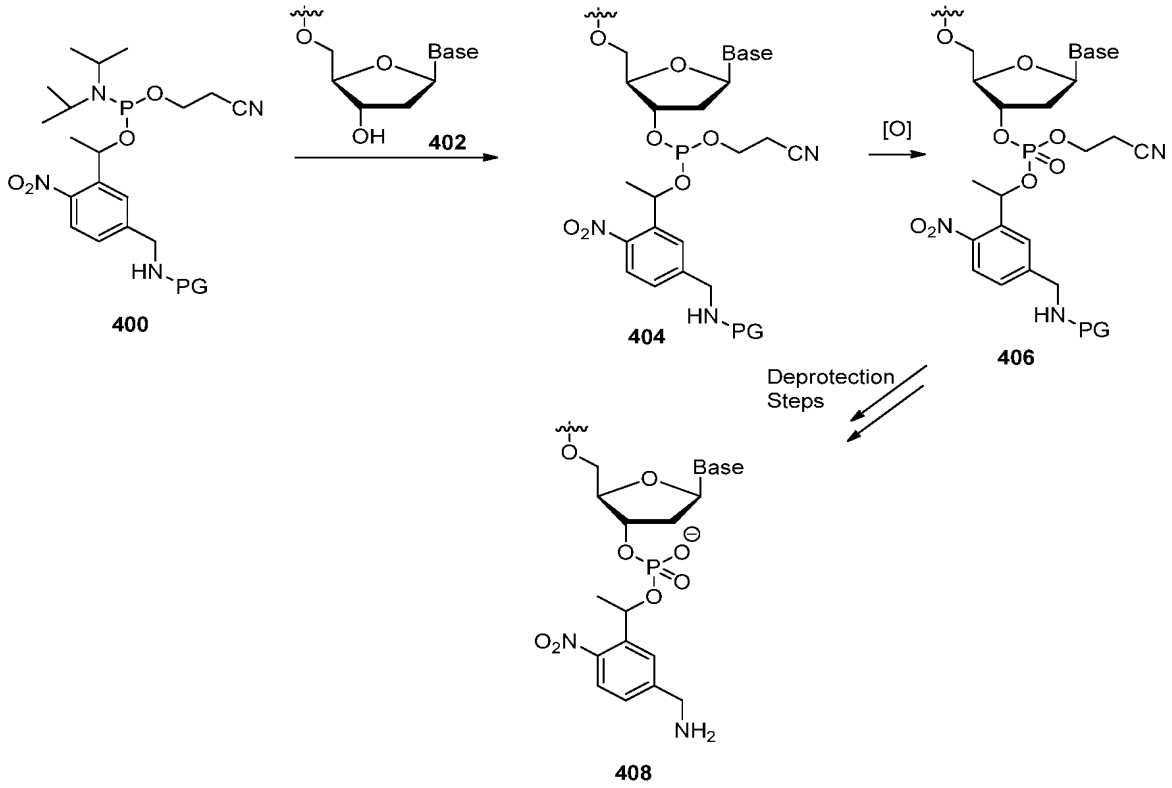


300 **302** **304**

Scheme 3

[0110] In some embodiments using an oligo/polynucleotide as the specific binding moiety, the photo-cleavable moiety may be chemically integrated into the oligo/polynucleotide by using phosphoramidite reagents comprising the photo-cleavable functional group. An oligonucleotide synthesizer may be used to integrate the photo-cleavable phosphoramidite. An exemplary embodiment is illustrated below in Scheme 4. As illustrated in Scheme 4, phosphoramidite reagent **400**, which comprises a photo-cleavable moiety, is coupled to nucleotide **402** of an oligo/polynucleotide to provide labeled oligo/polynucleotide **404,** which is then oxidized to corresponding phosphate **406**. The cyanoester moiety is then removed to provide unprotected phosphate **408**. As previously indicated, the photo-cleavable moiety may comprise other functional groups that may be further manipulated to be coupled with the tag sequence, unique sequence identifier, or other probe components disclosed herein. As illustrated in Scheme 4, and provided solely as an example, photo-cleavable moiety **400** further comprises a protected amine group –NH–PG,

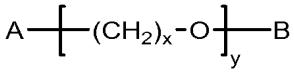
where PG is an amine protecting group. The protecting group may be removed to allow coupling the photo-cleavable moiety to one of more of the probe components disclosed herein.



Scheme 4

[0111] In some embodiments, the specific binding moiety may be bound to a detectable label using coupling conditions known in the art, such as Fc-specific coupling conditions, non-specific coupling conditions, and the like. In some embodiments, the detectable label may be coupled to an antibody through a thiol moiety of the antibody. In some embodiments, a detectable label may be coupled to an antibody through an aldehyde moiety generated on the antibody.

[0112] The detectable label also may be coupled to the specific binding moiety through a linker. Suitable linkers include, but are not limited to aliphatic linkers, alkylene oxide linkers, disulfide-containing linkers, commercially available linkers, polymeric linkers, and the like. One particular embodiment of a class of linkers for use with disclosed probe components is a polyalkyleneglycol linker having the general structure shown below:



[0113] wherein A and B include similar or different reactive groups suitable for reacting with the probe components disclosed herein (*e.g.*, carbonyl-reactive groups, amine-reactive group, thiol-reactive group, and the like, and that are suitable for reacting with a detectable label or a specific binding moiety), x is an integer from 2 to 10 (such as 2, 3 or 4), and y is an integer from 1 to 50, for example, from 2 to 30 such as from 3 to 20 or from 4 to 12. Exemplary linkers, such as PEG linkers, and methods for coupling a specific binding moiety to a detectable label using such linkers, are provided in U.S. Patent No. 7,695,929, which is incorporated herein by reference. These disclosed linkers also may be used to couple various components of the probe other than just coupling the detectable label and the specific binding moiety. For example, a linker may be used to couple the specific binding moiety to a detection tag and/or a detectable label, or one or more linkers may be used to couple the photo-cleavable moiety with one or more of the other probe components disclosed herein. Linkers may be cleaved using any of the primary and/or secondary cleavage techniques disclosed herein.

[0114] Some embodiments of the disclosed method concern using an irradiation event that photochemically induces formation of a compound or radical species that can cleave a chemically-cleavable linker. The compound or radical species formed by the irradiation event should be capable of interacting with one or more functional groups of a linker present in the probe (or a different probe species) in a manner that causes cleavage of the linker. Solely by way of example, a probe moiety may comprise a photo-cleavable moiety that can be cleaved by an irradiation event to provide a species, such as a radical species, that can then subsequently cleave another non-photo-labile linker in the probe. In other embodiments, the light-induced cleavage event can promote cleavage of a photo-cleavable moiety that comprises one or more moieties that can be cleaved and/or released from the photo-cleavable moiety to form a separate chemical species capable of interacting with a linker species

present in the probe (or a different probe).

C. Probe Modifications

[0115] One or more of the probe components disclosed herein may be configured to be more or less robust under certain conditions. For example, a linker and/or a photo-cleavable moiety present in the probe may be photochemically and/or chemically induced to become a species that is more or less robust after such photochemical and/or chemical manipulation. In some embodiments, specifically bound probes can be configured to have a linker and/or photo-cleavable moiety that become(s) more robust before or after being bound to the sample. These probes will therefore comprise detection tags that are not cleaved upon exposure to light of a particular wavelength. Probes that do not comprise the more robust photo-cleavable moiety are cleaved. A secondary cleavage event may then be implemented to cleave the specifically bound probes, such as by using light of a different wavelength, longer time periods of irradiation, more intense light sources, thermal cleavage, chemical cleavage, enzymatic cleavage, or combinations thereof.

[0116] Such techniques also may be applied to components of the probe that are not photochemically labile, but rather are chemically labile. For example, a linker (other than a photo-cleavable moiety) present in the probe may be configured to be more robust after the probe is bound to the sample (*e.g.*, through chemical modification of one or more functional group of the linker). The robust linker is then not susceptible to chemical cleavage under conditions where it would otherwise normally be cleaved. Manipulation and/or analysis of the sample may then be conducted using such conditions that do not chemically alter the robust linkers. The robust linker may then be subsequently cleaved (such as chemically cleaved) using conditions selected particularly for the robust linker (*e.g.*, conditions that are more basic and/or acidic than typical physiologically desired).

[0117] In some embodiments, the linker and/or photo-cleavable moiety may be configured to be so robust as to be impervious to a cleavage event (*e.g.*, photo-chemical and/or chemical cleavage) used for an initial analysis.

Accordingly, a sample comprising probes having these robust components can be stored without degrading. Such samples can then be analyzed at a later time.

[0118] The probe also may be configured to have components that facilitate cell wall permeation. In some embodiments, the probe is specifically designed to permeate a cell wall to facilitate analysis of cell components (*e.g.*, the nucleus, mitochondria, ribosomes, and the like). Such probes can be configured to comprise one or more components that promote lipid-solubility of the probe (*e.g.*, aliphatic linkers, aliphatic functional groups on a photo-cleavable moiety or other probe component).

[0119] In other disclosed embodiments, the hydrophobicity or hydrophilicity of the probe can be configured to attain solubility (or insolubility) in certain process solutions used for sample analysis. For example, the probe can be configured to be more or less hydrophilic, and therefore more or less soluble in developing solutions used with photoresist materials. A person of ordinary skill in the art will recognize that the type of developing solution that is being used will govern whether the probe should be configured to be more or less hydrophobic (or hydrophilic).

[0120] The hydrophobicity and/or hydrophilicity of the probe can also be configured to promote solubility in process solutions, such as the elution solutions used to isolate detection tags after a cleavage event. In these embodiments, if an aqueous buffer is used as the elution solution, the portion of the probe that will end up being cleaved (*e.g.*, detection tag) can be configured to be hydrophilic so as to promote its elution from the sample. If the cleaved portion is not to be isolated, the probe can be configured (prior to or after being added to the sample) to be less hydrophilic so that it does not elute with aqueous elution buffers.

IV. Targets

[0121] Embodiments of the probe and method disclosed herein may be used to identify and/or quantify many different biological targets. Throughout this

disclosure when reference is made to a target protein, it is understood that any polynucleotides associated with that protein can also be used as a target. The target may be a protein or nucleic acid molecule from a pathogen, such as a virus, bacteria, or intracellular parasite, such as from a viral genome. For example, a target protein may be produced from a target polynucleotide associated with (*e.g.*, correlated with, causally implicated in, etc.) a disease. In certain disclosed embodiments, the target (or targets) of interest may be a particular nucleic acid sequence that may comprise a genetic aberration, such as a single nucleotide polymorphism, promoter methylation, mRNA expression, siRNA, a particular copy number change, a mutation, a certain expression level, a rearrangement, or combination thereof. In some embodiments, the targets are soluble proteins obtained from biological samples, such as serum, plasma, and/or urine. Some embodiments of the disclosed method may be used to detect and quantify DNA, RNA, and proteins of the same target (*e.g.*, HER2) simultaneously from the same sample (*e.g.*, from the same tissue section).

[0122] The disclosed method may be used to detect microRNA (miRNA or miR). MicroRNAs are small, non-coding RNAs that negatively regulate gene expression, such as by translation repression. For example, miR-205 regulates epithelial to mesenchymal transition (EMT), a process that facilitates tissue remodeling during embryonic development. However, EMT also is an early step in tumor metastasis. Down-regulation of microRNAs, such as miR-205, may be an important step in tumor progression. For instance, expression of miR-205 is down-regulated or lost in some breast cancers. MiR-205 also can be used to stratify squamous cell and non-small cell lung carcinomas (*J. Clin. Oncol.*, 2009, 27(12):2030-7). Other microRNAs have been found to modulate angiogenic signaling cascades. Down-regulation of miR-126, for instance, may exacerbate cancer progression through angiogenesis and increased inflammation. Thus, microRNA expression levels may be indicative of a disease state.

[0123] A target nucleic acid sequence can vary substantially in size. Without limitation, the nucleic acid sequence can have a variable number of nucleic acid residues. For example a target nucleic acid sequence can have at

least about 10 nucleic acid residues, or at least about 20, 30, 50, 100, 150, 500, 1000 residues. Similarly, a target polypeptide can vary substantially in size. The target polypeptides typically include at least one epitope that binds to a peptide specific antibody, or fragment thereof. In some embodiments that polypeptide can include at least two epitopes that bind to a peptide specific antibody, or fragment thereof.

[0124] In specific, non-limiting examples, a target protein is produced by a target nucleic acid sequence (*e.g.,* genomic target nucleic acid sequence) associated with a neoplasm (for example, a cancer). Numerous chromosome abnormalities (including translocations and other rearrangements, amplification or deletion) have been identified in neoplastic cells, especially in cancer cells, such as B cell and T cell leukemias, lymphomas, breast cancer, colon cancer, neurological cancers and the like. In some examples, therefore, at least a portion of the target molecule is produced by a nucleic acid sequence (*e.g.,* genomic target nucleic acid sequence) amplified or deleted in at least a subset of cells in a sample. In one example, the genomic target nucleic acid sequence is selected to include a gene (*e.g.,* an oncogene) that is reduplicated in one or more malignancies (*e.g.,* a human malignancy). Oncogenes are known to be responsible for several human malignancies. For example, chromosomal rearrangements involving the SYT gene located in the breakpoint region of chromosome 18q11.2 are common among synovial sarcoma soft tissue tumors.

[0125] For example, HER2, also known as c-erbB2 or HER2/neu, is a gene that plays a role in the regulation of cell growth (a representative human HER2 genomic sequence is provided at GENBANK™ Accession No. NC_000017, nucleotides 35097919-35138441). The gene codes for a 185 kd transmembrane cell surface receptor that is a member of the tyrosine kinase family. HER2 is amplified in human breast, ovarian, and other cancers; therefore, a HER2 gene (or a region of chromosome 17 that includes the HER2 gene) can be used as a genomic target nucleic acid sequence. Other breast cancer relevant proteins include the estrogen receptor (ER) and progesterone receptor (PR).

[0126] In other examples, a target protein produced from a nucleic acid

sequence (*e.g.,* genomic target nucleic acid sequence) is selected that is a tumor suppressor gene that is deleted (lost) in malignant cells. For example, the p16 region (including D9S1749, D9S1747, p16(INK4A), p14(ARF), D9S1748, p15(INK4B), and D9S1752) located on chromosome 9p21 is deleted in certain bladder cancers. Chromosomal deletions involving the distal region of the short arm of chromosome 1 (that encompasses, for example, SHGC57243, TP73, EGFL3, ABL2, ANGPTL1, and SHGC-1322), and the pericentromeric region (*e.g.,* 19p13-19q13) of chromosome 19 (that encompasses, for example, MAN2B1, ZNF443, ZNF44, CRX, GLTSCR2, and GLTSCR1) are characteristic molecular features of certain types of solid tumors of the central nervous system.

[0127] The aforementioned examples are provided solely for purpose of illustration and are not intended to be limiting. Numerous other cytogenetic abnormalities that correlate with neoplastic transformation and/or growth are known to those of ordinary skill in the art. Target proteins that are produced by nucleic acid sequences (*e.g.,* genomic target nucleic acid sequences), which have been correlated with neoplastic transformation and which are useful in the disclosed methods, also include (GENBANK™ RefSeq Accession No. in parentheses) the EGFR gene (7p12; *e.g.,* NC_000007, nucleotides 55054219-55242525), the C-MYC gene (8q24.21; *e.g.,* NC_000008, nucleotides 128817498-128822856), D5S271 (5p15.2), lipoprotein lipase (LPL) gene (8p22; *e.g.,* NC_000008, nucleotides 19841058-19869049), RB1 (13q14; *e.g.,* NC_000013, nucleotides 47775912-47954023), p53 (17p13.1; *e.g.,* NC_000017, complement, nucleotides 7512464-7531642), N-MYC (2p24; *e.g.,* NC_000002, complement, nucleotides 151835231-151854620), CHOP (12q13; *e.g.,* NC_000012, complement, nucleotides 56196638-56200567), FUS (16p11.2; *e.g.,* NC_000016, nucleotides 31098954-31110601), FKHR (13p14; *e.g.,* NC_000013, complement, nucleotides 40027817-40138734), as well as, for example: ALK (2p23; *e.g.,* NC_000002, complement, nucleotides 29269144-29997936), Ig heavy chain, CCND1 (11q13; *e.g.,* NC_000011, nucleotides 69165054-69178423), BCL2 (18q21.3; *e.g.,* NC_000018, complement, nucleotides 58941559-59137593), BCL6 (3q27; *e.g.,* NC_000003, complement, nucleotides 188921859-188946169),

MALF1, AP1 (1p32-p31; *e.g.,* NC_000001, complement, nucleotides 59019051-59022373), TOP2A (17q21-q22; *e.g.,* NC_000017, complement, nucleotides 35798321-35827695), TMPRSS (21q22.3; *e.g.,* NC_000021, complement, nucleotides 41758351-41801948), ERG (21q22.3; *e.g.,* NC_000021, complement, nucleotides 38675671-38955488); ETV1 (7p21.3; *e.g.,* NC_000007, complement, nucleotides 13897379-13995289), EWS (22q12.2; *e.g.,* NC_000022, nucleotides 27994271-28026505); FLI1 (11q24.1-q24.3; *e.g.,* NC_000011, nucleotides 128069199-128187521), PAX3 (2q35-q37; *e.g.,* NC_000002, complement, nucleotides 222772851-222871944), PAX7 (1p36.2-p36.12; *e.g.,* NC_000001, nucleotides 18830087-18935219), PTEN (10q23.3; *e.g.,* NC_000010, nucleotides 89613175-89716382), AKT2 (19q13.1-q13.2; *e.g.,* NC_000019, complement, nucleotides 45431556-45483036), MYCL1 (1p34.2; *e.g.,* NC_000001, complement, nucleotides 40133685-40140274), REL (2p13-p12; *e.g.,* NC_000002, nucleotides 60962256-61003682) and CSF1R (5q33-q35; *e.g.,* NC_000005, complement, nucleotides 149413051-149473128).

[0128] In other examples, a target protein is selected from a virus or other microorganism associated with a disease or condition. Detection of the virus- or microorganism-derived target nucleic acid sequence (*e.g.,* genomic target nucleic acid sequence) in a cell or tissue sample is indicative of the presence of the organism. For example, the target peptide, polypeptide or protein can be selected from the genome of an oncogenic or pathogenic virus, a bacterium or an intracellular parasite (such as *Plasmodium falciparum* and other *Plasmodium* species, *Leishmania* (sp.), *Cryptosporidium parvum*, *Entamoeba histolytica*, and *Giardia lamblia*, as well as *Toxoplasma*, *Eimeria*, *Theileria*, and *Babesia* species).

[0129] In some examples, the target protein is produced from a nucleic acid sequence (*e.g.,* genomic target nucleic acid sequence) from a viral genome. Exemplary viruses and corresponding genomic sequences (GENBANK™ RefSeq Accession No. in parentheses) include human adenovirus A (NC_001460), human adenovirus B (NC_004001), human adenovirus C (NC_001405), human adenovirus D (NC_002067), human adenovirus E

(NC_003266), human adenovirus F (NC_001454), human astrovirus (NC_001943), human BK polyomavirus (V01109; GI:60851) human bocavirus (NC_007455), human coronavirus 229E (NC_002645), human coronavirus HKU1 (NC_006577), human coronavirus NL63 (NC_005831), human coronavirus OC43 (NC_005147), human enterovirus A (NC_001612), human enterovirus B (NC_001472), human enterovirus C (NC_001428), human enterovirus D (NC_001430), human erythrovirus V9 (NC_004295), human foamy virus (NC_001736), human herpesvirus 1 (Herpes simplex virus type 1) (NC_001806), human herpesvirus 2 (Herpes simplex virus type 2) (NC_001798), human herpesvirus 3 (Varicella zoster virus) (NC_001348), human herpesvirus 4 type 1 (Epstein-Barr virus type 1) (NC_007605), human herpesvirus 4 type 2 (Epstein-Barr virus type 2) (NC_009334), human herpesvirus 5 strain AD169 (NC_001347), human herpesvirus 5 strain Merlin Strain (NC_006273), human herpesvirus 6A (NC_001664), human herpesvirus 6B (NC_000898), human herpesvirus 7 (NC_001716), human herpesvirus 8 type M (NC_003409), human herpesvirus 8 type P (NC_009333), human immunodeficiency virus 1 (NC_001802), human immunodeficiency virus 2 (NC_001722), human metapneumovirus (NC_004148), human papillomavirus-1 (NC_001356), human papillomavirus-18 (NC_001357), human papillomavirus-2 (NC_001352), human papillomavirus-54 (NC_001676), human papillomavirus-61 (NC_001694), human papillomavirus-cand90 (NC_004104), human papillomavirus RTRX7 (NC_004761), human papillomavirus type 10 (NC_001576), human papillomavirus type 101 (NC_008189), human papillomavirus type 103 (NC_008188), human papillomavirus type 107 (NC_009239), human papillomavirus type 16 (NC_001526), human papillomavirus type 24 (NC_001683), human papillomavirus type 26 (NC_001583), human papillomavirus type 32 (NC_001586), human papillomavirus type 34 (NC_001587), human papillomavirus type 4 (NC_001457), human papillomavirus type 41 (NC_001354), human papillomavirus type 48 (NC_001690), human papillomavirus type 49 (NC_001591), human papillomavirus type 5 (NC_001531), human papillomavirus type 50 (NC_001691), human

papillomavirus type 53 (NC_001593), human papillomavirus type 60 (NC_001693), human papillomavirus type 63 (NC_001458), human papillomavirus type 6b (NC_001355), human papillomavirus type 7 (NC_001595), human papillomavirus type 71 (NC_002644), human papillomavirus type 9 (NC_001596), human papillomavirus type 92 (NC_004500), human papillomavirus type 96 (NC_005134), human parainfluenza virus 1 (NC_003461), human parainfluenza virus 2 (NC_003443), human parainfluenza virus 3 (NC_001796), human parechovirus (NC_001897), human parvovirus 4 (NC_007018), human parvovirus B19 (NC_000883), human respiratory syncytial virus (NC_001781) , human rhinovirus A (NC_001617), human rhinovirus B (NC_001490), human spumaretrovirus (NC_001795), human T-lymphotropic virus 1 (NC_001436), human T-lymphotropic virus 2 (NC_001488).

[0130] In certain examples, the target protein is produced from a nucleic acid sequence (*e.g.,* genomic target nucleic acid sequence) from an oncogenic virus, such as Epstein-Barr Virus (EBV) or a Human Papilloma Virus (HPV, *e.g.,* HPV16, HPV18). In other examples, the target protein produced from a nucleic acid sequence (*e.g.,* genomic target nucleic acid sequence) is from a pathogenic virus, such as a Respiratory Syncytial Virus, a Hepatitis Virus (*e.g.,* Hepatitis C Virus), a Coronavirus (*e.g.,* SARS virus), an Adenovirus, a Polyomavirus, a Cytomegalovirus (CMV), or a Herpes Simplex Virus (HSV).

V. Method of Use

[0131] Disclosed herein are several embodiments of a method of using probe embodiments described herein, as well as novel methods for isolating and/or analyzing a tissue sample, or other cellular specimen, while maintaining tissue context and integrity. In particular disclosed embodiments, the method may further comprise employing subsequent analytical techniques (*e.g.*, PCR, next generation sequencing, array hybridization, etc.) to provide improved multiplexing capabilities, increased sensitivity, relative quantification, and/or improved resolution.

A. General Method

[0132] In particular disclosed embodiments, the method comprises providing a biological sample, such as a tissue sample, comprising one or more targets of interest. The biological sample is contacted with a probe using conditions sufficient to facilitate binding of the specific binding moiety to the target. In some embodiments, the probe may comprise a specific binding moiety, a photo-cleavable moiety, and a detection tag. In other embodiments, the probe may not comprise a photo-cleavable moiety; rather, the sample is exposed to a probe and a composition comprising a separate component comprising a photo-cleavable moiety. The method may further comprise exposing particular selected sections of the biological sample to light having a wavelength and intensity suitable to cleave the photo-cleavable moiety thereby freeing the detection tag. The detection tag may then be detected. The presence and/or location of target may be determined by detecting (*e.g.*, visually and/or with instrumentation) a detectable label. The identity of the target also may be determined by photochemically isolating portions of the probe, such as a tag sequence, tag-template duplex, and/or a unique sequence identifier. The method may have one or more additional steps, one or more steps may be performed manually, or certain or all of the steps may be performed by an automated system.

[0133] FIG. 3 is a schematic that illustrates a method for performing contextual molecular diagnostics 300 showing a step of contacting a sample with a probe 301, removing unbound probe 302, selecting a region 303, irradiating the selected region 304, and detecting the detection tag 305.

[0134] FIG. 4 is a photomicrograph showing that a first region 401 and a second region 402 can be selected. In particular, an image of the sample can be acquired and stored as a digital file. The digital file can be manipulated by a user or through the application of image analysis algorithms to select one or more regions of interest. These selected regions can then been irradiated. In various embodiments, the regions can be irradiated concurrently or sequentially. The detection tags coming from the first and second region can

either be detected concurrently or sequentially.

[0135] FIG. 5(A) – (C) are schematic drawings showing several approaches to irradiating a first selected region 504 and/or a second selected region 503 of a sample 502 disposed on a slide 501. Referring now to FIG. 5(A), a microscope objective 511 can be used to view the sample so that a region of interest can be selected. The microscope objective can also be used to direct light irradiating the sample. For instance, region of interest 504 may be selected in the view of objective 511. Light from an appropriate source can then be directed through objective 511 to region region 504. Another similar approach would be to combine this with raster scanning of the specimen to illuminate the entire selected region. Raster scanning would be achieved by stage translation in the X and Y directions or deflection of the light with rotating mirrors. The intensity of the light beam can be modulated (*e.g.,* UV LED with electronic control of output intensity) as the beam position is scanned across a rectangular region of the specimen to create the illuminated region, or the beam can always be on and the raster scan is restricted to only the boundaries of the selected region. FIG. 5(C) shows a schematic depiction of a digital micromirror device (DMD) being used to selectively irradiate portions of the slide. This embodiment was used to make the elaborate and high resolution design shown in FIGS. 15 and 16.

[0136] FIG. 6(A) – (D) is a schematic depicting an approach to using photochemistry to facilitate an *in situ* polymerase extension reaction wherein polymerase enzymes may be added to tissue for *in situ* polymerase-mediated primer extension. Referring now to FIG. 6(A), oligonucleotide probe 600 is added to a sample having a target sequence 608. Probe 600 includes a photo-reactive moiety (PRM) 606 and optionally a linker 604 and label 602. FIG. 6(B) shows light 612 irradiating PRM 606 so as to enable extension of probe using a polymerase enzyme. In some embodiments, target 608 includes a mutation 610, wherein extension captures mutation information in probe extension 614. In some embodiments, an antibody 616 can be used to purify probe extension. Antibody 616 could be bound to a bead, plate, in gel, etc. and could be selected

to bind label 602. FIG. 6(D) shows an arrow 618 representing that probe extension 614 may be analyzed by an appropriate technology, such as next generation sequencing. The analyzed sequence may contain mutation information, a tag sequence (not shown), and the probe sequence.

[0137] FIG. 6(E) – (F) is analogous to FIG. 6(A) – (B) and can be used replaceably in the workflow of FIG. 6(A) – (D). According to this embodiment, a probe 1600 and a separate photo-activatable moiety 1606 are used for the photo-activated selection. According to this embodiment, photo-activatable moiety 1606 includes 1 or more of the 4-deoxynucleoside triphosphates (i.e. dNTP), rather than an oligonucleotide probe as shown in FIG. 6(A). Oligonucleotide probe 1600 is added to a sample having a target sequence 1608. Probe 1600 optionally includes a linker 1604 and label 1602. FIG. 6(F) shows light 1612 irradiating photo-activatable moiety 1606 so as to enable extension of probe using a polymerase enzyme. In some embodiments, target 1608 includes a mutation 1610, wherein extension captures mutation information in the probe extension (as shown in FIG. 6(C)). FIG. 6(D) shows an arrow 618 representing that any probe extension may be analyzed by an appropriate technology, such as next generation sequencing. The analyzed sequence may contain mutation information, a tag sequence (not shown), and the probe sequence.

[0138] The extent of polymerase chain extension that can be obtained using embodiments of the disclosed method ranges from hundreds of extended nucleotides to thousands of extended nucleotides. In some embodiments, the number of nucleotides added to the initial probe can range from about 10 nucleotides to about 2000 nucleotides (such as about 50 nucleotides to about 1500 nucleotides, or about 100 nucleotides to about 1000 nucleotides, or about 200 nucleotides to about 400 nucleotides). The extent of the chain extension can be determined using one or more reverse primers that have some degree of complementarity (*e.g.*, about 70% to about 99% complementarity) to certain regions of the extended probe. Polymerase-mediated chain extension may be carried out using a range of probe concentrations (*e.g.*, 100 ng/mL to about 100

μg/mL, such as about 10 ng/mL to about 10 μg/mL), and can be carried out for any suitable amount of time, such as from about 20 minutes to about 90 minutes.

[0139] FIG. 7 similarly shows a detection motif within the scope of the present application. In particular, a probe 700 can be brought into contact with a sample having a target 708. Probe 700 comprises specific binding oligonucleotide 701, a photo-cleavable moiety 702, and a detection tag 713. Probe 700 may optionally include a label 714, for example to aid in purification. Contacting probe 700 and sample, under the appropriate conditions results in hybridization of probe 700 to target 708 as indicated by an arrow 709. Subsequent to removal of unbound probe, irradiating a selected portion of the sample with light, as indicated by arrow 710, cleaves photo-cleavable moiety 702, freeing detection tag 713. Detection tag 713 can then be removed from the sample (*e.g.*, through pipetting, contact transfer, absorption, rinsing, microfluidics, rinsing, or other appropriate method), optionally purified (*e.g.*, through use of optional label 714), and analyzed by an appropriate technology (*e.g.,* an array, NGS, PCR, etc.).

[0140] Referring now to FIG. 8, shown is a schematic of a particular embodiment. Oligo/polynucleotide 800 is combined with sample 802. Probe 800 associates with a complimentary sequence sample 802. A composition comprising caged ATP 804 caged with photo-cleavable moiety 806 is added to the sample 802. Irradiation with light liberates caged ATP 804. Exposure to a suitable kinase then phosphorylates probe 800 at the 5' end to provide phosphorylated probe 808. Subsequent modification of probe 808 may be carried out in order to provide, for example, labeled probe 810.

[0141] In yet other embodiments, adjacent probes may be linked via a ligase and then a particular tag sequence or unique sequence identifier of one probe may become detectable through ligation to a probe comprising a detectable label. The ligated probes may be amplified using PCR (or ligase chain reaction), if greater sensitivity is desired. A particular embodiment is exemplified

schematically by FIG. 9. With reference to FIG. 9, oligo/polynucleotide probe 900 is added to sample 902. Probe 900 associates with a complimentary sequence of sample 902. Sample 902 is then exposed to caged-ATP 904. Activation with light releases the caged ATP. Adding a suitable kinase phosphorylates oligo/polynucleotide probe 900 to provide phosphorylated probe 906. A ligase then ligates oligo/polynucleotide probe 906 to an adjacent oligomer deposited on sample 902 comprising a detectable label 908, thereby forming ligated probe 910. Only those oligomer sequences that are adjacent to the phosphorylated end of oligo/polynucleotide probe 906 will be ligated to the probe. By forming ligated probe 910, a target sequence may be identified once the ligated probe is detected (through detectable label 908) and subsequently quantified using array hybridization or sequencing. In some embodiments, ligase chain reaction or PCR can be used to amplify ligated probe 910 if greater sensitivity is needed.

1. Selecting and Preparing the Sample

[0142] Samples used in various embodiments of the method, such as a tissue, a cell, or other biological sample (*e.g.*, serum, plasma, and/or urine), can be obtained from a subject for routine screening or from a subject that is suspected of having a disorder, such as a genetic abnormality, infection, or a neoplasia. Samples may be fresh, frozen, formalin-fixed, paraffin-embedded, or combinations thereof. In some embodiments, the method also can be applied to samples that do not have genetic abnormalities, diseases, disorders, etc., referred to as "normal" samples. Such normal samples are useful, among other things, as controls for comparison to other samples. The samples can be analyzed for many different purposes. For example, the samples can be used in a scientific study or for the diagnosis of a suspected malady, or as prognostic indicators for treatment success, survival, etc. Samples can include a single target, or can include multiple targets that can be specifically bound by one or more probe embodiments disclosed herein. While the methods may be used on a variety of samples, it is particularly advantageous for tissue samples which have inherent contextual information that is not present in solution based

samples. In particular, the relationship between cells, and indeed the size, shape, and distribution of cells provides an enormous amount of information that is lost if that tissue is lysed, digested, ground. Evidencing the importance of this contextual information is the significance of primary staining (i.e. staining with hematoxylin and eosin, H&E) to the diagnosis of cancer. Being non-specific stains, H&E staining provides only contextual information yet it is the gold-standard in diagnosing cancer.

[0143] In particular disclosed embodiments, the sample tissue sample is a formalin-fixed, paraffin-embedded tissue sample. Such tissue samples may be prepared by extracting tissue from a subject, exposing the tissue to a buffered formalin (or paraformaldehyde) solution, and then embedding the tissue with paraffin. The tissue sample(s) may then be placed on a substrate, such as a microscope slide or a microarray slide.

[0144] Other embodiments concern cell samples, or soluble protein samples. In some embodiments, the sample may be a smear sample, a cytology sample, a chromosomal smear, a biopsy sample, a fine needle aspirate sample, or a circulating tumor cell sample.

[0145] In some embodiments, the sample may be provided on a slide having a registration element suitable for identifying the sample and/or the location/position of the slide comprising the sample. For example, the registration element may be a physical registration element, such as a barcode, alignment holes, alignment protrusions, alignment keys, and the like. The registration elements can facilitate transfer of the slide throughout the system embodiments disclosed herein. For example, the registration elements of the slide can facilitate transfer of a slide comprising a sample to the imaging portion of the system whereby certain areas of the sample are selected and marked for further analysis. The slide can then be processed in other parts of the system, or it can be removed from the imaging component and saved for analysis at a later time. In these embodiments, the registration elements of the slide can be used to reset the slide in substantially the same position as it was taken from the imaging portion of the system, thereby bypassing the need for

an additional sample selection step. The slide can be added to the system and undergo subsequent analysis steps without being reimaged and marked.

[0146] In some embodiments, the registration elements can be used to translate image based actions (*e.g.*, selection and marking) into actions within the system, such as an irradiation event. The registration element can indicate to the system that the slide is correctly aligned for irradiating the sample in the selected (or non-selected) areas as disclosed herein.

[0147] Each slide can be labeled with a barcode number specific to that particular slide and sample. This barcode number can be used to correlate a slide specimen with particularly selected regions of the sample that can be stored electronically. This electronically stored information can be used for subsequent analysis of the particular sample. Accordingly, a number of slide specimens can undergo one or more initial selection steps and then be stored for later analysis.

2. Contacting the sample with the probe

[0148] After a sample is prepared, the biological sample can then be exposed to one or more probes. For example, a sample may be exposed to a probe by contacting the sample with a solution comprising the probe, or a mixture of probes. The solution containing the probe can be made with any biologically-acceptable composition, such as a buffer composition. In some embodiments, the buffer comprises 50 mM Tris/HCl, 10 mM KCl, 5 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$. In other embodiments, the buffer may be selected from those typically used in the art, such as TES, HEPES, Tris, and the like.

[0149] In particular disclosed embodiments, the probe is deposited onto a sample region and may be hybridized to, or chemically coupled to (*e.g.*, electrostatically coupled, covalently coupled, or ionically coupled), a particular target sequence or moiety present in the target, or that has also been associated with the target, such as a primary antibody. After the probe is associated with the target, the sample may be washed and any unassociated probe components can be washed away.

[0150] In some embodiments, the deposited probe may be detected and

particular portions of the sample comprising the deposited probe may be marked for analysis. In some embodiments, visual detection of a signal produced by a detectable label (*e.g.*, a chromophore, a fluorophore, a hapten, etc.) during a primary and/or secondary recognition event can be used to detect the particular portion of the sample to be marked. In some embodiments, one or more portions of the sample may be selected for further analysis by using a stain (*e.g.*, a primary stain, counterstain, or combination thereof). In other embodiments, fluorescence of a fluorophore deposited on one or more portions of the sample may be used to indicate where the sample should be marked for further analysis. In yet other embodiments, morphological characteristics may be used to distinguish portions of a sample for analysis, particularly in cell or tissue samples. For example, fluorescence microscopy and/or phase contrast microscopy techniques known in the art can be used to visualize the shape and/or arrangement of cell components thereby facilitating selection of these particular regions for analysis. Digital pathology techniques also can be used to image the sample and facilitate sample selection.

[0151] Once the portion of the sample that will be analyzed has been identified, it may be marked. For example, the portion of the sample that will be analyzed may be manually marked, such as marking a sample portion on a slide to which the sample is attached, or it may be marked on a separate digital viewing screen with the selected locations being stored electronically. In some embodiments, a marking/selection step may involve a substantially instantaneous irradiation event. In other embodiments, a marking/selection step may occur first, followed by a subsequent irradiation step. The act of marking/selecting a portion of the sample may be configured to initiate an irradiation event.

3. Exposing the sample to light

[0152] In illustrative embodiments, a system for performing contextual diagnostics using a photo-reactive probe further comprises a registration system configured to determine the position of the sample, thereby enabling

automatic relation of the location inputs to the location on the sample upon which the light source is configured to irradiate. In one embodiment, the system includes a sample platform that is configured to receive the sample adhered to a substrate and the registration system is configured to relate the position of the sample by determining the position of the substrate in relation to the light source. In another embodiment, the substrate is a glass microscope slide inscribed with a registration mark. In another embodiment, the registration system is configured to detect a registration mark. In yet another embodiment, the registration system is configured to detect two or more registration marks, the two or more registration marks inscribed on the substrate. In yet another embodiment, the registration system is configured to identify a location of the registration mark to within about 500 μm, to within about 300 μm, to within about 100 μm, to within about 50 μm, or to within about 10 μm. In yet other illustrative embodiments, the system uses the tissue as a registration mark.

[0153] In illustrative embodiments, the registration system is configured so that it enables the computer, which is configured to receive a data file associated with an image of the sample, to relate the image data to the position of the sample in relation to the light source. This in turn, enables user inputs that relate to the image data file to be actionable with respect to the sample. In one embodiment, the registration system is configured to detect a registration mark and the computer is configured to associate the registration mark with an imaged registration mark within the image of the sample. In another embodiment, the registration system is configured to detect two or more registration marks and the computer is configured to associate the two or more registration marks with two or more imaged registration marks within the image of the sample such that the computer is configured to relate the image of the sample with a position of the sample within an irradiation chamber. In yet another embodiment, the computer is configured to accept user inputs, the user inputs include selecting portions of the sample, *e.g.*, based on the image of the sample, for instructing the laser to ablate the sample. In another

embodiment, the computer is configured to accept artificial intelligence inputs, the artificial intelligence inputs include data from automated pattern recognition of the image of the sample. In another embodiment, the user is a pathologist and the user inputs are selections of regions of the sample selected by the pathologist from the image of the sample. In another embodiment, the computer is configured to accept a data file associated with an image of the sample and the location inputs are associated with user-selectable regions on the sample for irradiating.

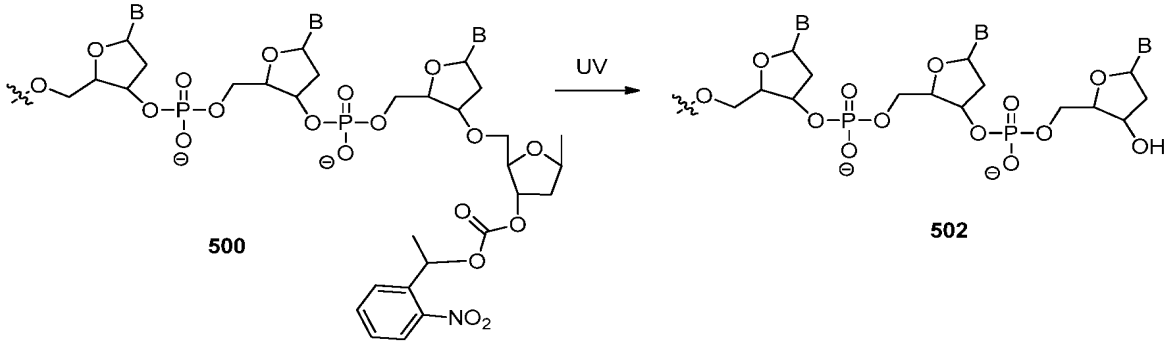
[0154] When the sample is exposed to light, the entire sample may be irradiated, or a portion (or portions) of the sample may be selectively irradiated. In some embodiments, the sample is selectively irradiated to facilitate isolation of only certain portions of the probe deposited on the sample (*e.g.*, the unique sequence identifier or tag sequence of a probe) in a targeted region of the sample. In some embodiments, the sample may be selectively irradiated to facilitate isolation of probes deposited on the sample in non-selected regions. That is, selected portions of the sample identified for analysis may be marked and then unmarked portions of the sample to which one or more non-selected probes are bound may then be irradiated, thereby cleaving the non-selected probes. The selected probes may then be isolated using a second irradiation step and/or secondary cleavage technique as disclosed herein and a suitable elution step. In other embodiments, the sample may be selectively irradiated to facilitate isolation of probes deposited on the sample in the selected regions. Selective irradiation can be performed iteratively to provide multiple different irradiation events. In such embodiments, targets and/or detection tags optionally can be isolated between each iterative irradiation step to provide a quantitative, multiplexed map of the sample and the various different targets therein.

[0155] Selective irradiation can be accomplished using a photomask, a focused beam of light from a light source, or by iteratively tuning (for a particular photo-cleavable moiety) light of selected wavelengths. For example, selective irradiation can be used to activate a photo-cleavable moiety that is

present in the probe and/or other component to which the sample is exposed. In some embodiments, activating the photo-cleavable moiety facilitates cleaving portions of the probe. For example, if the photo-cleavable moiety is used to link two components of a probe, selective irradiation separates the two components.

[0156] In some embodiments, irradiation may be used to generate a reactive species that is capable of reacting with other moieties present, such as by cleaving one or more linkers. Suitable reactive species do not interfere with sample analysis or destroy the sample. Suitable reactive species include, but are not limited to, photoinitiators such as azobisisobutyronitrile (AIBN), benzoyl peroxide (BPO). Embodiments of the reactive species can cleave linkers selected from, for example, photo-cleavable moieties used as linkers, aliphatic linkers, alkylene oxide linkers, disulfide linkers, commercially available linkers, polymeric linkers, and the like.

[0157] In some embodiments, the photo-cleavable moiety functions as a blocking agent. In these embodiments, selective irradiation cleaves the photo-cleavable moiety from the particular probe (*e.g.*, a primer) and frees at least one end of the probe for further manipulation. An exemplary embodiment is illustrated below in Scheme 5. Conjugate **500** includes a nitrophenyl moiety that readily absorbs light having wavelengths of about 300 nm to about 400 nm, such as about 325 nm to about 375 nm, or about 345 nm to about 370 nm. The nitrophenyl moiety is attached to at least one nucleotide of an oligo/polynucleotide via an ester group. UV irradiation cleaves the benzyl carbon-oxygen bond, followed by release of CO_2 to form conjugate **502**, which has a free hydroxyl group suitable for chain extension.



500 **502**

Scheme 5

[0158] Selective irradiation also can be used to cleave photo-cleavable moieties from a photo-activatable primer extension reagent, such as caged dATP molecules, which are disclosed herein. Other photo-activatable primer extension reagents can include activatable degenerate primers for whole genome amplification or specific genetic lesions, dGTP, dCTP, dTTP, and the like.

[0159] Disclosed herein are various different embodiments for irradiating the sample to provide selective target irradiation. During the irradiation step, the sample is exposed to light capable of providing the energy sufficient to induce a desired result, such as to activate the photo-cleavable moiety. In particular disclosed embodiments, the light is ultraviolet light having a wavelength ranging from about 200 nm to about 400 nm. In some embodiments, light is visible light having wavelengths ranging from about 400 nm to about 790 nm. Suitable light sources are capable of providing low intensity light to preserve sample integrity. In some embodiments, the light source provides from about 1 mW/cm^2 to about $1W/cm^2$, such as from about 1 mW/cm^2 to about 500 mW/cm^2, or from about 1 mW/cm^2 to about 200 mW/cm^2, or from about 5 mW/cm^2 to about 150 mW/cm^2, or from about 10 mW/cm^2 to about 150 mW/cm^2. Exemplary light sources include, but are not limited to, lasers, LEDs, or lamps (such as incandescent lamps, halogen lamps, mercury lamps, fluorescent lamps, and the like).

[0160] In illustrative embodiments, the irradiated tissue area would be selected dependent on the surface area of diseased tissue, more specifically the number of diseased cells. In many cases, such as biopsies, this area would range from a single cell (10-20 micrometers in diameter) to a few thousand cells (1-50 millimeters in diameter). In resected tissues, the diseased area may include hundreds of thousands to millions of cells covering the majority of a glass slide (75 x 25 mm). In yet other embodiments in which a DMD is used, the irradiated area could be about 10 μm x10 μm or larger regions made up by

10 μm x10 μm regions

[0161] Selective irradiation may comprise direct irradiation or indirect irradiation. Direct irradiation typically comprises exposing the sample to light, particularly a light source capable of producing light of any suitable wavelength (or wavelengths), without filtering the light produced by the light source. For indirect irradiation, an intermediate filter or masking material may be placed between the light source and the sample, thereby filtering or blocking some of the light produced by the light source from the sample.

[0162] Selective irradiation also may comprise exposing the sample to light positioned to selectively expose a particular portion of the sample. For example, an LCD screen may be used to direct light of a selected wavelength (or wavelengths) to a particular location in the sample. A tumor-containing portion of a tissue may be irradiated and a non-tumor-containing portion of the sample may be irradiated (typically sequentially). These particular embodiments provide the ability for collecting portions of a probe (*e.g.*, a tag-template duplex or unique sequence identifier) that are liberated by the selective irradiation (via cleavage of the photo-cleavable moiety) and sequencing a tag-template duplex or unique sequence identifier that correlates to the tumor portion of the sample and a different tag-template duplex or unique sequence identifier that correlates to the normal portion of the sample. In some embodiments, selective irradiation can be used to cleave a pull-down tag in an unselected tissue area (*e.g.*, an area that does not comprise a tumor). For example, a detectable label (*e.g.*, biotin) attached to a sample through one or more probe components can be cleaved from regions of the sample other than the tumor-containing area.

[0163] FIG. 5(A) – (C) are schematic drawings showing several approaches to irradiating a first selected region 504 and/or a second selected region 503 of a sample 502 disposed on a slide 501. Referring now to FIG. 5(A), a microscope objective 511 can be used to view the sample so that a region of interest can be selected. The microscope objective can also be used to direct light irradiating the sample. For instance, region of interest 504 may be selected in the view of objective 511. Light from an appropriate source can then be directed through

objective 511 to region 504. The effect of this approach is irradiation of the field of view, as such, the field of view for the microscope is the selected region. Referring now to FIG. 5(B), shown is an alternative approach that uses a light source 520 to irradiate a screen 521 having selective optical density. Screen 521can block incident irradiation to the bulk of sample 502 while allowing irradiation to impact regions 504 and 503 as depicted by beams 522 and 523 passing through the screen to the respective selected regions.

[0164] The selective optical density screen may be a liquid crystal display (LCD) screen. The sample regions of interest may be marked by configuring an LCD screen to have the desired area outlined (for example, the positions may be marked manually, and the particular area/pattern transferred to the LCD screen using a projector). The areas that are not to be irradiated are "blacked-out." Portions of the LCD screen that are not "blacked-out" enable passage of the light produced by the light source to the selected region. Referring now to FIG. 5(C), in another embodiment, a digital micromirror device 531 is used selectively irradiate regions 504 and 503. Beams 532 and 533 are generated by irradiating digital micromirror device 531 with light source 520. In particular disclosed embodiments, the digital micromirror device comprises a 1400 x 1050 array of micro-mirrors (12.88 μm^2).

[0165] In yet other embodiments, selective irradiation may be obtained by using a digital micromirror device, which can be used to focus light from the light source onto the sample. The digital micromirror device may comprise various different portions/mirror segments of a particular size (*e.g.*, squares having a size ranging from about 4.5x4.5 μm to about 15x15 μm) that can be adapted to irradiate particular locations and/or features within the sample.

[0166] The sample typically is exposed to the light for a time sufficient to cleave the desired photo-cleavable moieties, but not so long as to effect sample degradation. For example, the time period during which the sample is irradiated typically is no longer than about 5 minutes, about 4 minutes, about 3 minutes, or about 2 minutes. In some embodiments, the sample may be

irradiated for an effective period of time ranging from about 1 second to about 60 seconds, typically from about 10 seconds to about 60 seconds, more typically from about 20 seconds to about 60 seconds, with certain working embodiments irradiating the sample for about 20 seconds or less.

[0167] Referring now to FIG. 10, shown is a schematic that illustrates an embodiment of the present disclosure wherein a multiplexed image of the sample is constructed using the approaches described herein. Across a substrate 1000, for example a microscope slide labeled with a barcode 1001, a grid 1002 or other appropriate sectioning strategy is created to overlap the sample (not shown, but understood to be within the grid). Grid 1002 is arranged to correspond to the resolution of the chemical image that is sought. Illustrative of the method, the grid may be a series of raster locations on which the sample is iteratively irradiated. The arrangement and structure of the grid is not material, rather it can be variably designated according to the intended need. In some samples, specific cells are irradiated and the "grid" does not follow a geometric pattern, but instead matches a biologic pattern associated with the sample. Referring back to FIG. 10, arrow 1003 represents removal of detection tags from substrate 1000 to a multi-well plate 1010. Multi-well plates are well known in the art and come in various configurations having various numbers of wells from 2 to hundreds. While shown as a typical macroscopic (*e.g.*, 64-well plate), microscopic vessels are also within the scope of the present disclosure. For example, multi-well plate may be a microfluidic device having small volume chambers. Multi-well plate 1010 is shown to include a plurality of reaction wells 1011.

[0168] According to illustrative embodiments, a light is used to irradiate a selected region (*e.g.*, a single grid location), thereby releasing detection tags from the selected region. The detection tags are removed from substrate 1003, as represented by arrow 1003, and deposited into one of the plurality of reaction wells 1011. The process is repeated iteratively so that detection tags from each selected region are deposited into a different reaction well. While the detection tags present in each cell could be independently analyzed, a more

elegant solution is shown in FIG. 10. According to this solution, different forward primers (represented as arrows 1005) can be placed in each row of multi-well plate 1010. Similarly, different reverse primers (represented as arrows 1004) can be placed in each column of multi-well plate 1010. Within each well, detection tag 1008 (represented in a block diagram), includes a portion 1014 corresponding to a portion of forward primer 1005 and a portion 1013 corresponding to a portion of reverse primer 1004. Arrow 1007 represents a PCR step which amplifies the detection tags in each well and incorporates forward and reverse primers to create a PCR product 1009. PCR product 1009 may include a first adapter sequence 1017 (*e.g.*, for NGS), a unique row identifier 1016, the portion 1014 corresponding to the portion of the forward primer 1005, a unique identification tag sequence (*e.g.*, which would include target or reagent information), the portion 1013 corresponding to the portion of the reverse primer 1004, a unique column identifier 1018, and a second adapter sequence 1018 (*e.g.*,also for NGS). The elegance of this approach is that all the wells could then be pooled and analyzed concurrently. The location from where the detection tag came on the sample is encoded in the sequence. As such, all detection tags from each location on the sample could be binned together to read out a highly multi-plexed chemical image of each selected region of the sample.

4. Detecting the tag and/or the target

[0169] The target may be detected using various detection schemes now known or hereafter developed. In particular disclosed embodiments, detecting may first comprise obtaining an isolated portion of the probe and analyzing/sequencing the isolated portion of the probe. For example, after the photo-cleavable moiety is cleaved, a portion of the probe may be released and the released portion isolated from the sample, such as by washing the sample with an appropriate buffer solution. If the isolated portion includes information encoded by the nucleotide sequence, the isolated portion is then sequenced, such as by using next generation sequencing. If necessary, or desirable, an isolated portion of the probe may be amplified, such as by performing a

polymerase chain reaction. In particular disclosed embodiments, amplifying and sequencing steps may be combined. In some embodiments, portions of the sample that are irradiated may undergo amplification while portions of the sample that are not irradiated are not amplified.

[0170] Next generation sequencing provides the ability to sequence a large number of samples, wherein each sample may comprise multiple targets and/or specific binding moieties bound to unique sequence identifiers. As each unique sequence identifier is different, each of the multiple different targets and/or specific binding moieties can be bound to its own unique sequence identifier. Once isolated, unique sequence identifiers can be differentiated during sequencing and the particular target and/or specific binding moiety identified based on its correlated unique sequence identifier. Thus, the identity of the target and/or specific binding moiety may be detected with increased accuracy as compared to typical tissue analytical techniques. Additionally, the quantity of specific binding moieties bound in a particular portion of a sample may be determined based on the isolated unique sequence identifiers.

[0171] In some embodiments, the liberated unique sequence identifier may be collected out of a suitable microfluidic platform after sample irradiation. The microfluidic platform may be a microarray platform, such as a microarray slide. In some embodiments, a microarray slide can be used comprising a surface designed to spatially arrange particular targets in locations that facilitate probe addition and selective retrieval of moieties that are cleaved from the probe (*e.g.*, unique sequence identifiers, tag sequences, primers, and the like). The microarray slide can comprise multiple, spatially arranged locations that can comprise one particular target, multiple targets, or subsets of targets. Due to the spatial arrangement on the microarray slide, a selected region containing a portion of the moieties that are cleaved from the probe after selective irradiation can be eluted without eluting substantial amounts of non-selected regions. The selected region may be exposed to an elution buffer and/or a denaturing buffer. These buffers facilitate removal of the desired moieties within the selected region (*e.g.*, unique sequence identifier, tag sequence, primer, etc.). The

microarray slide typically may be used in combination with a microarray scanner that is configured to scan the microarray slide's surface and identify the region of interest. A dispensing instrument also may be used, which can be configured to receive input from the microarray scanner and dispense the desired elution and/or denaturing buffers to a particular region of the slide. An exemplary microarray system comprising these components (and other suitable components) is disclosed in U.S. Patent Publication No. 2010/0076185, which is incorporated herein by reference in its entirety.

B. Photochemical Immunohistochemical Analysis

[0172] In particular disclosed embodiments, the method may comprise a histochemical assay suitable for detecting one or more targets in a sample. The probes used in these embodiments typically comprise an antibody as the specific binding moiety, a photo-cleavable moiety, and a tag sequence or a unique sequence identifier. Particular embodiments of the method concern exposing a biological sample, such as a tissue sample (*e.g.*, an FFPE tissue sample), to one or more probes comprising an antibody that can either bind directly to a target in the sample, or indirectly to a target. For example, particular embodiments concern first contacting the sample with a primary antibody, followed by contacting the sample with a probe comprising a secondary antibody, a photo-cleavable moiety and a unique sequence identifier (or tag sequence), where the secondary antibody specifically recognizes the primary antibody. If a unique sequence identifier is used, then it can be released via selective irradiation and the sequence determined to identify the target detected. This embodiment also can be used to quantify the number of bound antibodies in the tissue as the number of sequencing reads can equate directly to the number of bound antibody-unique sequence identifier conjugates.

C. Photochemical In Situ Hybridization

[0173] Also disclosed herein are embodiments of a method wherein one or more targets are detected using a photochemical-based *in situ* hybridization assay. In particular disclosed embodiments, the method concerns exposing a

sample to a probe comprising an oligo/polynucleotide, a photo-cleavable moiety, and a tag sequence or unique sequence identifier. The probe is allowed to hybridize to the tissue sample and then portions of the sample are irradiated using a light source.

[0174] In other disclosed embodiments, the method may concern exposing a sample to a probe comprising an oligo/polynucleotide, a photo-cleavable moiety, and a first member of a binding pair, typically selected from biotin, streptavidin, or a hapten. After the probe is hybridized to the particular target of interest, the sample is washed and then certain portions are irradiated with light to dissociate a first member of the specific binding pair from any hybridized probes within the particular zone of irradiation. After performing a dehybridization step on the sample, the de-hybridized probes are passed through a column comprising an immobilized second member of a specific binding pair. Probes that were irradiated elute through the column, as specific binding cannot occur between the probe and the second member of the specific binding pair present on the column. Those probes that were not selectively irradiated will not elute through the column, as the still-attached first member of the specific binding pair binds with the second member and is retained on the column. In some embodiments, the method facilitates isolation of the probes located in the particular regions of interest.

D. In Situ Polymerase Extension

[0175] In particular disclosed embodiments, photochemistry may be used to facilitate an *in situ* polymerase extension reaction wherein polymerase enzymes may be added to tissue for *in situ* polymerase-mediated extension, *in situ* PCR, and/or *in situ* nick-translation. For example, in particular embodiments, a nucleic acid target with a point mutation may be analyzed by exposing the nucleic acid target to a probe comprising an oligo/polynucleotide and a photo-cleavable moiety that blocks one end, typically the 3' end, of the oligo/polynucleotide. The oligo/polynucleotide may further comprise a detectable label.

[0176] In another embodiment one or more of the deoxynucleoside

triphposphates are blocked with a photo-cleavable moiety and the oligo/polynucleotide may or may not be blocked with a photo-cleavable moiety.

[0177] The extent of polymerase chain extension that can be obtained using embodiments of the disclosed method ranges from hundreds of extended nucleotides to thousands of extended nucleotides. In some embodiments, the number of nucleotides added to the initial probe can range from about 10 nucleotides to about 2000 nucleotides (such as about 50 nucleotides to about 1500 nucleotides, or about 100 nucleotides to about 1000 nucleotides, or about 200 nucleotides to about 400 nucleotides). The extent of the chain extension can be determined using one or more reverse primers that have some degree of complementarity (*e.g.*, about 70% to about 99% complementarity) to certain regions of the extended probe. Polymerase-mediated chain extension may be carried out using a range of probe concentrations (*e.g.*, 100 ng/mL to about 100 μg/mL, such as about 10 ng/mL to about 10 μg/mL), and can be carried out for any suitable amount of time, such as from about 20 minutes to about 90 minutes.

E. Multiplexing

[0178] The components, methods, kits, and systems described herein primarily provide description for one or two probes used on a particular sample. While useful in this context, one significant advantage of the present approach is that the detection modality is based on oligonucleotides. As such, the detection and quantification of binding events is performed using robust and well-known oligonucleotide approaches which are capable (and routinely applied to) high levels of multiplexing. In illustrative embodiments, between about 2 and about 100,000 targets are detected. In other embodiments, between about 10 and about 100,000 targets are detected. In another embodiment, between about 10 and about 1000 targets are detected. In yet another embodiment, greater than about 5 targets are detected. In another embodiment, greater than about 10 targets are detected.

VI. System Embodiments

[0179] Also disclosed is a system for implementing any of the method embodiments disclosed herein. In some embodiments, the system comprises a slide tray configured to accept a slide comprising a sample, a dispenser configured to dispense disclosed probe embodiments, such as a probe comprising a specific binding moiety, a photo-cleavable moiety, and a detection tag. The system also includes a light source configured to selectively irradiate the sample, thereby cleaving the photo-cleavable moiety and releasing the detection tag. The slide used in the disclosed system may be a single specimen slide, an array slide, or a microarray slide. The light source may be selected from any of the light sources disclosed herein, such as a laser, an LED, or a lamp.

[0180] In some embodiments, the disclosed system further comprises an imaging device configured to capture a modifiable sample image, a digital screen configured to accept the modifiable sample image from the imaging device, and a selection device configured to mark a selected portion of the modifiable sample image for analysis. The imaging device may be a digital camera or a video camera.

[0181] The disclosed system also may further comprise a dispensing instrument configured to dispense an elution buffer onto the slide to elute the detection tag, and a recovery instrument configured to recover the detection tag. The recovery instrument may be configured to deliver the detection tag directly to an additional portion of the system. For example, the system may further comprise a PCR instrument configured to accept and amplify the detection tag. Other embodiments further comprise a sequencing instrument configured to accept and sequence the detection tag, and a computer configured to translate information generated from the sequencing instrument to the digital screen, thereby correlating the detection tag to the selected region from which it was cleaved. Yet other embodiments further comprise an array platform comprising a modified surface configured to accept and bind the

detection tag.

[0182] Exemplary automated systems available through Ventana Medical Systems, Inc., Tucson, AZ include SYMPHONY® Staining System, catalog #: 900-SYM3, VENTANA® BenchMark Automated Slide Preparation Systems, catalog #s: N750-BMKXT-FS, N750-BMKU-FS, VENTANA, and VENTANA® BenchMark Special Stains automated slide stainer. These systems employ a microprocessor controlled system including a revolving carousel supporting radially positioned slides. A stepper motor rotates the carousel placing each slide under one of a series of reagent dispensers positioned above the slides. Bar codes on the slides and reagent dispensers permits the computer controlled positioning of the dispensers and slides so that different reagent treatments can be performed for each of the various tissue samples by appropriate programming of the computer.

[0183] Illustrative instrumentation systems are designed to sequentially apply reagents to tissue sections mounted on one by three inch glass microscope slides under controlled environmental conditions. The instrument must perform several basic functions such as reagent application, washing (to remove a previously applied reagent), jet draining (a technique to reduce the residual buffer volume on a slide subsequent to washing), application of a light oil used to contain reagents and prevent evaporation, and other instrument functions. Exemplary staining instruments process slides on a rotating carousel. The slides maintain a stationary position and a dispenser carousel rotates the reagents above the fixed slides. The processes described herein can be performed using various physical configurations. The process of clarifying and staining tissue on a slide consists of the sequential repetition of basic instrument functions described above. Essentially a reagent is applied to the tissue then incubated for a specified time at a specific temperature. When the incubation time is completed the reagent is washed off the slide and the next reagent is applied, incubated, and washed off, *etc.*, until all of the reagents have been applied and the staining process is complete.

[0184] For related disclosure, reference is made to Richards et al. U.S. Pat.

No. 6,296,809, assigned to Ventana Medical Systems, which describes an apparatus and methods for automatically staining or treating multiple tissue samples mounted on microscope slides so that each sample can receive an individualized staining or treatment protocol even when such protocols require different temperature parameters. More specifically, what is described is an apparatus comprising a computer controlled, bar code driven, staining instrument that automatically applies chemical and biological reagents to tissue or cells mounted or affixed to standard glass microscope slides. A plurality of slides are mounted in a circular array on a carousel which rotates, as directed by the computer, to a dispensing location placing each slide under one of a series of reagent dispensers on a second rotating carousel positioned above the slides. Each slide receives the selected reagents (e.g. DNA probe) and is washed, mixed, and/or heated in an optimum sequence and for the required period of time.The instrument would have a micro-fluidic device that covers the slide using a gasket to keep volumes low. It would also contain micro-pipettes to aspirate the liquid off the slide and into a collection vial following UV illumination.

VII. Kits

[0185] Also disclosed herein are embodiments of a kit for use in disclosed systems. An exemplary kit for performing contextual diagnostics are is disclosed herein. The kit can comprise obvious components used to practice disclosed embodiments, such as a probe reagent solution comprising a specific binding moiety, a photo-cleavable moiety, and a detection tag.

VIII. Advancements over the Art

[0186] Embodiments of the methods disclosed herein provide advancements over methods currently used in the art for histopathology. In some embodiments, the disclosed method provides quantifiable results because labels and probes used in embodiments of the method are directly detected and do not require further amplification. Embodiments of the disclosed method are suitable for multiplexing assays as the detection tags used in certain

embodiments of the method are uniquely distinct with an infinite number of different species available for use in the disclosed probes. Many multiplexing methods known in the art are limited by the number of detectable tags that can be attached to ISH probes and/or the type of detection system used for immunohistochemistry (IHC) analysis (*e.g.*, colorimetric and/or fluorescent detection schemes). While methods exist to address such limitations, these methods typically require disassociating and lysing the tissue sample, thereby destroying all tissue context. The method disclosed herein avoids these undesirable steps and provides a level of maintained tissue contextual information that is important in IHC and ISH analysis. The disclosed method also provides the benefit of preserving tissue samples for repeated analyses, archiving, and/or conformational analysis. The disclosed method not only facilitates multiplexing assays, but certain embodiments further provide the ability to repeat multiple distinct multiplexing assays on different portions of a particular sample. Methods currently known in the art are unable to address tissue heterogeneity in the same way.

[0187] Embodiments of the disclosed method also offer levels of specificity and sensitivity that far exceed methods currently known in the art. For example, some embodiments concern using primer extension techniques to produce detectable moieties that can be isolated from a sample for further analysis. Such *in situ* extension can be conducted solely on the particular gene of interest, whereas all other sequence reads can be ignored.

[0188] In addition, embodiments of the disclosed method provide higher spatial resolution than is achieved with tissue microdissection methods used in the art. Methods disclosed herein also facilitate probing cellular structure and rather than solely the cell itself. The methods disclosed herein also can be automated, thereby providing fast, efficient, methods that can be conducted on bench-top automated devices.

IX. Working Embodiments

Probe Design:

[0189] In some embodiments, oligonucleotide DNA probes were synthesized with a 3' binding sequence to either a chromosome 17 alpha satellite repeat sequence (SEQ ID NO: 4) or an ALU repetitive sequence (SEQ ID NO: 5), and a 5' non-binding sequence. In some experiments, a nick-translated, DIG labeled clone of the α7t1alpha satellite repeat from the centromeric region of chromosome 7 was used as a non-UV sensitive control probe. Probes used for demonstration of tissue irradiation were designed with a 5' non-binding portion that contained two DNP (dinitrophenol) haptens (DNP-TATTTT-DNP-TATTTT), whereas probes used to demonstrate PCR detection of a DNA barcode were designed with a non-binding 5' tail that contained priming sites for M13 primers (Forward: 5' being SEQ ID NO: 6, Reverse: 5' being SEQ ID NO: 7**)** flanking a DNA unique sequence identifier (SEQ ID NO: 1). The two sequence elements were linked via a photo-cleavable moiety (Ambergen).

In Situ Hybridization:

[0190] *In Situ* Hybridization (ISH) procedures were executed on a BENCHMARK XT automated slide stainer. Probes were hybridized at 42 °C for two hours in a formamide containing buffer. Probes were washed, and slides were removed from the staining instrument and irradiated. Following UV irradiation, the hybridized probes were detected using Ventana SISH detection on a Benchmark XT instrument.

UV Irradiation:

[0191] An LED light source emitting at 365nm (XeLED-Ni1UV-R3-365, Xenopus Electronix) was used to irradiate the tissue at a distance of 2 cm for 30 seconds. Aluminum foil was placed over the tissue section to block the UV light. Subsequent irradiation methods used UV light from a metal-halide lamp (X-Cite, Lumen Dynamics) passed through a 20x microscope objective. UV energy doses ranged from 50 - 10 mW for ten seconds to one minute. Intensity of the DMD device UV source was 130 mW/cm2. The slides were irradiated for

20 seconds.

[0192] UV irradiation using an LCD screen also was used to shine a pattern on the tissue. The field diaphragm of a microscope using Kohler irradiation was replaced with a liquid crystal panel to turn the microscope into a projector. This allowed controlling which regions of the slide to irradiate by masking specific regions of this tissue from irradiation. For tissue irradiation, the LCD panel was transitioned into a checkerboard pattern. Slides were irradiated using 100 mW for 30 to 60 seconds.

Example 1

[0193] To determine whether UV-cleavable DNA oligonucleotides are suitable DNA probes, a CHR-17 oligonucleotide was designed with a photo-cleavable moiety between the target binding region and the detection region, which comprised detectable DNP labels. The probes were hybridized to a tissue sample; the slides were taken off the instrument and then irradiated with UV light. The slides were then put back on the instrument and the presence of the UV-cleavable DNA tag was determined using a silver ISH [SISH] detection kit. A control also was used with a nick-translate DIG CHR-7 probe.

[0194] The results indicated the ability of a probe containing a photo-cleavable moiety to be selectively irradiated and therefore useful in detecting a DNA target in a tissue sample. No DNP tags were detected in the control sample as the portion of the sample was exposed to light and all DNP tags were cleaved with the photo-cleavable moiety. In comparison, the masked areas (*i.e.*, non-irradiated portions) of the sample did contain DNP-labeled probes as the photo-cleavable moiety in these probes were not cleaved. The DNP labels attached to the probe through the photo-cleavable moiety were therefore detectable using SISH detection.

Example 2

[0195] In this embodiment, an ALU DNA ISH probe (comprising an oligonucleotide, photo-cleavable moiety, and tag sequence, labeled with DNP label) was designed to deposit a significant amount of SISH signal off of a DNA hapten attached to a DNA detection region of a DNA probe. The probe was

hybridized to various spots on the tissue sample, excess probe was washed, and the slides comprising the tissue samples were removed from the automated stainer and irradiated using one of three methods. In one embodiment, a sample was irradiated at 50 mW for approximately 20 seconds using a 20x objective. FIG. 11 provides an image of a control ALU probe that was not irradiated with light.

[0196] An additional, different sample slide was exposed to a different UV irradiation method. In this embodiment, an LCD screen was placed between a UV lamp and the slide. The LCD display was transitioned to provide a pattern by projecting a patterned slide and thereby making the UV light in the blacked-out portions of the pattern. The slide was irradiated for approximately 60 seconds and provided the patterned sample, to which a photomicrograph is shown in FIG. 12. As shown in FIG. 12, portions of the sample that were exposed to light did not provide a detectable signal, as the DNP tags attached to the probe through a photo-cleavable linker were detached from the sample, whereas those portions that were not irradiated did provide a detectable signal produced by SISH detection of the DNP tags. Additional embodiments are illustrated in FIGS. 13-16, which also show that embodiments of the disclosed method can be used to obtain patterned samples using selective irradiation. In particular, FIGS. 13 and 14 show two levels of magnification for a region that had been directly irradiated using a microscope objective. The spherical aspect of the irradiation is clearly observed in FIG. 13, while the contrast and sharpness of the interface between the irradiated and non-irradiated sections can be seen in FIG. 14.

[0197] FIGS. 15 and 16 show the versatility in shape and provide an indication of the resolution which can be accomplished using a light source with a digital micromirror device (DMD). In particular, FIG. 15 shows that an elaborate symbol can be illuminated while FIG. 16 shows the resolution of the approach, where each of the speckles represents quality control patterns associated with the DMD.

Example 3

[0198] In this particular embodiment, a CHR-17 oligonucleotide was used as a primer for DNA polymerase I. The probe was hybridized to tissue, which was then washed. Using nick-translation protocol, DNA polymerase I (*e.g.*, 10, 1, or 0.1 units of DNA polymerase per slide) and other extension reagents were added to the sample to extend the 3' end of the CHR-17 oligonucleotide, with an extension time of 2 hours. DIG-dUTP was then added to detect primer extension using an anti-DIG, AP Red detection kit. The results for these particular examples indicated that polymerase chain extension had occurred as the AP red labels were detected in the sample.

Example 4

[0199] In this particular embodiment, a DNA probe, comprising a photo-cleavable moiety and a unique sequence identifier (comprising a primer portion, a sequence identifier portion, and a second primer portion) was used. The DNA probe was hybridized to tonsil tissue on a BenchMark XT instrument. The slides were then removed and irradiated with UV light at 50 mW for approximately 20 seconds using a 20x objective. Liquid on the slides containing the liberated unique sequence identifiers was collected and used as a template in subsequent PCR reactions.

Example 5

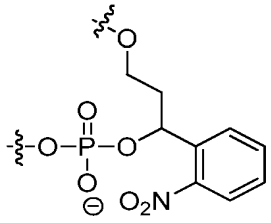
[0200] In one embodiment, a biotinylated secondary antibody was bound to a biotinylated tag sequence using streptavidin. An antigen retrieved target was first labeled with a primary antibody, followed by binding with the biotinylated anti-species secondary antibody. Streptavidin was used to connect the biotinylated tag with the biotinylated antibody. After a stringent wash, the tag was hybridized to a template oligonucleotide. The slide was then treated with exonuclease. All steps were performed on a VMSI Benchmark XT stainer. After washing and air-drying, the template was transferred into a solution for PCR quantification.

[0201] Series dilutions of two particular template oligonucleotides were made and the quantification dynamic range determined by using corresponding

primers. A seven orders of magnitude linear range was achieved with the detection of 120 molecules possible (2 µL of 0.1 fM). Run-to-run reproducibility also was demonstrated. There was little discrepancy between the two template oligonucleotides and therefore it is possible to develop multiplexed assays by using different template oligonucleotides with the same probe. The detection and quantification of the template oligonucleotides with sufficient sensitivity and linear dynamic range have been demonstrated. In some embodiments, a specific template was quantified even in the presence of other template oligonucleotides using a corresponding unique primer.

Example 6

[0202] In this particular embodiment, ALU and CHR17 oligo/polynucleotide sequences (SEQ ID NO: 5 and SEQ ID NO: 4, respectively) attached to the three unique sequence identifier sequences (SEQ ID NOS: 21-23) via a photocleavable linker (commercially available from TriLink, structure provided below) were designed. The probes (1 µg/ml final concentration) were hybridized to tonsil tissue and were exposed to UV irradiation using a digital mirror display (DMD) irradiation technique. Irradiation was carried out using the entire DMD array field (19,151 x 14,363 µm) with a UV intensity of 130 mW/cm^2. Control slides were left untreated with UV light.



Photocleavable Linker

[0203] After the irradiation, the liberated unique sequence identifiers were released into 200 µL of TE buffer solution and collected using a microfluidic platform. PCR amplification was performed in a 50 µL of total reaction using 20 µL of the solution comprising the unique sequence identifiers as a template and NGS specific (Illumina) adaptor sequences (SEQ ID NOS: 26 and 27) as primers. Amplification was performed using the following thermal cycling conditions: 98

°C for 10 seconds, 60 °C for 10 s, 72 °C for 1 min (15-20 cycles). The amplified unique sequence identifiers with Illumina adaptors were analyzed using gel electrophoresis (Invitrogen 4% EZ Gel). The results are shown in FIG. 17. As shown in FIG. 17, no signal was obtained for control samples that were not irradiated (and therefore did not provide unique sequence identifiers capable of being isolated) and for control samples that were not exposed to the probe comprising the unique sequence identifiers. Bands were obtained, however, for the unique sequence identifiers (attached to the NGS specific adaptor sequences) that were exposed to light.

[0204] Furthermore, we demonstrated the detection of single and multiplexed DNA barcodes corresponding to diagnostic biomarkers via Illumina Sequencing. In particular, ALU, CHR7 and CHR17 probes attached to the three unique barcode sequences were hybridized to the tonsil tissue separately and/or in multiplex. After washing, the barcodes were then released into the solution by UV illumination. The Illumina adaptors were attached to the barcodes in a PCR reaction using the solution containing the barcodes as the template (as described previously). Each PCR sample was purified, and the gel band of the target size (211 bp) was selected on 2% E-gel Size Select agarose gel (Invitrogen) resulting in sequencing-ready library for Illumina paired-end sequencing. The library was then quantified using the PicoGreen assay (Invitrogen Molecular Probes) and multiplexed with other libraries and ran on a Illumina MiSeq instrument (150 x 150 paired end).

[0205] The counts from the sequencing run were mapped using BOWTIE2 with a 68.79 % overall alignment rate (Table 1). In this particular experiment, barcode counts corresponding to CHR7 and CHR17 are in multiplex, *i.e.*, the probe cocktail containing CHR7 and CHR17 probes were hybridized simultaneously on the same tissue slide. The barcode counts corresponding to ALU is from a separate single ISH experiment. The libraries were multiplexed and analyzed on a single MiSeq run.

Table 1

Barcode	Centromeric Probe	MiSeq Counts	Binding events/cell
BC1	CHR7	37602	~200-1000
BC2	CHR17	118013	~3000
BC3	ALU	1398884	~ 1,000000

[0206] The barcode counts corresponding to CHR7 and CHR17 in multiplex correlate very well with the binding events demonstrating the sensitivity of the NGS sequencing to quantify biomarkers in multiplex.

Example 7

[0207] In this particular embodiment, the ability to extend off of target sequences within the tissue was tested using PE1.0-CHR7 (SEQ. ID. NO. 45) and PE1.0-CHR17 (SEQ. ID. NO. 46) probes. These experiments demonstrate polymerase extension without the added complexity of a photo-reactive step. The sequences were:

PE1.0-CHR7 (SEQ. ID. NO: 45):

- **5'-**AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG CTCTTCCGATCT**TCTAGCCATTTGATGCCAACAGTAGAAAGGG-3'**

PE1.0-CHR17 (SEQ. ID. NO: 46):

- **5'**AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCT CTTCCGATCT**CATTCAAATCCCCGAGTTGAACTTTCCTTTCAAAGTTC ACGT-3'**

[0208] The probes were hybridized to tissue with increasing concentrations of 10 ng/ml, 0.1 µg/ml and 1 µg/ml per slide. Following deparaffinization and cell conditioning on the Benchmark XT, single or a cocktail of PE1.0-CHR7 and PE1.0-CHR17 probes, dNTPs and FastTaq DNA PoLymerase (Roche Applied Sciences) were added on the slides. Following hybridization and extension at 60°C for 1 hour in a buffer comprising 50 mM Tris/HCl, 10 mM KCl, 5 mM

$(NH_4)_2SO_4$, 2 mM $MgCl_2$, the probe sequences were collected off of the slides following a high temperature denaturation. The extended sequences were selectively amplified using a forward primer within the Illumina NGS adaptor sequence (SEQ ID NO: 26). The distance of the *in situ* polymerase extension was detected using reverse primers (SEQ ID NOS: 29-33) complementary to the regions further from the probe sequences (FIG. 18). Amplification was performed using the following thermal cycling conditions: 98 °C for 10 seconds, 60 °C for 10 s, 72 °C for 1 min (15-20 cycles), with modifications of the annealing temperatures based on the melting temperatures of the primer pairs. The products were analyzed using gel electrophoresis (Invitrogen 4% EZ Gel), and the results are provided in FIGS. 19 and 20.

[0209] As illustrated in FIGS. 19 and 20, as the concentration of each probe is increased, the obtained signal increases. FIGS. 19 and 20 also show signals obtained from the reverse primer extension, which indicates that the extension reaction is capable of producing extended products having high numbers of base pairs. As illustrated in FIG. 20, the method produced an extended product having about 230 base pairs. The results obtained from this particular embodiment illustrate the ability to extend primers that are hybridized to tissue without having to first isolate a section of the tissue sample. Rather, the primer can be added to a tissue sample, extended, amplified, and detected without ever having to destroy tissue context by physically excising and/or manipulating the tissue sample. Furthermore, it demonstrates the ability to change the detection length. The three extension lengths were selected to show an ability to extend to various lengths.

Example 8

[0210] In this particular embodiment, CHR7 and CHR17 primers (SEQ ID NOs: 20 and 4, respectively) were hybridized to a tissue sample simultaneously with concentrations of 0.1 μg/ml each. The collected, extended sequences were amplified via duplex PCR using both of the extended CHR7 and CHR17 reverse primers (FIG. 21). This particular embodiment establishes that the disclosed methods are suitable for multiplexing assays. Again, these were done

in the absence of a photo-reactive step to show robustness of the underlying extension approach.

Example 9

[0211] Referring now to FIGS. 22(A) – (B), FIG 22(A) is a schematic showing that an approach by which the sequences could be purified using a biotin molecule at the 5'end (square). PS1 and CHR 7R5 (which is complementary to the region further from the probe sequence) were used to detect *in situ* polymerase extension. FIG. 22(B) is a photograph of a gel showing that subsequent PCR amplification, a 230 bp extended product was detected with the control probe Biotin-PS1-CHR7 with no caged thymidines (Lanes 1-2).

[0212] In this particular embodiment, light selective *in situ* polymerase extension using light-activated oligos (LANs) is demonstrated. In particular, light-activated oligonucleotides (LANs) (TriLink technologies) were used to demonstrate selective polymerase extension in tissue. LANs contain photolabile nucleobase modifications, nitropiperonyloxymethyl (NPOM), to one or more of the thymidines. The NPOM-caging group interrupts hybridization by sterically blocking Watson-Crick base pairing. Reference is made to Young et al., 2008 Light triggered polymerase chain reaction, *Chem. Commun.*, 2008, 462-464, which is hereby incorporated by reference, herein, in its entirety for disclosure relate to *in situ* polymerase extension.

[0213] CHR7 probe was synthesized with a nonhuman priming sequence (PS1: SEQ. ID NO. 42) and a biotin molecule at the 5'-end and 'caged' thymidines distributed towards to 3'-end. The non-caged version of the probe was used as a positive control (FIG. 22(B), lanes 1-2).

[0214] The probes were hybridized to tissues on BenchMark XT instrument at 60˚C and 65˚C for 1 hour. After the hybridization, slides that will be illuminated by UV were taken off the instrument (FIG. 22(B), lanes 5-6) while the rest were kept dark (FIG. 22(B), lanes 3-4) in the instrument. The slides were treated with UV for 1 min with a UV lamp for slide illumination and put back in the instrument. The unbound material was removed from the slide with a stringency wash. The extension buffer containing Sequenase V2.0 DNA

polymerase (Affymetrix, Inc) and dNTPs (Roche) were added on the slides to allow polymerase extension. The probe sequences were collected from the slides following a high temperature denaturation and purified with Dynabeads M-280 Streptavidin beads (Invitrogen). Purified sequences bound to the streptavidin beads were used as a template for detection PCR. Light selective in-situ polymerase extension was demonstrated by using PS1 as a forward primer and CHR7R5 reverse primer complementary to the regions 230 bp further from the probe sequence. The products were analyzed using gel electrophoresis (Invitrogen 4% EZ Gel).

[0215] While the probes from the slides without UV illumination failed to produce extended product due to the caged thymidines (Lanes 3-4), the probes on the UV illuminated slides were successfully extended and produced the expected size band (Lanes 5-6) at both hybridization temperatures. Primer and probe sequences used in this study:

- PS1 (SEQ. ID NO: 42): 5'-TAACTTACGGAGTCGCTCTACG-3'
- CHR7 (SEQ. ID NO: 43): 5'-TCTAGCCATTTGATGCCAACAGTAG AAAGGG-3'
- Biotin-PS1-CHR7LAN5C (SEQ. ID NO: 44): 5'/{Biotin}/ TAACTTACGGA GTCGCTCTACGTCTAGCCAT*T*T*GAT*GCCAACAGT*AGAAAGGG-3'
- *T denotes positions of caged thymidines in caged versions of the probes.

Example 10

[0216] In this embodiment, we demonstrate photo-selective *in situ* polymerase extension of single genes. Referring now to FIG. 23(A)-(B), probes specific for BRAF Exon 13 (FIG. 23(A)) and for HER2 Exon 10 (FIG. 23(B)), was designed with and without caged thymidines. Both primers were hybridized and polymerase extended at 60 ˚C for 1 hour. The slide which was hybridized with the caged thymidines was treated with UV light, while the remainder was kept in the dark. Lane 1 (exclusive of the marker lanes) shows the control probe with no caged thymidines; thus, extension products are generated and

detected as expected. Lane 2 shows the caged thymidine containing probe which had not been subjected to light. It is apparent from the gel that no extension occurred. Lane 3 shows the caged thymidine containing probe subjected to UV light and the extended product is detected at a level similar to that expected from the control probe (i.e. probe without caged thymidine). The light activated nucleotides were included at 5 different locations in both primers. Both exon primers were biotin labelled at the 5' end, and extended primers were purified with streptavidin beads following denaturation at 98 °C for 3-4 min. The aliquots of the purified products were PCR amplified with the corresponding primer pairs and detected with Invitrogen 4 % E-gel. The following probe sequences used in this study:

- Biotin-HER2Exon10 (SEQ. ID NO: 47): 5'-Biotin-TAACTTACGGAGTC GCTCTACGGCATGGAGCACT*T*GCGAG AGGT*GAGGGCAGT*T*A-3'
- BiotinBRAFExon13 (SEQ. ID NO: 48): 5'-Biotin-TAACTTACGGAG TCGCTCTACGAGTGGTGTGAGGGCTCCAGCTTGTAT*CACCAT*CT*A T*T*G-3'
- *T denotes positions of caged thymidines in caged versions of the probes.

SEQUENCE LISTING

The nucleic and amino acid sequences provided herein are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids, as defined in 37 C.F.R. § 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

SEQ ID NO: 1 is an exemplary sequence for a unique sequence identifier.
ATGGTATCCTTACATCGTCATTTATC

SEQ ID NO: 2 is another exemplary sequence for a unique sequence identifier.
CCTCTCTATGGGCAGTCGGTGATTCCTATTCAGATCACGATGGTATCCTTACATCT
GAGTCGGAGACACGCAGGGATGAGATGG

SEQ ID NO: 3 is yet another exemplary sequence for a unique sequence identifier.
CCTCTCTATGGGCAGTCGGTGATTCCTATTCAGATGGTATCCTTACATCTGAGTCG
GAGACACGCAGGGATGAGATGG

SEQ ID NO: 4 is an exemplary sequence for an oligo/polynucleotide specific binding moiety.
CATTCAAATCCCCGAGTTGAACTTTCCTTTCAAAGTTCACGT

SEQ ID NO: 5 is another exemplary sequence for an oligo/polynucleotide specific binding moiety.
CGGGAGGCGGAGGTTGCAGTGAGCC

SEQ ID NO: 6 is an exemplary primer sequence for use with a probe disclosed herein.
GTAAAACGACGGCCAGT

SEQ ID NO: 7 is another exemplary primer sequence for use with a probe disclosed herein.

CAGGAAACAGCTATGAC

SEQ ID NO: 8 is a Her1 amino acid sequence (NCBI Reference Sequence No. NP_005219.2).

MRPSGTAGAALLALLAALCPASRALEEKKVCQGTSNKLTQLGTFEDHFLSLQRMFN
NCEVVLGNLEITYVQRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNMYYEN
SYALAVLSNYDANKTGLKELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSD
FLSNMSMDFQNHLGSCQKCDPSCPNGSCWGAGEENCQKLTKIICAQQCSGRCRG
KSPSDCCHNQCAAGCTGPRESDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDV
NPEGKYSFGATCVKKCPRNYVVTDHGSCVRACGADSYEMEEDGVRKCKKCEGPCR
KVCNGIGIGEFKDSLSINATNIKHFKNCTSISGDLHILPVAFRGDSFTHTPPLDPQELDI
LKTVKEITGFLLIQAWPENRTDLHAFENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLK
EISDGDVIISGNKNLCYANTINWKKLFGTSGQKTKIISNRGENSCKATGQVCHALCSP
EGCWGPEPRDCVSCRNVSRGRECVDKCNLLEGEPREFVENSECIQCHPECLPQAM
NITCTGRGPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYADAGHVCHLCHP
NCTYGCTGPGLEGCPTNGPKIPSIATGMVGALLLLLVVALGIGLFMRRRHIVRKRTLR
RLLQERELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKI
PVAIKELREATSPKANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCL
LDYVREHKDNIGSQYLLNWCVQIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKIT
DFGLAKLLGAEEKEYHAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTF
GSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELIIEFSK
MARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIPQQGFFS
SPSTSRTPLLSSLSATSNNSTVACIDRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDD
TFLPVPEYINQSVPKRPAGSVQNPVYHNQPLNPAPSRDPHYQDPHSTAVGNPEYLN
TVQPTCVNSTFDSPAHWAQKGSHQISLDNPDYQQDFFPKEAKPNGIFKGSTAENAE
YLRVAPQSSEFIGA

SEQ ID NO: 9 is an exemplary cDNA sequence encoding Her1 (NCBI Reference Sequence No. NM_005228.3).

CCCCGGCGCAGCGCGGCCGCAGCAGCCTCCGCCCCCCGCACGGTGTGAGCGC
CCGACGCGGCCGAGGCGGCCGGAGTCCCGAGCTAGCCCCGGCGGCCGCCGCC
GCCCAGACCGGACGACAGGCCACCTCGTCGGCGTCCGCCCGAGTCCCCGCCTC
GCCGCCAACGCCACAACCACCGCGCACGGCCCCCTGACTCCGTCCAGTATTGA
TCGGGAGAGCCGGAGCGAGCTCTTCGGGGAGCAGCGATGCGACCCTCCGGGAC
GGCCGGGGCAGCGCTCCTGGCGCTGCTGGCTGCGCTCTGCCCGGCGAGTCGGG
CTCTGGAGGAAAAGAAAGTTTGCCAAGGCACGAGTAACAAGCTCACGCAGTTG
GGCACTTTTGAAGATCATTTTCTCAGCCTCCAGAGGATGTTCAATAACTGTGAGG
TGGTCCTTGGGAATTTGGAAATTACCTATGTGCAGAGGAATTATGATCTTTCCTTC
TTAAAGACCATCCAGGAGGTGGCTGGTTATGTCCTCATTGCCCTCAACACAGTGG
AGCGAATTCCTTTGGAAAACCTGCAGATCATCAGAGGAAATATGTACTACGAAA
ATTCCTATGCCTTAGCAGTCTTATCTAACTATGATGCAAATAAAACCGGACTGAA
GGAGCTGCCCATGAGAAATTTACAGGAAATCCTGCATGGCGCCGTGCGGTTCAG
CAACAACCCTGCCCTGTGCAACGTGGAGAGCATCCAGTGGCGGGACATAGTCA
GCAGTGACTTTCTCAGCAACATGTCGATGGACTTCCAGAACCACCTGGGCAGCT
GCCAAAAGTGTGATCCAAGCTGTCCCAATGGGAGCTGCTGGGGTGCAGGAGAG
GAGAACTGCCAGAAACTGACCAAAATCATCTGTGCCCAGCAGTGCTCCGGGCG
CTGCCGTGGCAAGTCCCCCAGTGACTGCTGCCACAACCAGTGTGCTGCAGGCTG
CACAGGCCCCCGGGAGAGCGACTGCCTGGTCTGCCGCAAATTCCGAGACGAAG
CCACGTGCAAGGACACCTGCCCCCCACTCATGCTCTACAACCCCACCACGTACC
AGATGGATGTGAACCCCGAGGGCAAATACAGCTTTGGTGCCACCTGCGTGAAGA
AGTGTCCCCGTAATTATGTGGTGACAGATCACGGCTCGTGCGTCCGAGCCTGTG
GGGCCGACAGCTATGAGATGGAGGAAGACGGCGTCCGCAAGTGTAAGAAGTGC
GAAGGGCCTTGCCGCAAAGTGTGTAACGGAATAGGTATTGGTGAATTTAAAGAC
TCACTCTCCATAAATGCTACGAATATTAAACACTTCAAAAACTGCACCTCCATCA
GTGGCGATCTCCACATCCTGCCGGTGGCATTTAGGGGTGACTCCTTCACACATAC
TCCTCCTCTGGATCCACAGGAACTGGATATTCTGAAAACCGTAAAGGAAATCAC
AGGGTTTTTGCTGATTCAGGCTTGGCCTGAAAACAGGACGGACCTCCATGCCTTT
GAGAACCTAGAAATCATACGCGGCAGGACCAAGCAACATGGTCAGTTTTCTCTT
GCAGTCGTCAGCCTGAACATAACATCCTTGGGATTACGCTCCCTCAAGGAGATA
AGTGATGGAGATGTGATAATTTCAGGAAACAAAAATTTGTGCTATGCAAATACAA

TAAACTGGAAAAAACTGTTTGGGACCTCCGGTCAGAAAACCAAAATTATAAGCA
ACAGAGGTGAAAACAGCTGCAAGGCCACAGGCCAGGTCTGCCATGCCTTGTGC
TCCCCCGAGGGCTGCTGGGGCCCGGAGCCCAGGGACTGCGTCTCTTGCCGGAA
TGTCAGCCGAGGCAGGGAATGCGTGGACAAGTGCAACCTTCTGGAGGGTGAGC
CAAGGGAGTTTGTGGAGAACTCTGAGTGCATACAGTGCCACCCAGAGTGCCTGC
CTCAGGCCATGAACATCACCTGCACAGGACGGGGACCAGACAACTGTATCCAG
TGTGCCCACTACATTGACGGCCCCCACTGCGTCAAGACCTGCCCGGCAGGAGTC
ATGGGAGAAAACAACACCCTGGTCTGGAAGTACGCAGACGCCGGCCATGTGTG
CCACCTGTGCCATCCAAACTGCACCTACGGATGCACTGGGCCAGGTCTTGAAGG
CTGTCCAACGAATGGGCCTAAGATCCCGTCCATCGCCACTGGGATGGTGGGGGC
CCTCCTCTTGCTGCTGGTGGTGGCCCTGGGGATCGGCCTCTTCATGCGAAGGCGC
CACATCGTTCGGAAGCGCACGCTGCGGAGGCTGCTGCAGGAGAGGGAGCTTGT
GGAGCCTCTTACACCCAGTGGAGAAGCTCCCAACCAAGCTCTCTTGAGGATCTT
GAAGGAAACTGAATTCAAAAAGATCAAAGTGCTGGGCTCCGGTGCGTTCGGCAC
GGTGTATAAGGGACTCTGGATCCCAGAAGGTGAGAAAGTTAAAATTCCCGTCGC
TATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAACAAGGAAATCCTCG
ATGAAGCCTACGTGATGGCCAGCGTGGACAACCCCCACGTGTGCCGCCTGCTG
GGCATCTGCCTCACCTCCACCGTGCAGCTCATCACGCAGCTCATGCCCTTCGGCT
GCCTCCTGGACTATGTCCGGGAACACAAAGACAATATTGGCTCCCAGTACCTGC
TCAACTGGTGTGTGCAGATCGCAAAGGGCATGAACTACTTGGAGGACCGTCGCT
TGGTGCACCGCGACCTGGCAGCCAGGAACGTACTGGTGAAAACACCGCAGCAT
GTCAAGATCACAGATTTTGGGCTGGCCAAACTGCTGGGTGCGGAAGAGAAAGA
ATACCATGCAGAAGGAGGCAAAGTGCCTATCAAGTGGATGGCATTGGAATCAAT
TTTACACAGAATCTATACCCACCAGAGTGATGTCTGGAGCTACGGGGTGACCGTT
TGGGAGTTGATGACCTTTGGATCCAAGCCATATGACGGAATCCCTGCCAGCGAG
ATCTCCTCCATCCTGGAGAAAGGAGAACGCCTCCCTCAGCCACCCATATGTACC
ATCGATGTCTACATGATCATGGTCAAGTGCTGGATGATAGACGCAGATAGTCGC
CCAAAGTTCCGTGAGTTGATCATCGAATTCTCCAAAATGGCCCGAGACCCCCAG
CGCTACCTTGTCATTCAGGGGGATGAAAGAATGCATTTGCCAAGTCCTACAGACT
CCAACTTCTACCGTGCCCTGATGGATGAAGAAGACATGGACGACGTGGTGGATG
CCGACGAGTACCTCATCCCACAGCAGGGCTTCTTCAGCAGCCCCTCCACGTCAC

GGACTCCCCTCCTGAGCTCTCTGAGTGCAACCAGCAACAATTCCACCGTGGCTT
GCATTGATAGAAATGGGCTGCAAAGCTGTCCCATCAAGGAAGACAGCTTCTTGC
AGCGATACAGCTCAGACCCCACAGGCGCCTTGACTGAGGACAGCATAGACGAC
ACCTTCCTCCCAGTGCCTGAATACATAAACCAGTCCGTTCCCAAAAGGCCCGCT
GGCTCTGTGCAGAATCCTGTCTATCACAATCAGCCTCTGAACCCCGCGCCCAGC
AGAGACCCACACTACCAGGACCCCCACAGCACTGCAGTGGGCAACCCCGAGTA
TCTCAACACTGTCCAGCCCACCTGTGTCAACAGCACATTCGACAGCCCTGCCCA
CTGGGCCCAGAAAGGCAGCCACCAAATTAGCCTGGACAACCCTGACTACCAGC
AGGACTTCTTTCCCAAGGAAGCCAAGCCAAATGGCATCTTTAAGGGCTCCACAG
CTGAAAATGCAGAATACCTAAGGGTCGCGCCACAAAGCAGTGAATTTATTGGAG
CATGACCACGGAGGATAGTATGAGCCCTAAAAATCCAGACTCTTTCGATACCCA
GGACCAAGCCACAGCAGGTCCTCCATCCCAACAGCCATGCCCGCATTAGCTCTT
AGACCCACAGACTGGTTTTGCAACGTTTACACCGACTAGCCAGGAAGTACTTCC
ACCTCGGGCACATTTTGGGAAGTTGCATTCCTTTGTCTTCAAACTGTGAAGCATTT
ACAGAAACGCATCCAGCAAGAATATTGTCCCTTTGAGCAGAAATTTATCTTTCAA
AGAGGTATATTTGAAAAAAAAAAAAAGTATATGTGAGGATTTTTATTGATTGGGG
ATCTTGGAGTTTTTCATTGTCGCTATTGATTTTTACTTCAATGGGCTCTTCCAACAA
GGAAGAAGCTTGCTGGTAGCACTTGCTACCCTGAGTTCATCCAGGCCCAACTGT
GAGCAAGGAGCACAAGCCACAAGTCTTCCAGAGGATGCTTGATTCCAGTGGTTC
TGCTTCAAGGCTTCCACTGCAAAACACTAAAGATCCAAGAAGGCCTTCATGGCC
CCAGCAGGCCGGATCGGTACTGTATCAAGTCATGGCAGGTACAGTAGGATAAGC
CACTCTGTCCCTTCCTGGGCAAAGAAGAAACGGAGGGGATGGAATTCTTCCTTA
GACTTACTTTTGTAAAAATGTCCCCACGGTACTTACTCCCCACTGATGGACCAGT
GGTTTCCAGTCATGAGCGTTAGACTGACTTGTTTGTCTTCCATTCCATTGTTTTGAA
ACTCAGTATGCTGCCCCTGTCTTGCTGTCATGAAATCAGCAAGAGAGGATGACA
CATCAAATAATAACTCGGATTCCAGCCCACATTGGATTCATCAGCATTTGGACCA
ATAGCCCACAGCTGAGAATGTGGAATACCTAAGGATAGCACCGCTTTTGTTCTCG
CAAAAACGTATCTCCTAATTTGAGGCTCAGATGAAATGCATCAGGTCCTTTGGGG
CATAGATCAGAAGACTACAAAAATGAAGCTGCTCTGAAATCTCCTTTAGCCATCA
CCCCAACCCCCCCAAAATTAGTTTGTGTTACTTATGGAAGATAGTTTTCTCCTTTTA
CTTCACTTCAAAAGCTTTTTACTCAAAGAGTATATGTTCCCTCCAGGTCAGCTGCC

CCCAAACCCCCTCCTTACGCTTTGTCACACAAAAAGTGTCTCTGCCTTGAGTCAT
CTATTCAAGCACTTACAGCTCTGGCCACAACAGGGCATTTTACAGGTGCGAATG
ACAGTAGCATTATGAGTAGTGTGGAATTCAGGTAGTAAATATGAAACTAGGGTTT
GAAATTGATAATGCTTTCACAACATTTGCAGATGTTTTAGAAGGAAAAAAGTTCC
TTCCTAAAATAATTTCTCTACAATTGGAAGATTGGAAGATTCAGCTAGTTAGGAG
CCCACCTTTTTTCCTAATCTGTGTGTGCCCTGTAACCTGACTGGTTAACAGCAGTC
CTTTGTAAACAGTGTTTTAAACTCTCCTAGTCAATATCCACCCCATCCAATTTATC
AAGGAAGAAATGGTTCAGAAAATATTTTCAGCCTACAGTTATGTTCAGTCACACA
CACATACAAAATGTTCCTTTTGCTTTTAAAGTAATTTTTGACTCCCAGATCAGTCA
GAGCCCCTACAGCATTGTTAAGAAAGTATTTGATTTTTGTCTCAATGAAAATAAA
ACTATATTCATTTCCACTCTAAAAAAAAAAAAAAAAA

SEQ ID NO: 10 is a Her2 amino acid sequence (NCBI Reference Sequence No. NP_004439.2).

MELAALCRWGLLLALLPPGAASTQVCTGTDMKLRLPASPETHLDMLRHLYQGCQV
VQGNLELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTQLFEDNY
ALAVLDNGDPLNNTTPVTGASPGGLRELQLRSLTEILKGGVLIQRNPQLCYQDTILW
KDIFHKNNQLALTLIDTNRSRACHPCSPMCKGSRCWGESSEDCQSLTRTVCAGGCA
RCKGPLPTDCCHEQCAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFES
MPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVCPLHNQEVTAEDGTQRCEKCSKP
CARVCYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQP
EQLQVFETLEEITGYLYISAWPDSLPDLSVFQNLQVIRGRILHNGAYSLTLQGLGISWL
GLRSLRELGSGLALIHHNTHLCFVHTVPWDQLFRNPHQALLHTANRPEDECVGEGL
ACHQLCARGHCWGPGPTQCVNCSQFLRGQECVEECRVLQGLPREYVNARHCLPC
HPECQPQNGSVTCFGPEADQCVACAHYKDPPFCVARCPSGVKPDLSYMPIWKFPD
EEGACQPCPINCTHSCVDLDDKGCPAEQRASPLTSIISAVVGILLVVVLGVVFGILIKR
RQQKIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKVLGSGAFGTV
YKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSPYVSRLLGICLTST
VQLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVRLVHRDLA
ARNVLVKSPNHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFTHQSD
VWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWMI

DSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDMGDL
VDAEEYLVPQQGFFCPDPAPGAGGMVHHRHRSSSTRSGGGDLTLGLEPSEEEAPRS
PLAPSEGAGSDVFDGDLGMGAAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVA
PLTCSPQPEYVNQPDVRPQPPSPREGPLPAARPAGATLERPKTLSPGKNGVVKDVFA
FGGAVENPEYLTPQGGAAPQPHPPPAFSPAFDNLYYWDQDPPERGAPPSTFKGTPT
AENPEYLGLDVPV

SEQ ID NO: 11 is an exemplary cDNA sequence encoding Her2 (NCBI Reference Sequence No. NM_004448.2).

GGAGGAGGTGGAGGAGGAGGGCTGCTTGAGGAAGTATAAGAATGAAGTTGTGA
AGCTGAGATTCCCCTCCATTGGGACCGGAGAAACCAGGGGAGCCCCCCGGGCA
GCCGCGCGCCCCTTCCCACGGGGCCCTTTACTGCGCCGCGCGCCCGGCCCCCA
CCCCTCGCAGCACCCCGCGCCCCGCGCCCTCCCAGCCGGGTCCAGCCGGAGCC
ATGGGGCCGGAGCCGCAGTGAGCACCATGGAGCTGGCGGCCTTGTGCCGCTGG
GGGCTCCTCCTCGCCCTCTTGCCCCCCGGAGCCGCGAGCACCCAAGTGTGCACC
GGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACCTGGACAT
GCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTCA
CCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAGGATATCCAGGAGGTGC
AGGGCTACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGG
CTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTG
CTAGACAATGGAGACCCGCTGAACAATACCACCCCTGTCACAGGGGCCTCCCC
AGGAGGCCTGCGGGAGCTGCAGCTTCGAAGCCTCACAGAGATCTTGAAAGGAG
GGGTCTTGATCCAGCGGAACCCCCAGCTCTGCTACCAGGACACGATTTTGTGGA
AGGACATCTTCCACAAGAACAACCAGCTGGCTCTCACACTGATAGACACCAACC
GCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCTCCCGCTGCTGGG
GAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCGGTGGCT
GTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTG
CCGGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACC
ACAGTGGCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACA
CGTTTGAGTCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTG
TGACTGCCTGTCCCTACAACTACCTTTCTACGGACGTGGGATCCTGCACCCTCGT

CTGCCCCCTGCACAACCAAGAGGTGACAGCAGAGGATGGAACACAGCGGTGTG
AGAAGTGCAGCAAGCCCTGTGCCCGAGTGTGCTATGGTCTGGGCATGGAGCACT
TGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCCAGGAGTTTGCTGGCTGCA
AGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTGATGGGGACCCAG
CCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGACTCTGG
AAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACC
TCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCG
CCTACTCGCTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCAC
TGAGGGAACTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCT
TCGTGCACACGGTGCCCTGGGACCAGCTCTTTCGGAACCCGCACCAAGCTCTGC
TCCACACTGCCAACCGGCCAGAGGACGAGTGTGTGGGCGAGGGCCTGGCCTGC
CACCAGCTGTGCGCCCGAGGGCACTGCTGGGGTCCAGGGCCCACCCAGTGTGT
CAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGGAGGAATGCCGAGTACT
GCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGCCGTGCCACCC
TGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAGGCTGACCA
GTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCC
AGCGGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAG
GAGGGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTG
GATGACAAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCAT
CTCTGCGGTGGTTGGCATTCTGCTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCC
TCATCAAGCGACGGCAGCAGAAGATCCGGAAGTACACGATGCGGAGACTGCTG
CAGGAAACGGAGCTGGTGGAGCCGCTGACACCTAGCGGAGCGATGCCCAACC
AGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAGGTGAAGGTGCTT
GGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGATGGGGAG
AATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAA
GCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTCCCCA
TATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACAC
AGCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCC
TGGGCTCCCAGGACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCT
ACCTGGAGGATGTGCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTG
GTCAAGAGTCCCAACCATGTCAAAATTACAGACTTCGGGCTGGCTCGGCTGCTG

GACATTGACGAGACAGAGTACCATGCAGATGGGGGCAAGGTGCCCATCAAGTG
GATGGCGCTGGAGTCCATTCTCCGCCGGCGGTTCACCCACCAGAGTGATGTGTG
GAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCCAAACCTTACGAT
GGGATCCCAGCCCGGGAGATCCCTGACCTGCTGGAAAAGGGGGAGCGGCTGCC
CCAGCCCCCCATCTGCACCATTGATGTCTACATGATCATGGTCAAATGTTGGATG
ATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCA
TGGCCAGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCAG
CCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGG
GGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCCAGCAGGGCTTCTTCTGTC
CAGACCCTGCCCCGGGCGCTGGGGGCATGGTCCACCACAGGCACCGCAGCTCA
TCTACCAGGAGTGGCGGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGA
GGAGGCCCCCAGGTCTCCACTGGCACCCTCCGAAGGGGCTGGCTCCGATGTATT
TGATGGTGACCTGGGAATGGGGGCAGCCAAGGGGCTGCAAAGCCTCCCCACAC
ATGACCCCAGCCCTCTACAGCGGTACAGTGAGGACCCCACAGTACCCCTGCCCT
CTGAGACTGATGGCTACGTTGCCCCCCTGACCTGCAGCCCCCAGCCTGAATATG
TGAACCAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGC
CTGCTGCCCGACCTGCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAG
GGAAGAATGGGGTCGTCAAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACC
CCGAGTACTTGACACCCCAGGGAGGAGCTGCCCCTCAGCCCCACCCTCCTCCTG
CCTTCAGCCCAGCCTTCGACAACCTCTATTACTGGGACCAGGACCCACCAGAGC
GGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAGAACCCAGA
GTACCTGGGTCTGGACGTGCCAGTGTGAACCAGAAGGCCAAGTCCGCAGAAGC
CCTGATGTGTCCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTGGCATCAAGAGG
TGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACCTGTC
CTAAGGAACCTTCCTTCCTGCTTGAGTTCCCAGATGGCTGGAAGGGGTCCAGCCT
CGTTGGAAGAGGAACAGCACTGGGGAGTCTTTGTGGATTCTGAGGCCCTGCCCA
ATGAGACTCTAGGGTCCAGTGGATGCCACAGCCCAGCTTGGCCCTTTCCTTCCA
GATCCTGGGTACTGAAAGCCTTAGGGAAGCTGGCCTGAGAGGGGAAGCGGCCC
TAAGGGAGTGTCTAAGAACAAAAGCGACCCATTCAGAGACTGTCCCTGAAACCT
AGTACTGCCCCCCATGAGGAAGGAACAGCAATGGTGTCAGTATCCAGGCTTTGT
ACAGAGTGCTTTTCTGTTTAGTTTTTACTTTTTTTGTTTTGTTTTTTTAAAGATGAAAT

AAAGACCCAGGGGGAGAATGGGTGTTGTATGGGGAGGCAAGTGTGGGGGGTCC
TTCTCCACACCCACTTTGTCCATTTGCAAATATATTTTGGAAAACAGCTA

SEQ ID NO: 12 is a Her3 amino acid sequence (NCBI Reference Sequence No. NP_001973.2).

MRANDALQVLGLLFSLARGSEVGNSQAVCPGTLNGLSVTGDAENQYQTLYKLYER
CEVVMGNLEIVLTGHNADLSFLQWIREVTGYVLVAMNEFSTLPLPNLRVVRGTQVY
DGKFAIFVMLNYNTNSSHALRQLRLTQLTEILSGGVYIEKNDKLCHMDTIDWRDIVR
DRDAEIVVKDNGRSCPPCHEVCKGRCWGPGSEDCQTLTKTICAPQCNGHCFGPNP
NQCCHDECAGGCSGPQDTDCFACRHFNDSGACVPRCPQPLVYNKLTFQLEPNPHT
KYQYGGVCVASCPHNFVVDQTSCVRACPPDKMEVDKNGLKMCEPCGGLCPKACE
GTGSGSRFQTVDSSNIDGFVNCTKILGNLDFLITGLNGDPWHKIPALDPEKLNVFRTV
REITGYLNIQSWPPHMHNFSVFSNLTTIGGRSLYNRGFSLLIMKNLNVTSLGFRSLKEI
SAGRIYISANRQLCYHHSLNWTKVLRGPTEERLDIKHNRPRRDCVAEGKVCDPLCSS
GGCWGPGPGQCLSCRNYSRGGVCVTHCNFLNGEPREFAHEAECFSCHPECQPME
GTATCNGSGSDTCAQCAHFRDGPHCVSSCPHGVLGAKGPIYKYPDVQNECRPCHE
NCTQGCKGPELQDCLGQTLVLIGKTHLTMALTVIAGLVVIFMMLGGTFLYWRGRRIQ
NKRAMRRYLERGESIEPLDPSEKANKVLARIFKETELRKLKVLGSGVFGTVHKGVWIP
EGESIKIPVCIKVIEDKSGRQSFQAVTDHMLAIGSLDHAHIVRLLGLCPGSSLQLVTQY
LPLGSLLDHVRQHRGALGPQLLLNWGVQIAKGMYYLEEHGMVHRNLAARNVLLK
SPSQVQVADFGVADLLPPDDKQLLYSEAKTPIKWMALESIHFGKYTHQSDVWSYGV
TVWELMTFGAEPYAGLRLAEVPDLLEKGERLAQPQICTIDVYMVMVKCWMIDENIR
PTFKELANEFTRMARDPPRYLVIKRESGPGIAPGPEPHGLTNKKLEEVELEPELDLDL
DLEAEEDNLATTTLGSALSLPVGTLNRPRGSQSLLSPSSGYMPMNQGNLGESCQES
AVSGSSERCPRPVSLHPMPRGCLASESSEGHVTGSEAELQEKVSMCRSRSRSRSPRP
RGDSAYHSQRHSLLTPVTPLSPPGLEEEDVNGYVMPDTHLKGTPSSREGTLSSVGLS
SVLGTEEEDEDEEYEYMNRRRRHSPPHPPRPSSLEELGYEYMDVGSDLSASLGSTQ
SCPLHPVPIMPTAGTTPDEDYEYMNRQRDGGGPGGDYAAMGACPASEQGYEEMR
AFQGPGHQAPHVHYARLKTLRSLEATDSAFDNPDYWHSRLFPKANAQRT

SEQ ID NO: 13 is an exemplary cDNA sequence encoding Her3 (NCBI

Reference Sequence No. NM_001982.3).

ACTCCAGCCTCGCGCGGGAGGGGGCGCGGCCGTGACTCACCCCCTTCCCTCTG
CGTTCCTCCCTCCCTCTCTCTCTCTCTCTCACACACACACACCCCTCCCCTGCCAT
CCCTCCCCGGACTCCGGCTCCGGCTCCGATTGCAATTTGCAACCTCCGCTGCCGT
CGCCGCAGCAGCCACCAATTCGCCAGCGGTTCAGGTGGCTCTTGCCTCGATGTC
CTAGCCTAGGGGCCCCCGGGCCGGACTTGGCTGGGCTCCCTTCACCCTCTGCGG
AGTCATGAGGGCGAACGACGCTCTGCAGGTGCTGGGCTTGCTTTTCAGCCTGGC
CCGGGGCTCCGAGGTGGGCAACTCTCAGGCAGTGTGTCCTGGGACTCTGAATGG
CCTGAGTGTGACCGGCGATGCTGAGAACCAATACCAGACACTGTACAAGCTCTA
CGAGAGGTGTGAGGTGGTGATGGGGAACCTTGAGATTGTGCTCACGGGACACA
ATGCCGACCTCTCCTTCCTGCAGTGGATTCGAGAAGTGACAGGCTATGTCCTCGT
GGCCATGAATGAATTCTCTACTCTACCATTGCCCAACCTCCGCGTGGTGCGAGG
GACCCAGGTCTACGATGGGAAGTTTGCCATCTTCGTCATGTTGAACTATAACACC
AACTCCAGCCACGCTCTGCGCCAGCTCCGCTTGACTCAGCTCACCGAGATTCTG
TCAGGGGGTGTTTATATTGAGAAGAACGATAAGCTTTGTCACATGGACACAATTG
ACTGGAGGGACATCGTGAGGGACCGAGATGCTGAGATAGTGGTGAAGGACAAT
GGCAGAAGCTGTCCCCCCTGTCATGAGGTTTGCAAGGGGCGATGCTGGGGTCCT
GGATCAGAAGACTGCCAGACATTGACCAAGACCATCTGTGCTCCTCAGTGTAAT
GGTCACTGCTTTGGGCCCAACCCCAACCAGTGCTGCCATGATGAGTGTGCCGGG
GGCTGCTCAGGCCCTCAGGACACAGACTGCTTTGCCTGCCGGCACTTCAATGAC
AGTGGAGCCTGTGTACCTCGCTGTCCACAGCCTCTTGTCTACAACAAGCTAACTT
TCCAGCTGGAACCCAATCCCCACACCAAGTATCAGTATGGAGGAGTTTGTGTAG
CCAGCTGTCCCCATAACTTTGTGGTGGATCAAACATCCTGTGTCAGGGCCTGTCC
TCCTGACAAGATGGAAGTAGATAAAAATGGGCTCAAGATGTGTGAGCCTTGTGG
GGGACTATGTCCCAAAGCCTGTGAGGGAACAGGCTCTGGGAGCCGCTTCCAGA
CTGTGGACTCGAGCAACATTGATGGATTTGTGAACTGCACCAAGATCCTGGGCA
ACCTGGACTTTCTGATCACCGGCCTCAATGGAGACCCCTGGCACAAGATCCCTG
CCCTGGACCCAGAGAAGCTCAATGTCTTCCGGACAGTACGGGAGATCACAGGTT
ACCTGAACATCCAGTCCTGGCCGCCCCACATGCACAACTTCAGTGTTTTTTCCAA
TTTGACAACCATTGGAGGCAGAAGCCTCTACAACCGGGGCTTCTCATTGTTGATC
ATGAAGAACTTGAATGTCACATCTCTGGGCTTCCGATCCCTGAAGGAAATTAGTG

CTGGGCGTATCTATATAAGTGCCAATAGGCAGCTCTGCTACCACCACTCTTTGAA
CTGGACCAAGGTGCTTCGGGGGCCTACGGAAGAGCGACTAGACATCAAGCATA
ATCGGCCGCGCAGAGACTGCGTGGCAGAGGGCAAAGTGTGTGACCCACTGTGC
TCCTCTGGGGGATGCTGGGGCCCAGGCCCTGGTCAGTGCTTGTCCTGTCGAAATT
ATAGCCGAGGAGGTGTCTGTGTGACCCACTGCAACTTTCTGAATGGGGAGCCTC
GAGAATTTGCCCATGAGGCCGAATGCTTCTCCTGCCACCCGGAATGCCAACCCA
TGGAGGGCACTGCCACATGCAATGGCTCGGGCTCTGATACTTGTGCTCAATGTG
CCCATTTTCGAGATGGGCCCCACTGTGTGAGCAGCTGCCCCCATGGAGTCCTAG
GTGCCAAGGGCCCAATCTACAAGTACCCAGATGTTCAGAATGAATGTCGGCCCT
GCCATGAGAACTGCACCCAGGGGTGTAAAGGACCAGAGCTTCAAGACTGTTTAG
GACAAACACTGGTGCTGATCGGCAAAACCCATCTGACAATGGCTTTGACAGTGA
TAGCAGGATTGGTAGTGATTTTCATGATGCTGGGCGGCACTTTTCTCTACTGGCGT
GGGCGCCGGATTCAGAATAAAAGGGCTATGAGGCGATACTTGGAACGGGGTGA
GAGCATAGAGCCTCTGGACCCCAGTGAGAAGGCTAACAAAGTCTTGGCCAGAA
TCTTCAAAGAGACAGAGCTAAGGAAGCTTAAAGTGCTTGGCTCGGGTGTCTTTGG
AACTGTGCACAAAGGAGTGTGGATCCCTGAGGGTGAATCAATCAAGATTCCAGT
CTGCATTAAAGTCATTGAGGACAAGAGTGGACGGCAGAGTTTTCAAGCTGTGAC
AGATCATATGCTGGCCATTGGCAGCCTGGACCATGCCCACATTGTAAGGCTGCT
GGGACTATGCCCAGGGTCATCTCTGCAGCTTGTCACTCAATATTTGCCTCTGGGTT
CTCTGCTGGATCATGTGAGACAACACCGGGGGGCACTGGGGCCACAGCTGCTG
CTCAACTGGGGAGTACAAATTGCCAAGGGAATGTACTACCTTGAGGAACATGGT
ATGGTGCATAGAAACCTGGCTGCCCGAAACGTGCTACTCAAGTCACCCAGTCAG
GTTCAGGTGGCAGATTTTGGTGTGGCTGACCTGCTGCCTCCTGATGATAAGCAGC
TGCTATACAGTGAGGCCAAGACTCCAATTAAGTGGATGGCCCTTGAGAGTATCC
ACTTTGGGAAATACACACACCAGAGTGATGTCTGGAGCTATGGTGTGACAGTTTG
GGAGTTGATGACCTTCGGGGCAGAGCCCTATGCAGGGCTACGATTGGCTGAAGT
ACCAGACCTGCTAGAGAAGGGGGAGCGGTTGGCACAGCCCCAGATCTGCACAA
TTGATGTCTACATGGTGATGGTCAAGTGTTGGATGATTGATGAGAACATTCGCCC
AACCTTTAAAGAACTAGCCAATGAGTTCACCAGGATGGCCCGAGACCCACCAC
GGTATCTGGTCATAAAGAGAGAGAGTGGGCCTGGAATAGCCCCTGGGCCAGAG
CCCCATGGTCTGACAAACAAGAAGCTAGAGGAAGTAGAGCTGGAGCCAGAACT

AGACCTAGACCTAGACTTGGAAGCAGAGGAGGACAACCTGGCAACCACCACAC
TGGGCTCCGCCCTCAGCCTACCAGTTGGAACACTTAATCGGCCACGTGGGAGCC
AGAGCCTTTTAAGTCCATCATCTGGATACATGCCCATGAACCAGGGTAATCTTGG
GGAGTCTTGCCAGGAGTCTGCAGTTTCTGGGAGCAGTGAACGGTGCCCCCGTCC
AGTCTCTCTACACCCAATGCCACGGGGATGCCTGGCATCAGAGTCATCAGAGGG
GCATGTAACAGGCTCTGAGGCTGAGCTCCAGGAGAAAGTGTCAATGTGTAGGAG
CCGGAGCAGGAGCCGGAGCCCACGGCCACGCGGAGATAGCGCCTACCATTCC
CAGCGCCACAGTCTGCTGACTCCTGTTACCCCACTCTCCCCACCCGGGTTAGAG
GAAGAGGATGTCAACGGTTATGTCATGCCAGATACACACCTCAAAGGTACTCCC
TCCTCCCGGGAAGGCACCCTTTCTTCAGTGGGTCTCAGTTCTGTCCTGGGTACTG
AAGAAGAAGATGAAGATGAGGAGTATGAATACATGAACCGGAGGAGAAGGCA
CAGTCCACCTCATCCCCCTAGGCCAAGTTCCCTTGAGGAGCTGGGTTATGAGTAC
ATGGATGTGGGGTCAGACCTCAGTGCCTCTCTGGGCAGCACACAGAGTTGCCCA
CTCCACCCTGTACCCATCATGCCCACTGCAGGCACAACTCCAGATGAAGACTAT
GAATATATGAATCGGCAACGAGATGGAGGTGGTCCTGGGGGTGATTATGCAGCC
ATGGGGGCCTGCCCAGCATCTGAGCAAGGGTATGAAGAGATGAGAGCTTTTCAG
GGGCCTGGACATCAGGCCCCCCATGTCCATTATGCCCGCCTAAAAACTCTACGT
AGCTTAGAGGCTACAGACTCTGCCTTTGATAACCCTGATTACTGGCATAGCAGGC
TTTTCCCCAAGGCTAATGCCCAGAGAACGTAACTCCTGCTCCCTGTGGCACTCAG
GGAGCATTTAATGGCAGCTAGTGCCTTTAGAGGGTACCGTCTTCTCCCTATTCCCT
CTCTCTCCCAGGTCCCAGCCCCTTTTCCCCAGTCCCAGACAATTCCATTCAATCTT
TGGAGGCTTTTAAACATTTTGACACAAAATTCTTATGGTATGTAGCCAGCTGTGCA
CTTTCTTCTCTTTCCCAACCCCAGGAAAGGTTTTCCTTATTTTGTGTGCTTTCCCAG
TCCCATTCCTCAGCTTCTTCACAGGCACTCCTGGAGATATGAAGGATTACTCTCC
ATATCCCTTCCTCTCAGGCTCTTGACTACTTGGAACTAGGCTCTTATGTGTGCCTTT
GTTTCCCATCAGACTGTCAAGAAGAGGAAAGGGAGGAAACCTAGCAGAGGAAA
GTGTAATTTTGGTTTATGACTCTTAACCCCCTAGAAAGACAGAAGCTTAAAATCT
GTGAAGAAAGAGGTTAGGAGTAGATATTGATTACTATCATAATTCAGCACTTAAC
TATGAGCCAGGCATCATACTAAACTTCACCTACATTATCTCACTTAGTCCTTTATC
ATCCTTAAAACAATTCTGTGACATACATATTATCTCATTTTACACAAAGGGAAGTC
GGGCATGGTGGCTCATGCCTGTAATCTCAGCACTTTGGGAGGCTGAGGCAGAAG

GATTACCTGAGGCAAGGAGTTTGAGACCAGCTTAGCCAACATAGTAAGACCCCC
ATCTCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAACTTTAGAACTGGGTGCAG
TGGCTCATGCCTGTAATCCCAGCCAGCACTTTGGGAGGCTGAGATGGGAAGATC
ACTTGAGCCCAGAATTAGAGATAAGCCTATGGAAACATAGCAAGACACTGTCTC
TACAGGGGAAAAAAAAAAAAGAAACTGAGCCTTAAAGAGATGAAATAAATTAA
GCAGTAGATCCAGGATGCAAAATCCTCCCAATTCCTGTGCATGTGCTCTTATTGT
AAGGTGCCAAGAAAAACTGATTTAAGTTACAGCCCTTGTTTAAGGGGCACTGTTT
CTTGTTTTTGCACTGAATCAAGTCTAACCCCAACAGCCACATCCTCCTATACCTA
GACATCTCATCTCAGGAAGTGGTGGTGGGGGTAGTCAGAAGGAAAAATAACTGG
ACATCTTTGTGTAAACCATAATCCACATGTGCCGTAAATGATCTTCACTCCTTATC
CGAGGGCAAATTCACAAGGATCCCCAAGATCCACTTTTAGAAGCCATTCTCATC
CAGCAGTGAGAAGCTTCCAGGTAGGACAGAAAAAAGATCCAGCTTCAGCTGCA
CACCTCTGTCCCCTTGGATGGGGAACTAAGGGAAAACGTCTGTTGTATCACTGAA
GTTTTTTGTTTTGTTTTTATACGTGTCTGAATAAAAATGCCAAAGTTTTTTTTCAGCA
AAAAAAA

SEQ ID NO: 14 is a Her4 amino acid sequence (Genbank Accession No. AAI43750).

MKPATGLWVWVSLLVAAGTVQPSDSQSVCAGTENKLSSLSDLEQQYRALRKYYEN
CEVVMGNLEITSIEHNRDLSFLRSVREVTGYVLVALNQFRYLPLENLRIIRGTKLYEDR
YALAIFLNYRKDGNFGLQELGLKNLTEILNGGVYVDQNKFLCYADTIHWQDIVRNP
WPSNLTLVSTNGSSGCGRCHKSCTGRCWGPTENHCQTLTRTVCAEQCDGRCYGPY
VSDCCHRECAGGCSGPKDTDCFACMNFNDSGACVTQCPQTFVYNPTTFQLEHNFN
AKYTYGAFCVKKCPHNFVVDSSSCVRACPSSKMEVEENGIKMCKPCTDICPKACDG
IGTGSLMSAQTVDSSNIDKFINCTKINGNLIFLVTGIHGDPYNAIEAIDPEKLNVFRTVR
EITGFLNIQSWPPNMTDFSVFSNLVTIGGRVLYSGLSLLILKQQGITSLQFQSLKEISAG
NIYITDNSNLCYYHTINWTTLFSTINQRIVIRDNRKAENCTAEGMVCNHLCSSDGCW
GPGPDQCLSCRRFSRGRICIESCNLYDGEFREFENGSICVECDPQCEKMEDGLLTCH
GPGPDNCTKCSHFKDGPNCVEKCPDGLQGANSFIFKYADPDRECHPCHPNCTQGCI
GSSIEDCIGLMDRTPLIAAGVIGGLFILVIVGLTFAVYVRRKSIKKKRALRRFLETELVEP
LTPSGTAPNQAQLRILKETELKRVKVLGSGAFGTVYKGIWVPEGETVKIPVAIKILNETT

GPKANVEFMDEALIMASMDHPHLVRLLGVCLSPTIQLVTQLMPHGCLLEYVHEHKD
NIGSQLLLNWCVQIAKGMMYLEERRLVHRDLAARNVLVKSPNHVKITDFGLARLLE
GDEKEYNADGGKMPIKWMALECIHYRKFTHQSDVWSYGVTIWELMTFGGKPYDGI
PTREIPDLLEKGERLPQPPICTIDVYMVMVKCWMIDADSRPKFKELAAEFSRMARDP
QRYLVIQGDDRMKLPSPNDSKFFQNLLDEEDLEDMMDAEEYLVPQAFNIPPPIYTS
RARIDSNRNQFVYRDGGFAAEQGVSVPYRAPTSTIPEAPVAQGATAEIFDDSCCNGT
LRKPVAPHVQEDSSTQRYSADPTVFAPERSPRGELDEEGYMTPMRDKPKQEYLNPV
EENPFVSRRKNGDLQALDNPEYHNASNGPPKAEDEYVNEPLYLNTFANTLGKAEYL
KNNILSMPEKAKKAFDNPDYWNHSLPPRSTLQHPDYLQEYSTKYFYKQNGRIRPIV
AENPEYLSEFSLKPGTVLPPPPYRHRNTVV

SEQ ID NO: 15 is an exemplary cDNA sequence encoding Her4 (GenBank Accession No. BC143749.1).

CAGCACGGGATCTGAGACTTCCAAAAAATGAAGCCGGCGACAGGACTTTGGGT
CTGGGTGAGCCTTCTCGTGGCGGCGGGGACCGTCCAGCCCAGCGATTCTCAGTC
AGTGTGTGCAGGAACGGAGAATAAACTGAGCTCTCTCTCTGACCTGGAACAGCA
GTACCGAGCCTTGCGCAAGTACTATGAAAACTGTGAGGTTGTCATGGGCAACCT
GGAGATAACCAGCATTGAGCACAACCGGGACCTCTCCTTCCTGCGGTCTGTTCG
AGAAGTCACAGGCTACGTGTTAGTGGCTCTTAATCAGTTTCGTTACCTGCCTCTG
GAGAATTTACGCATTATTCGTGGGACAAAACTTTATGAGGATCGATATGCCTTGG
CAATATTTTTAAACTACAGAAAAGATGGAAACTTTGGACTTCAAGAACTTGGATT
AAAGAACTTGACAGAAATCCTAAATGGTGGAGTCTATGTAGACCAGAACAAATT
CCTTTGTTATGCAGACACCATTCATTGGCAAGATATTGTTCGGAACCCATGGCCTT
CCAACTTGACTCTTGTGTCAACAAATGGTAGTTCAGGATGTGGACGTTGCCATAA
GTCCTGTACTGGCCGTTGCTGGGGACCCACAGAAAATCATTGCCAGACTTTGAC
AAGGACGGTGTGTGCAGAACAATGTGACGGCAGATGCTACGGACCTTACGTCA
GTGACTGCTGCCATCGAGAATGTGCTGGAGGCTGCTCAGGACCTAAGGACACAG
ACTGCTTTGCCTGCATGAATTTCAATGACAGTGGAGCATGTGTTACTCAGTGTCC
CCAAACCTTTGTCTACAATCCAACCACCTTTCAACTGGAGCACAATTTCAATGCA
AAGTACACATATGGAGCATTCTGTGTCAAGAAATGTCCACATAACTTTGTGGTAG
ATTCCAGTTCTTGTGTGCGTGCCTGCCCTAGTTCCAAGATGGAAGTAGAAGAAAA

TGGGATTAAAATGTGTAAACCTTGCACTGACATTTGCCCAAAAGCTTGTGATGGC
ATTGGCACAGGATCATTGATGTCAGCTCAGACTGTGGATTCCAGTAACATTGACA
AATTCATAAACTGTACCAAGATCAATGGGAATTTGATCTTTCTAGTCACTGGTATT
CATGGGGACCCTTACAATGCAATTGAAGCCATAGACCCAGAGAAACTGAACGTC
TTTCGGACAGTCAGAGAGATAACAGGTTTCCTGAACATACAGTCATGGCCACCA
AACATGACTGACTTCAGTGTTTTTTCTAACCTGGTGACCATTGGTGGAAGAGTACT
CTATAGTGGCCTGTCCTTGCTTATCCTCAAGCAACAGGGCATCACCTCTCTACAG
TTCCAGTCCCTGAAGGAAATCAGCGCAGGAAACATCTATATTACTGACAACAGC
AACCTGTGTTATTATCATACCATTAACTGGACAACACTCTTCAGCACAATCAACC
AGAGAATAGTAATCCGGGACAACAGAAAAGCTGAAAATTGTACTGCTGAAGGA
ATGGTGTGCAACCATCTGTGTTCCAGTGATGGCTGTTGGGGACCTGGGCCAGAC
CAATGTCTGTCGTGTCGCCGCTTCAGTAGAGGAAGGATCTGCATAGAGTCTTGTA
ACCTCTATGATGGTGAATTTCGGGAGTTTGAGAATGGCTCCATCTGTGTGGAGTG
TGACCCCCAGTGTGAGAAGATGGAAGATGGCCTCCTCACATGCCATGGACCGG
GTCCTGACAACTGTACAAAGTGCTCTCATTTTAAAGATGGCCCAAACTGTGTGGA
AAAATGTCCAGATGGCTTACAGGGGGCAAACAGTTTCATTTTCAAGTATGCTGAT
CCAGATCGGGAGTGCCACCCATGCCATCCAAACTGCACCCAAGGGTGCATAGG
CTCAAGTATTGAAGACTGCATCGGCCTGATGGATAGAACTCCCCTGATTGCAGCT
GGAGTAATTGGTGGGCTCTTCATTCTGGTCATTGTGGGTCTGACATTTGCTGTTTAT
GTTAGAAGGAAGAGCATCAAAAAGAAAAGAGCCTTGAGAAGATTCTTGGAAAC
AGAGTTGGTGGAACCATTAACTCCCAGTGGCACAGCACCCAATCAAGCTCAACT
TCGTATTTTGAAAGAAACTGAGCTGAAGAGGGTAAAAGTCCTTGGCTCAGGTGCT
TTTGGAACGGTTTATAAAGGTATTTGGGTACCTGAAGGAGAAACTGTGAAGATTC
CTGTGGCTATTAAGATTCTTAATGAGACAACTGGTCCCAAGGCAAATGTGGAGTT
CATGGATGAAGCTCTGATCATGGCAAGTATGGATCATCCACACCTAGTCCGGTTG
CTGGGTGTGTGTCTGAGCCCAACCATCCAGCTGGTTACTCAACTTATGCCCCATG
GCTGCCTGTTGGAGTATGTCCACGAGCACAAGGATAACATTGGATCACAACTGC
TGCTTAACTGGTGTGTCCAGATAGCTAAGGGAATGATGTACCTGGAAGAAAGAC
GACTCGTTCATCGGGATTTGGCAGCCCGTAATGTCTTAGTGAAATCTCCAAACCA
TGTGAAAATCACAGATTTTGGGCTAGCCAGACTCTTGGAAGGAGATGAAAAAGA
GTACAATGCTGATGGAGGAAAGATGCCAATTAAATGGATGGCTCTGGAGTGTAT

ACATTACAGGAAATTCACCCATCAGAGTGACGTTTGGAGCTATGGAGTTACTATA
TGGGAACTGATGACCTTTGGAGGAAAACCCTATGATGGAATTCCAACGCGAGAA
ATCCCTGATTTATTAGAGAAAGGAGAACGTTTGCCTCAGCCTCCCATCTGCACTA
TTGACGTTTACATGGTCATGGTCAAATGTTGGATGATTGATGCTGACAGTAGACC
TAAATTTAAGGAACTGGCTGCTGAGTTTTCAAGGATGGCTCGAGACCCTCAAAG
ATACCTAGTTATTCAGGGTGATGATCGTATGAAGCTTCCCAGTCCAAATGACAGC
AAGTTCTTTCAGAATCTCTTGGATGAAGAGGATTTGGAAGATATGATGGATGCTG
AGGAGTACTTGGTCCCTCAGGCTTTCAACATCCCACCTCCCATCTATACTTCCAG
AGCAAGAATTGACTCGAATAGGAACCAGTTTGTGTACCGAGATGGAGGTTTTGCT
GCTGAACAAGGAGTGTCTGTGCCCTACAGAGCCCCAACTAGCACAATTCCAGAA
GCTCCTGTGGCACAGGGTGCTACTGCTGAGATTTTTGATGACTCCTGCTGTAATG
GCACCCTACGCAAGCCAGTGGCACCCCATGTCCAAGAGGACAGTAGCACCCAG
AGGTACAGTGCTGACCCCACCGTGTTTGCCCCAGAACGGAGCCCACGAGGAGA
GCTGGATGAGGAAGGTTACATGACTCCTATGCGAGACAAACCCAAACAAGAAT
ACCTGAATCCAGTGGAGGAGAACCCTTTTGTTTCTCGGAGAAAAAATGGAGACC
TTCAAGCATTGGATAATCCCGAATATCACAATGCATCCAATGGTCCACCCAAGG
CCGAGGATGAGTATGTGAATGAGCCACTGTACCTCAACACCTTTGCCAACACCTT
GGGAAAAGCTGAGTACCTGAAGAACAACATACTGTCAATGCCAGAGAAGGCCA
AGAAAGCGTTTGACAACCCTGACTACTGGAACCACAGCCTGCCACCTCGGAGC
ACCCTTCAGCACCCAGACTACCTGCAGGAGTACAGCACAAAATATTTTTATAAA
CAGAATGGGCGGATCCGGCCTATTGTGGCAGAGAATCCTGAATACCTCTCTGAG
TTCTCCCTGAAGCCAGGCACTGTGCTGCCGCCTCCACCTTACAGACACCGGAAT
ACTGTGGTGTAAGCTCAGTTGTGGTTTTTTAGGTGGAGAGACACACCTGCTCCAA
TTTCCCCACCCC

SEQ ID NO: 16 is an Estrogen Receptor amino acid sequence (NCBI Reference Sequence No. NP_000116.2).

MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGEVYLDSSKPAVYNYPE
GAAYEFNAAAAANAQVYGQTGLPYGPGSEAAAFGSNGLGGFPPLNSVSPSPLMLL
HPPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQGGRERLAS
TNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHNDYM

CPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGIRKDRRGGRMLKHKRQRDDG
EGRGEVGSAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLDAEPPILYSE
YDPTRPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEI
LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQ
GEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQH
QRLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRG
GASVEETDQSHLATAGSTSSHSLQKYYITGEAEGFPATV

SEQ ID NO: 17 is an exemplary cDNA sequence encoding an Estrogen Receptor (NCBI Reference Sequence No. NM_000125.3).

AGGAGCTGGCGGAGGGCGTTCGTCCTGGGACTGCACTTGCTCCCGTCGGGTCGC
CCGGCTTCACCGGACCCGCAGGCTCCCGGGGCAGGGCCGGGGCCAGAGCTCG
CGTGTCGGCGGGACATGCGCTGCGTCGCCTCTAACCTCGGGCTGTGCTCTTTTTC
CAGGTGGCCCGCCGGTTTCTGAGCCTTCTGCCCTGCGGGGACACGGTCTGCACC
CTGCCCGCGGCCACGGACCATGACCATGACCCTCCACACCAAAGCATCTGGGA
TGGCCCTACTGCATCAGATCCAAGGGAACGAGCTGGAGCCCCTGAACCGTCCG
CAGCTCAAGATCCCCCTGGAGCGGCCCCTGGGCGAGGTGTACCTGGACAGCAG
CAAGCCCGCCGTGTACAACTACCCCGAGGGCGCCGCCTACGAGTTCAACGCCG
CGGCCGCCGCCAACGCGCAGGTCTACGGTCAGACCGGCCTCCCCTACGGCCCC
GGGTCTGAGGCTGCGGCGTTCGGCTCCAACGGCCTGGGGGGTTTCCCCCCACTC
AACAGCGTGTCTCCGAGCCCGCTGATGCTACTGCACCCGCCGCCGCAGCTGTCG
CCTTTCCTGCAGCCCCACGGCCAGCAGGTGCCCTACTACCTGGAGAACGAGCCC
AGCGGCTACACGGTGCGCGAGGCCGGCCCGCCGGCATTCTACAGGCCAAATTC
AGATAATCGACGCCAGGGTGGCAGAGAAAGATTGGCCAGTACCAATGACAAGG
GAAGTATGGCTATGGAATCTGCCAAGGAGACTCGCTACTGTGCAGTGTGCAATG
ACTATGCTTCAGGCTACCATTATGGAGTCTGGTCCTGTGAGGGCTGCAAGGCCTT
CTTCAAGAGAAGTATTCAAGGACATAACGACTATATGTGTCCAGCCACCAACCA
GTGCACCATTGATAAAAACAGGAGGAAGAGCTGCCAGGCCTGCCGGCTCCGCA
AATGCTACGAAGTGGGAATGATGAAAGGTGGGATACGAAAAGACCGAAGAGGA
GGGAGAATGTTGAAACACAAGCGCCAGAGAGATGATGGGGAGGGCAGGGGTG
AAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAAGCCCGCTCA

TGATCAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCCTGACGGCCGACCAGA
TGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCATACTCTATTCCGAGTATGATCCT
ACCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGAC
AGGGAGCTGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAGGCTTTGTGGAT
TTGACCCTCCATGATCAGGTCCACCTTCTAGAATGTGCCTGGCTAGAGATCCTGA
TGATTGGTCTCGTCTGGCGCTCCATGGAGCACCCAGGGAAGCTACTGTTTGCTCC
TAACTTGCTCTTGGACAGGAACCAGGGAAAATGTGTAGAGGGCATGGTGGAGAT
CTTCGACATGCTGCTGGCTACATCATCTCGGTTCCGCATGATGAATCTGCAGGGA
GAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAATTCTGGAGTGTACACATTT
CTGTCCAGCACCCTGAAGTCTCTGGAAGAGAAGGACCATATCCACCGAGTCCTG
GACAAGATCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTGACCCTG
CAGCAGCAGCACCAGCGGCTGGCCCAGCTCCTCCTCATCCTCTCCCACATCAGG
CACATGAGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGT
GGTGCCCCTCTATGACCTGCTGCTGGAGATGCTGGACGCCCACCGCCTACATGC
GCCCACTAGCCGTGGAGGGGCATCCGTGGAGGAGACGGACCAAAGCCACTTGG
CCACTGCGGGCTCTACTTCATCGCATTCCTTGCAAAAGTATTACATCACGGGGGA
GGCAGAGGGTTTCCCTGCCACGGTCTGAGAGCTCCCTGGCTCCCACACGGTTCA
GATAATCCCTGCTGCATTTTACCCTCATCATGCACCACTTTAGCCAAATTCTGTCT
CCTGCATACACTCCGGCATGCATCCAACACCAATGGCTTTCTAGATGAGTGGCC
ATTCATTTGCTTGCTCAGTTCTTAGTGGCACATCTTCTGTCTTCTGTTGGGAACAGC
CAAAGGGATTCCAAGGCTAAATCTTTGTAACAGCTCTCTTTCCCCCTTGCTATGTT
ACTAAGCGTGAGGATTCCCGTAGCTCTTCACAGCTGAACTCAGTCTATGGGTTGG
GGCTCAGATAACTCTGTGCATTTAAGCTACTTGTAGAGACCCAGGCCTGGAGAG
TAGACATTTTGCCTCTGATAAGCACTTTTTAAATGGCTCTAAGAATAAGCCACAG
CAAAGAATTTAAAGTGGCTCCTTTAATTGGTGACTTGGAGAAAGCTAGGTCAAGG
GTTTATTATAGCACCCTCTTGTATTCCTATGGCAATGCATCCTTTTATGAAAGTGGT
ACACCTTAAAGCTTTTATATGACTGTAGCAGAGTATCTGGTGATTGTCAATTCATT
CCCCCTATAGGAATACAAGGGGCACACAGGGAAGGCAGATCCCCTAGTTGGCA
AGACTATTTTAACTTGATACACTGCAGATTCAGATGTGCTGAAAGCTCTGCCTCTG
GCTTTCCGGTCATGGGTTCCAGTTAATTCATGCCTCCCATGGACCTATGGAGAGC
AGCAAGTTGATCTTAGTTAAGTCTCCCTATATGAGGGATAAGTTCCTGATTTTTGT

TTTTATTTTTGTGTTACAAAAGAAAGCCCTCCCTCCCTGAACTTGCAGTAAGGTCA
GCTTCAGGACCTGTTCCAGTGGGCACTGTACTTGGATCTTCCCGGCGTGTGTGTG
CCTTACACAGGGGTGAACTGTTCACTGTGGTGATGCATGATGAGGGTAAATGGT
AGTTGAAAGGAGCAGGGGCCCTGGTGTTGCATTTAGCCCTGGGGCATGGAGCTG
AACAGTACTTGTGCAGGATTGTTGTGGCTACTAGAGAACAAGAGGGAAAGTAGG
GCAGAAACTGGATACAGTTCTGAGGCACAGCCAGACTTGCTCAGGGTGGCCCTG
CCACAGGCTGCAGCTACCTAGGAACATTCCTTGCAGACCCCGCATTGCCCTTTG
GGGGTGCCCTGGGATCCCTGGGGTAGTCCAGCTCTTCTTCATTTCCCAGCGTGGC
CCTGGTTGGAAGAAGCAGCTGTCACAGCTGCTGTAGACAGCTGTGTTCCTACAAT
TGGCCCAGCACCCTGGGGCACGGGAGAAGGGTGGGGACCGTTGCTGTCACTAC
TCAGGCTGACTGGGGCCTGGTCAGATTACGTATGCCCTTGGTGGTTTAGAGATAA
TCCAAAATCAGGGTTTGGTTTGGGGAAGAAAATCCTCCCCCTTCCTCCCCCGCCC
CGTTCCCTACCGCCTCCACTCCTGCCAGCTCATTTCCTTCAATTTCCTTTGACCTAT
AGGCTAAAAAAGAAAGGCTCATTCCAGCCACAGGGCAGCCTTCCCTGGGCCTTT
GCTTCTCTAGCACAATTATGGGTTACTTCCTTTTTCTTAACAAAAAAGAATGTTTG
ATTTCCTCTGGGTGACCTTATTGTCTGTAATTGAAACCCTATTGAGAGGTGATGTC
TGTGTTAGCCAATGACCCAGGTGAGCTGCTCGGGCTTCTCTTGGTATGTCTTGTTT
GGAAAAGTGGATTTCATTCATTTCTGATTGTCCAGTTAAGTGATCACCAAAGGAC
TGAGAATCTGGGAGGGCAAAAAAAAAAAAAAAGTTTTTATGTGCACTTAAATTTG
GGGACAATTTTATGTATCTGTGTTAAGGATATGTTTAAGAACATAATTCTTTTGTTG
CTGTTTGTTTAAGAAGCACCTTAGTTTGTTTAAGAAGCACCTTATATAGTATAATAT
ATATTTTTTTGAAATTACATTGCTTGTTTATCAGACAATTGAATGTAGTAATTCTGTT
CTGGATTTAATTTGACTGGGTTAACATGCAAAAACCAAGGAAAAATATTTAGTTTT
TTTTTTTTTTTTGTATACTTTTCAAGCTACCTTGTCATGTATACAGTCATTTATGCCT
AAAGCCTGGTGATTATTCATTTAAATGAAGATCACATTTCATATCAACTTTTGTAT
CCACAGTAGACAAAATAGCACTAATCCAGATGCCTATTGTTGGATACTGAATGA
CAGACAATCTTATGTAGCAAAGATTATGCCTGAAAAGGAAAATTATTCAGGGCA
GCTAATTTTGCTTTTACCAAAATATCAGTAGTAATATTTTTGGACAGTAGCTAATG
GGTCAGTGGGTTCTTTTTAATGTTTATACTTAGATTTTCTTTTAAAAAAATTAAAAT
AAAACAAAAAAAAAATTTCTAGGACTAGACGATGTAATACCAGCTAAAGCCAAAC
AATTATACAGTGGAAGGTTTTACATTATTCATCCAATGTGTTTCTATTCATGTTAAG

ATACTACTACATTTGAAGTGGGCAGAGAACATCAGATGATTGAAATGTTCGCCCA
GGGGTCTCCAGCAACTTTGGAAATCTCTTTGTATTTTTACTTGAAGTGCCACTAAT
GGACAGCAGATATTTTCTGGCTGATGTTGGTATTGGGTGTAGGAACATGATTTAA
AAAAAAACTCTTGCCTCTGCTTTCCCCCACTCTGAGGCAAGTTAAAATGTAAAAG
ATGTGATTTATCTGGGGGGCTCAGGTATGGTGGGGAAGTGGATTCAGGAATCTG
GGGAATGGCAAATATATTAAGAAGAGTATTGAAAGTATTTGGAGGAAAATGGTT
AATTCTGGGTGTGCACCAGGGTTCAGTAGAGTCCACTTCTGCCCTGGAGACCAC
AAATCAACTAGCTCCATTTACAGCCATTTCTAAAATGGCAGCTTCAGTTCTAGAG
AAGAAAGAACAACATCAGCAGTAAAGTCCATGGAATAGCTAGTGGTCTGTGTTT
CTTTTCGCCATTGCCTAGCTTGCCGTAATGATTCTATAATGCCATCATGCAGCAAT
TATGAGAGGCTAGGTCATCCAAAGAGAAGACCCTATCAATGTAGGTTGCAAAAT
CTAACCCCTAAGGAAGTGCAGTCTTTGATTTGATTTCCCTAGTAACCTTGCAGATA
TGTTTAACCAAGCCATAGCCCATGCCTTTTGAGGGCTGAACAAATAAGGGACTTA
CTGATAATTTACTTTTGATCACATTAAGGTGTTCTCACCTTGAAATCTTATACACTG
AAATGGCCATTGATTTAGGCCACTGGCTTAGAGTACTCCTTCCCCTGCATGACAC
TGATTACAAATACTTTCCTATTCATACTTTCCAATTATGAGATGGACTGTGGGTAC
TGGGAGTGATCACTAACACCATAGTAATGTCTAATATTCACAGGCAGATCTGCTT
GGGGAAGCTAGTTATGTGAAAGGCAAATAGAGTCATACAGTAGCTCAAAAGGC
AACCATAATTCTCTTTGGTGCAGGTCTTGGGAGCGTGATCTAGATTACACTGCAC
CATTCCCAAGTTAATCCCCTGAAAACTTACTCTCAACTGGAGCAAATGAACTTTG
GTCCCAAATATCCATCTTTTCAGTAGCGTTAATTATGCTCTGTTTCCAACTGCATTT
CCTTTCCAATTGAATTAAAGTGTGGCCTCGTTTTTAGTCATTTAAAATTGTTTTCTA
AGTAATTGCTGCCTCTATTATGGCACTTCAATTTTGCACTGTCTTTTGAGATTCAAG
AAAAATTTCTATTCTTTTTTTTGCATCCAATTGTGCCTGAACTTTTAAAATATGTAA
ATGCTGCCATGTTCCAAACCCATCGTCAGTGTGTGTGTTTAGAGCTGTGCACCCT
AGAAACAACATATTGTCCCATGAGCAGGTGCCTGAGACACAGACCCCTTTGCAT
TCACAGAGAGGTCATTGGTTATAGAGACTTGAATTAATAAGTGACATTATGCCAG
TTTCTGTTCTCTCACAGGTGATAAACAATGCTTTTTGTGCACTACATACTCTTCAGT
GTAGAGCTCTTGTTTTATGGGAAAAGGCTCAAATGCCAAATTGTGTTTGATGGATT
AATATGCCCTTTTGCCGATGCATACTATTACTGATGTGACTCGGTTTTGTCGCAGC
TTTGCTTTGTTTAATGAAACACACTTGTAAACCTCTTTTGCACTTTGAAAAAGAAT

CCAGCGGGATGCTCGAGCACCTGTAAACAATTTTCTCAACCTATTTGATGTTCAA
ATAAAGAATTAAACTAAA

SEQ ID NO: 18 is a Progesterone Receptor amino acid sequence (Genbank Accession No. AAD01587.1).

MTELKAKGPRAPHVAGGPPSPEVGSPLLCRPAAGPFPGSQTSDTLPEVSAIPISLDGL
LFPRPCQGQDPSDEKTQDQQSLSDVEGAYSRAEATRGAGGSSSSPPEKDSGLLDSV
LDTLLAPSGPGQSQPSPPACEVTSSWCLFGPELPEDPPAAPATQRVLSPLMSRSGCK
VGDSSGTAAAHKVLPRGLSPARQLLLPASESPHWSGAPVKPSPQAAAVEVEEEDGS
ESEESAGPLLKGKPRALGGAAAGGGAAAVPPGAAAGGVALVPKEDSRFSAPRVALV
EQDAPMAPGRSPLATTVMDFIHVPILPLNHALLAARTRQLLEDESYDGGAGAASAF
APPRSSPCASSTPVAVGDFPDCAYPPDAEPKDDAYPLYSDFQPPALKIKEEEEGAEA
SARSPRSYLVAGANPAAFPDFPLGPPPPLPPRATPSRPGEAAVTAAPASASVSSASS
SGSTLECILYKAEGAPPQQGPFAPPPCKAPGASGCLLPRDGLPSTSASAAAAGAAPA
LYPALGLNGLPQLGYQAAVLKEGLPQVYPPYLNYLRPDSEASQSPQYSFESLPQKIC
LICGDEASGCHYGVLTCGSCKVFFKRAMEGQHNYLCAGRNDCIVDKIRRKNCPACR
LRKCCQAGMVLGGRKFKKFNKVRVVRALDAVALPQPLGVPNESQALSQRFTFSPG
QDIQLIPPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLP
GFRNLHIDDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESS
FYSLCLTMWQIPQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRE
LIKAIGLRQKGVVSSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEM
MSEVIAAQLPKILAGMVKPLLFHKK

SEQ ID NO: 19 is an exemplary cDNA sequence encoding a Progesterone Receptor (GenBank Accession No. AF016381.1).

CTGACCAGCGCCGCCCTCCCCCGCCCCCGACCCAGGAGGTGGAGATCCCTCCG
GTCCAGCCACATTCAACACCCACTTTCTCCTCCCTCTGCCCCTATATTCCCGAAA
CCCCCTCCTCCTTCCCTTTTCCCTCCTCCCTGGAGACGGGGGAGGAGAAAAGGG
GAGTCCAGTCGTCATGACTGAGCTGAAGGCAAAGGGTCCCCGGGCTCCCCACG
TGGCGGGCGGCCCGCCCTCCCCCGAGGTCGGATCCCCACTGCTGTGTCGCCCA
GCCGCAGGTCCGTTCCCGGGGAGCCAGACCTCGGACACCTTGCCTGAAGTTTCG

GCCATACCTATCTCCCTGGACGGGCTACTCTTCCCTCGGCCCTGCCAGGGACAG
GACCCCTCCGACGAAAAGACGCAGGACCAGCAGTCGCTGTCGGACGTGGAGG
GCGCATATTCCAGAGCTGAAGCTACAAGGGGTGCTGGAGGCAGCAGTTCTAGTC
CCCCAGAAAAGGACAGCGGACTGCTGGACAGTGTCTTGGACACTCTGTTGGCGC
CCTCAGGTCCCGGGCAGAGCCAACCCAGCCCTCCCGCCTGCGAGGTCACCAGC
TCTTGGTGCCTGTTTGGCCCCGAACTTCCCGAAGATCCACCGGCTGCCCCCGCC
ACCCAGCGGGTGTTGTCCCCGCTCATGAGCCGGTCCGGGTGCAAGGTTGGAGAC
AGCTCCGGGACGGCAGCTGCCCATAAAGTGCTGCCCCGGGGCCTGTCACCAGC
CCGGCAGCTGCTGCTCCCGGCCTCTGAGAGCCCTCACTGGTCCGGGGCCCCAGT
GAAGCCGTCTCCGCAGGCCGCTGCGGTGGAGGTTGAGGAGGAGGATGGCTCTG
AGTCCGAGGAGTCTGCGGGTCCGCTTCTGAAGGGCAAACCTCGGGCTCTGGGTG
GCGCGGCGGCTGGAGGAGGAGCCGCGGCTGTCCCGCCGGGGGCGGCAGCAGG
AGGCGTCGCCCTGGTCCCCAAGGAAGATTCCCGCTTCTCAGCGCCCAGGGTCGC
CCTGGTGGAGCAGGACGCGCCGATGGCGCCCGGGCGCTCCCCGCTGGCCACCA
CGGTGATGGATTTCATCCACGTGCCTATCCTGCCTCTCAATCACGCCTTATTGGC
AGCCCGCACTCGGCAGCTGCTGGAAGACGAAAGTTACGACGGCGGGGCCGGG
GCTGCCAGCGCCTTTGCCCCGCCGCGGAGTTCACCCTGTGCCTCGTCCACCCCG
GTCGCTGTAGGCGACTTCCCCGACTGCGCGTACCCGCCCGACGCCGAGCCCAA
GGACGACGCGTACCCTCTCTATAGCGACTTCCAGCCGCCCGCTCTAAAGATAAA
GGAGGAGGAGGAAGGCGCGGAGGCCTCCGCGCGCTCCCCGCGTTCCTACCTTG
TGGCCGGTGCCAACCCCGCAGCCTTCCCGGATTTCCCGTTGGGGCCACCGCCCC
CGCTGCCGCCGCGAGCGACCCCATCCAGACCCGGGGAAGCGGCGGTGACGGC
CGCACCCGCCAGTGCCTCAGTCTCGTCTGCGTCCTCCTCGGGGTCGACCCTGGA
GTGCATCCTGTACAAAGCGGAGGGCGCGCCGCCCCAGCAGGGCCCGTTCGCGC
CGCCGCCCTGCAAGGCGCCGGGCGCGAGCGGCTGCCTGCTCCCGCGGGACGG
CCTGCCCTCCACCTCCGCCTCTGCCGCCGCCGCCGGGGCGGCCCCCGCGCTCTA
CCCTGCACTCGGCCTCAACGGGCTCCCGCAGCTCGGCTACCAGGCCGCCGTGCT
CAAGGAGGGCCTGCCGCAGGTCTACCCGCCCTATCTCAACTACCTGAGGCCGG
ATTCAGAAGCCAGCCAGAGCCCACAATACAGCTTCGAGTCATTACCTCAGAAGA
TTTGTTTAATCTGTGGGGATGAAGCATCAGGCTGTCATTATGGTGTCCTTACCTGT
GGGAGCTGTAAGGTCTTCTTTAAGAGGGCAATGGAAGGGCAGCACAACTACTTA

TGTGCTGGAAGAAATGACTGCATCGTTGATAAAATCCGCAGAAAAAACTGCCCA
GCATGTCGCCTTAGAAAGTGCTGTCAGGCTGGCATGGTCCTTGGAGGTCGAAAA
TTTAAAAAGTTCAATAAAGTCAGAGTTGTGAGAGCACTGGATGCTGTTGCTCTCC
CACAGCCATTGGGCGTTCCAAATGAAAGCCAAGCCCTAAGCCAGAGATTCACTT
TTTCACCAGGTCAAGACATACAGTTGATTCCACCACTGATCAACCTGTTAATGAG
CATTGAACCAGATGTGATCTATGCAGGACATGACAACACAAAACCTGACACCTC
CAGTTCTTTGCTGACAAGTCTTAATCAACTAGGCGAGAGGCAACTTCTTTCAGTA
GTCAAGTGGTCTAAATCATTGCCAGGTTTTCGAAACTTACATATTGATGACCAGA
TAACTCTCATTCAGTATTCTTGGATGAGCTTAATGGTGTTTGGTCTAGGATGGAGA
TCCTACAAACATGTCAGTGGGCAGATGCTGTATTTTGCACCTGATCTAATACTAA
ATGAACAGCGGATGAAAGAATCATCATTCTATTCATTATGCCTTACCATGTGGCA
GATCCCACAGGAGTTTGTCAAGCTTCAAGTTAGCCAAGAAGAGTTCCTCTGTATG
AAAGTATTGTTACTTCTTAATACAATTCCTTTGGAAGGGCTACGAAGTCAAACCC
AGTTTGAGGAGATGAGGTCAAGCTACATTAGAGAGCTCATCAAGGCAATTGGTTT
GAGGCAAAAAGGAGTTGTGTCGAGCTCACAGCGTTTCTATCAACTTACAAAACTT
CTTGATAACTTGCATGATCTTGTCAAACAACTTCATCTGTACTGCTTGAATACATTT
ATCCAGTCCCGGGCACTGAGTGTTGAATTTCCAGAAATGATGTCTGAAGTTATTG
CTGCACAATTACCCAAGATATTGGCAGGGATGGTGAAACCCCTTCTCTTTCATAA
AAAGTGAATGTCATCTTTTTCTTTTAAAGAATTAAATTTTGTGG

SEQ ID NO: 20 is an exemplary oligo/polynucleotide probe sequence.
TCTAGCCATTTGATGCCAACAGTAGAAAGGG

SEQ ID NO: 21 is an exemplary unique sequence identifier sequence.
CTTGACTGAGCGACTGAGC

SEQ ID NO: 22 is another exemplary unique sequence identifier sequence.
CTCCAGGGTTAGGCAGATC

SEQ ID NO: 23 is yet another exemplary unique sequence identifier sequence.
CAGCGGGATAGTGCGATTC

SEQ ID NO: 24 is an exemplary oligo/polynucleotide sequence that can be used in *in situ* polymerase extension methods.
TAACTTACGGAGTCGCTCTACG

SEQ ID NO: 25 is an exemplary oligo/polynucleotide sequence that can be used in *in situ* polymerase extension.
GGATGGGATTCTTTAGGTCCTG

SEQ ID NO: 26 is an exemplary forward primer sequence comprising an NGS adaptor sequence for use in *in situ* polymerase extension.
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCG ATCT

SEQ ID NO: 27 is an exemplary reverse primer sequence comprising an NGS adaptor sequence for use in *in situ* polymerase extension.
CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTT CCGATCT

SEQ ID NO: 28 is an exemplary primer sequence that can be used for PCR analysis of the SEQ ID NO: 27.
CGACCACCGAGATCTAC

SEQ ID NO: 29 is an exemplary reverse primer that can be used in PCR analysis of isolated probes that have been extended using *in situ* polymerase extension.
ACGTGAACTTTGAAAGG

SEQ ID NO: 30 is an exemplary reverse primer that can be used in PCR analysis of isolated probes that have been extended using *in situ* polymerase extension.
GTAGATCCTGCAAAAAGAGTG

SEQ ID NO: 31 is an exemplary reverse primer that can be used in PCR analysis of isolated probes that have been extended using *in situ* polymerase extension. CCTTTCTACTGTTGGCATC

SEQ ID NO: 32 is an exemplary reverse primer that can be used in PCR analysis of isolated probes that have been extended using *in situ* polymerase extension. CTGAGAATGATTCTGTCTGG

SEQ ID NO: 33 is an exemplary reverse primer that can be used in PCR analysis of isolated probes that have been extended using *in situ* polymerase extension. CAATGCGGTCCATATATCC

SEQ ID NO: 34 is an exemplary oligo/polynucleotide sequence comprising a detectable label, an oligo/polynucleotide sequence, and photocleavable moiety. {Biotin}/CATTCAAATCCCCGAGTTGAACTTTCCTTTCAAAGTTCACGT/{PCBiotin}

SEQ ID NO: 35 is an exemplary oligo/polynucleotide sequence comprising a detectable label, an oligo/polynucleotide sequence, and photocleavable moiety. {Biotin}/TCTAGCCATTTGATGCCAACAGTAGAAAGGG/{PCBiotin}

SEQ ID NO: 36 is an exemplary oligo/polynucleotide sequence that can be labeled with a photocleavable moiety for use in the methods disclosed herein. {PCBiotin}/TAACTTACGGAGTCGCTCTACGCATTCAAATCCCCGAGTTGAACTTTC CTTTCAAAGTTCACGT

SEQ ID NO: 37 is an exemplary oligo/polynucleotide sequence that can be labeled with a photocleavable moiety for use in the methods disclosed herein. {PCBiotin}/TAACTTACGGAGTCGCTCTACG TCTAGCCATTTGATGCCAACAGTAGAAAGGG

SEQ ID NO: 38 is an exemplary oligo/polynucleotide sequence that is modified

to comprise a caged thymidine and further comprises a biotin moiety. {Biotin}/TAACTTACGGAGTCGCTCTACGCATTCAAATCCCCGAGTTGAACTTTCCT TTCAAAGT**T**CACGT

SEQ ID NO: 39 is an exemplary oligo/polynucleotide sequence that is modified to comprise a caged thymidine and further comprises a biotin moiety. {Biotin}/TAACTTACGGAGTCGCTCTACGCATTCAAATCCCCGAGTTGAACTTTCCT TTCAAAGT**T**CACG**T**T

SEQ ID NO: 40 is an exemplary oligo/polynucleotide sequence that is modified to comprise a caged thymidine and further comprises a biotin moiety. {Biotin}/TAACTTACGGAGTCGCTCTACGTCTAGCCATTTGATGCCAACAGTAGAA AGGGAAA**T**A

SEQ ID NO: 41 is an exemplary oligo/polynucleotide sequence that is modified to comprise a caged thymidine and further comprises a biotin moiety. {Biotin}/TAACTTACGGAGTCGCTCTACGTCTAGCCATTTGATGCCAACAGTAGAA AGGGAAA**T**ATC**T**T

SEQ ID NO. 42: is an exemplary primer sequence (PS1) according to Example 9. 5'-TAACTTACGGAGTCGCTCTACG-3'

SEQ ID NO. 43: is another exemplary sequence for Chr7 according to Example 9. 5'-TCTAGCCATTTGATGCCAACAGTAGAAAGGG-3'

SEQ ID NO. 44 is a Biotin-PS1-CHR7LAN5C probe according to Example 9. 5'- {Biotin}/TAACTTACGGAGTCGCTCTACGTCTAGCCATTTGATGCCAACAGTAGAA AGGG-3'

SEQ ID NO. 45 is a probe according to Example 7 (PE1.0-CHR7).

5'AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC
GATCTTCTAGCCATTTGATGCCAACAGTAGAAAGGG-3'

SEQ ID NO. 46 is a probe according to Example 7 (PE1.0-CHR17).

5'AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC
GATCTCATTCAAATCCCCGAGTTGAACTTTCCTTTCAAAGTTCACGT-3'

SEQ ID NO. 47 is a BiotinHER2Exon10 probe according to Example 10, wherein *T denotes positions of caged thymidines in caged versions of the probes:

5'-{Biotin}/TAACTTACGGAGTC
GCTCTACGGCATGGAGCACT*T*GCGAGAGGT*GAGGGCAGT*T*A-3'

SEQ ID NO.48 is a BiotinBRAFExon13 probe according to Example 10, wherein *T denotes positions of caged thymidines in caged versions of the probes.

5'-{Biotin}/TAACTTACGGAG
TCGCTCTACGAGTGGTGTGAGGGCTCCAGCTTGTAT*CACCAT*CT*A T*T*G-3'

We claim:

1. A method for performing contextual molecular diagnostics, comprising:

contacting a sample with a probe comprising a photo-cleavable moiety, a specific binding moiety to a target, and a detection tag using conditions sufficient to facilitate binding of the specific binding moiety to a target in the sample;

removing probe that does not bind to the sample;

selecting a region of the sample for irradiation based on contextual information;

irradiating the selected region of the sample with light of a wavelength and an intensity sufficient to cleave the photo-cleavable moiety, thereby freeing the detection tag from the selected region of the sample; and

detecting the detection tag.

2. The method of claim 1, wherein

contacting includes contacting the sample with multiple probes having photo-cleavable moieties and specific binding moieties to multiple targets, the multiple different probes comprising unique detection tags and unique specific binding moieties and

detecting the detection tag includes detecting the multiple unique detection tags.

3. The method of claim 1 or 2, wherein selecting further comprises selecting additional regions; the method further comprising:

iteratively and separately irradiating the additional regions.

4. The method of claim 3, the method further comprising iteratively depositing the detection tags from the additional regions in additional separate reaction vials.

5. The method of claim 4, wherein the additional separate reaction vials contain PCR primers with distinct ID sequences

6. The method of claim 5, further comprising appending the PCR primers to the detection tags.

7. The method of claim 4, further comprising iteratively and separately detecting the detection tag from the additional selected regions.

8. The method of claim 3, further comprising relating the detecting of the detection tag from the selected regions to the sample so as to create a molecular profile of the sample, wherein the molecular profile relates each selected area with the detection tags detected from the selected area across the sample.

9. The method of any of claims 1 to 8, wherein detecting the detection tag includes quantifying the number of targets in a sample.

10. The method of ay of claims 1 to 9, wherein the specific binding moiety is selected from an amino acid, a peptide, a protein, an antibody, a nucleotide, an oligonucleotide, a polynucleotide, a primer, an aptamer, a binding variable region, a plasmid, DNA, RNA, miRNA, or combinations thereof.

11. The method of claim 1, wherein the method further comprises:

irradiating a second selected region of the sample to cleave the photo-cleavable moiety, thereby freeing the detection tag from the second selected region of the sample; and

detecting the detection tag that was bound to the target within the second selected region of the sample.

12. The method of any of claim 1 to 11, wherein the detection tag is a unique sequence identifier or a tag sequence.

13. The method of claim 12, wherein the unique sequence identifier is configured for analysis using an array.

14. The method of claim 12, comprising sequencing the unique sequence identifier or the tag.

15. The method of claim 12, wherein the unique sequence identifier is configured for sequencing using next generation sequencing.

16. The method of claim 12, wherein the unique sequence identifier comprises a reagent barcode, an adapter sequence, an alignment sequence, a subject index sequence, or combinations thereof.

17. The method of claim 12, wherein the unique sequence identifier or the tag sequence encodes information.

18. The method of claim 16, wherein the reagent barcode encodes lot data, manufacturing data, date information, source information, or combinations thereof.

19. The method of claim 16, wherein the reagent barcode comprises from about 6 nucleotides to about 200 nucleotides, the adapter sequence comprises from about 5 nucleotides to about 50 nucleotides, the alignment sequence comprises from about 3 nucleotides to at least about 20 nucleotides, and the subject index sequence comprises from at least 5 nucleotides to about 10 nucleotides.

20. The method of claim 12, wherein the tag sequence is capable of hybridizing with a template sequence to form a duplex.

21. The method of claim 12, wherein the tag sequence comprises one or more peptide nucleic acids, locked-nucleic acids, or combinations thereof.

22. The method of any of claims 1 to 11, wherein the detection tag further comprises a detectable label.

23. The method of claim 22, wherein the detectable label is a hapten.

24. The method of claim 23, wherein the method further includes purifying the detection tag using by binding the hapten to an anti-hapten anti-body.

25. The method of any of claims 1 to 24, wherein the probe further comprises at least one linker that links two probe components.

26. The method of any of claims 1 to 25, wherein the photo-cleavable moiety is cleaved by light having a wavelength ranging from about 200 nm to about 400 nm.

27. The method of any of claims 1 to 25, wherein the photo-cleavable moiety is cleaved by light having a wavelength ranging from about 400 nm to about 790 nm.

28. The method of any of claims 1 to 25, wherein the photo-cleavable moiety is selected from a photo-cleavable phosphoramidite, photo-cleavable biotin, or photo-cleavable amine.

29. The method of any of claims 1 to 28, wherein the sample is irradiated with light indirectly.

30. The method of any of claims 1 to 29, wherein the sample is irradiated with light by using a digital micromirror device.

31. The method of any of claims 1 to 28, wherein the sample is irradiated with light directly.

32. The method of claim 29, wherein the sample is irradiated with light indirectly by using a liquid crystal display.

33. The method of any of claims 1 to 32, wherein the sample is irradiated with light for a time sufficient to cleave the photo-cleavable moiety.

34. The method of any of claims 1 to 33, wherein the sample is irradiated for a time period ranging from about 1 second to about 60 seconds.

35. The method of any of claims 1 to 34, wherein the light is produced by a light source selected from a laser, an LED, or a lamp.

36. The method of any of claims 1 to 35, wherein the conditions sufficient to bind the specific binding moiety of the probe to the sample are suitable for inducing hybridization and/or chemical coupling of the probe to the sample.

37. The method of any of claims 1 to 36, wherein removing probe that does not bind to the sample includes washing the sample with a suitable washing solution.

38. The method of any of claims 1 to 37, wherein selecting the region of the sample comprises manually marking one or more portions of the sample or digitally marking one or more portions of the sample.

39. The method of any of claims 1 to 38, wherein the contextual information used for selecting is created using a combination of one or more of primary staining, secondary staining, fluorescence imaging, phase contrast imaging, morphological sample characteristics imaging, image analysis, digital image analysis, cell selection, or grouping selection.

40. The method of any of claims 1 to 39, wherein the method further comprises contacting the sample with a visually detectable reagent and selecting includes the contextual information provided by the visually detectable reagent.

41. The method of claim 40, wherein the detectable reagent includes a fluorescent or chromogenic reagent.

42. The method of any of claims 1 to 41, wherein detection of the detection tag indicates that the target is present within the region.

43. The method of claim 1, wherein detection of the detection tag further comprises sequencing the detection tag using next generation sequencing, analyzing the detection tag using PCR, or analyzing the detection tag using an array platform.

44. The method of claim 1, further comprising contacting the sample with a second probe comprising a second specific binding moiety, a second photo-cleavable moiety, and a second detection tag.

45. The method of claim 44, wherein the photo-cleavable moiety and the second photo-cleavable moiety are the same.

46. The method of claim 44 or 45, wherein the detection tag and the second detection tag are different and

the specific binding moiety and the second specific binding moiety are different.

47. The method of any of claims 44 to 46, further comprising relating the detecting of the detection tag from the selected regions to the sample so as to create a molecular profile of the sample, wherein the molecular profile relates each selected area with the detection tags detected from the selected area across the sample and a unique specific binding moiety

48. The method of any of claims 44 to 47 wherein contacting the sample with the probe occurs simultaneously with contacting the sample with the second probe.

49. The method of any of claims 1 to 48, wherein selecting further comprises selecting a control region; the method further comprising:

separately irradiating the control region and

separately detecting whether the detection tag is present in the control region.

50. A method for performing time-elapsed contextual molecular diagnostics, comprising:

performing the method of claim 1 on a first selected region,

archiving the sample,

removing the sample from archival, and

performing the method of claim 1 on a second selected region .

51. A method for performing contextual molecular diagnostics, the method comprising:

contacting a sample with a probe comprising a photo-reactive moiety, a specific binding moiety, and a detection tag using conditions sufficient to facilitate binding of the specific binding moiety to a target in the sample;

removing probe that does not bind to the sample;

selecting a region of the sample for irradiation based on contextual information;

irradiating the selected region of the sample with light of a wavelength and an intensity sufficient to cause the photo-reactive moiety to react, thereby freeing the detection sequence for further reaction;

exposing the tissue sample to an enzyme that acts on the detection sequence to effect a change; and

detecting the change.

52. The method of claim 51, wherein the detection tag is an extendable primer, the photo-reactive moiety is a photo-cleavable blocking group, the enzyme is a polymerase, and the change is an extension of a primer sequence in the presence of dATP, dCTP, dTTP, and dGTP.

53. The method of claim 51 or 52, further comprising ligating adjacent detection tag after irradiating the selected region.

54. The method of claim 51, wherein the photo-cleavable moiety includes a caged-ATP molecule.

55. The method of claim 54, wherein the caged-ATP molecule is converted to an ATP molecule capable of modifying a probe present on the sample with a phosphate group.

56. The method of claim 55, further comprising modifying the phosphate group with a detectable label, a member of a specific binding pair, or a combination thereof.

57. The method of claim 56, further comprising detecting the detectable label or isolating the probe using the member of the specific binding pair.

58. The method of any of claims 51 to 57, wherein contacting includes multiple probes having specific binding moieties to multiple targets, the multiple different probes comprising unique detection tags and unique specific binding moieties and detecting the detection tag includes detecting the multiple unique detection tags.

59. The method of any of claims 51 to 58, wherein selecting further comprises selecting additional regions; the method further comprising:

iteratively and separately irradiating the additional regions and

iteratively and separately detecting the detection tag from the additional selected regions.

60. The method of claim 59, further comprising relating the detecting of the detection tags from the selected regions to the sample so as to create a molecular profile of the sample, wherein the molecular profile relates each selected area with the detection tags detected from the selected area across the sample.

61. The method of any of claims 51 to 60, wherein detecting the detection tag includes quantifying the number of targets in a sample.

62. The method of any of claims 51 to 61, wherein the specific binding moiety is selected from an amino acid, a peptide, a protein, an antibody, a nucleotide, an oligonucleotide, a polynucleotide, a primer, an aptamer, a binding variable region, a plasmid, DNA, RNA, miRNA, or combinations thereof.

63. The method of claim 51, wherein the method further comprises:

irradiating a second selected region of the sample to cause the photo-reactive moiety to react, thereby freeing the detection sequence for further reaction from the second selected region of the sample; and

detecting the detection tag that was bound to the target within the second selected region of the sample.

64. The method of claim 51, wherein the detection tag is a unique sequence identifier or a tag sequence.

65. The method of claim 51, wherein selecting the region of the sample comprises manually marking one or more portions of the sample or digitally marking one or more portions of the sample.

66. The method of any of claims 51 to 65, wherein the contextual information used for selecting is created using a combination of one or more of primary staining, secondary staining, fluorescence imaging, phase contrast imaging, morphological sample characteristics imaging, image analysis, digital image analysis, cell selection, or grouping selection.

67. The method of any of claims 51 to 65, wherein the method further comprises contacting the sample with a visually detectable reagent and selecting includes the contextual information provided by the visually detectable reagent.

68. The method of any of claims 51 to 67, wherein detection of the detection tag indicates that the target is present within the region.

69. The method of claim 51, wherein detection of the detection tag further comprises sequencing the detection tag using next generation sequencing, analyzing the detection tag using PCR, or analyzing the detection tag using an array platform.

70. The method according to claim 51, further comprising contacting the sample with a second probe comprising a second specific binding moiety, a second photo-reactive moiety, and a second detection tag.

71. The method according to claim 70, wherein the photo-reactive moiety and the second photo-reactive moiety are the same, the detection tag and the second detection tag are different, and the specific binding moiety and the second specific binding moiety are different.

72. A system for performing contextual diagnostics using a photo-activated probe, comprising:

a slide configured to accept a sample;

a vessel containing the photo-activated probe;

a slide imaging device configured to receive the slide, and

a light source configured to selectively irradiate the sample.

73. The system of claim 72, wherein the photo-activated probe comprises a specific binding moiety, a photo-activated moiety, and a detection tag, wherein the photo-activated probe is configured to be detectable upon selective irradiation.

74. The system of claim 72 or 73, wherein the slide is a single specimen slide, an array slide, or a microarray slide.

75. The system of any of claims 72 to 74, wherein the light source is selected from a laser, an LED, or a lamp.

76. The system of any of claims 72 to 75, wherein
the slide imaging device is configured to capture a sample image.

77. The system of claim 76, the system further comprising

a digital screen in communication with the imaging device for receiving and displaying the sample image from the imaging device; and

a selection device in communication with the digital screen for marking a selected portion of the sample image for analysis.

78. The system of claim 77, wherein the light source is coupled to the selection device for irradiating the selected portion.

79. The system of any of claims 76 to 78, wherein the slide imaging device is a digital camera or a video camera.

80. The system of any of claims 73 to 79, further comprising:

a dispensing instrument configured to dispense an elution buffer onto the slide thereby facilitating isolation of the detection tag; and

a recovery instrument configured to recover the detection tag.

81. The system of any of claims 73 to 80, further comprising a PCR instrument configured to accept and amplify the detection tag.

82. The system of any of claims 73 to 81, further comprising:

a sequencing instrument configured to receive and sequence the detection tag; and

a computer configured to translate information generated from the sequencing instrument to the digital screen, thereby correlating the detection tag to the selected region from which it was cleaved.

83. The system of any of claims 73 to 80, further comprising an array platform comprising a modified surface configured to accept and bind the detection tag.

84. A kit for performing contextual diagnostics, comprising a probe reagent

solution comprising a specific binding moiety, a photo-cleavable moiety, and a detection tag.

85. A method for performing contextual molecular diagnostics, the method comprising:

contacting a sample with a probe comprising a specific binding moiety, using conditions sufficient to facilitate binding of the specific binding moiety to a target in the sample;

removing probe that does not bind to the sample;

contacting a sample with a photo-reactive moiety;

selecting a region of the sample for irradiation based on contextual information;

irradiating the selected region of the sample with light of a wavelength and an intensity sufficient to cause the photo-reactive moiety to react;

exposing the tissue sample to an enzyme that acts on the detection oligonucleotide to effect a change; and

detecting the change.

86. The method according to claim 85, wherein the detection tag is an extendable primer, the photo-reactive moiety is one or more caged deoxynucleoside triphosphate molecule(s), taken from the group of deoxynucleoside triphosphates including dATP, dCTP, dGTP, dTTP, and dUTP or derivatives thereof, with exclusion of the corresponding uncaged nucleoside triphosphates.

87. The method of claim 86, wherein the caged-nucleoside triphophate molecule(s) is/are converted to nucleoside triphosphate molecule(s) capable of being incorporated by a polymerase to extend the extendable primer.

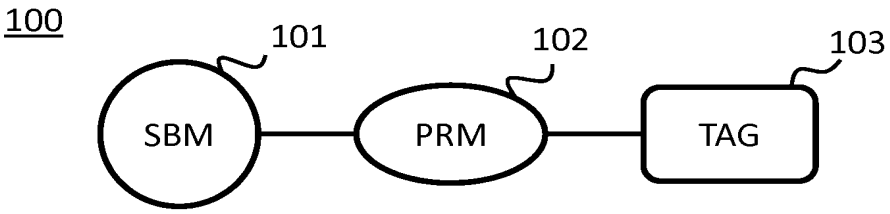


FIG. 1(A)

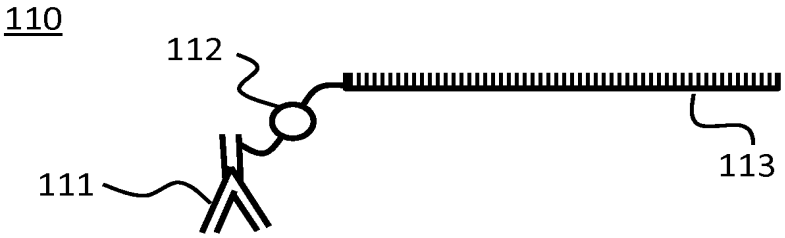


FIG. 1(B)

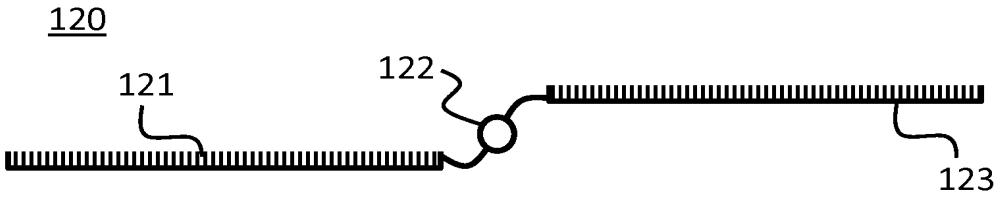


FIG. 1(C)

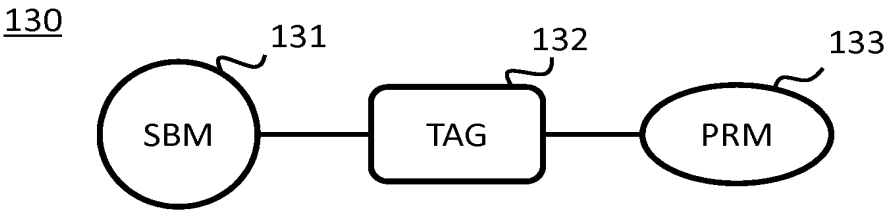


FIG. 1(D)

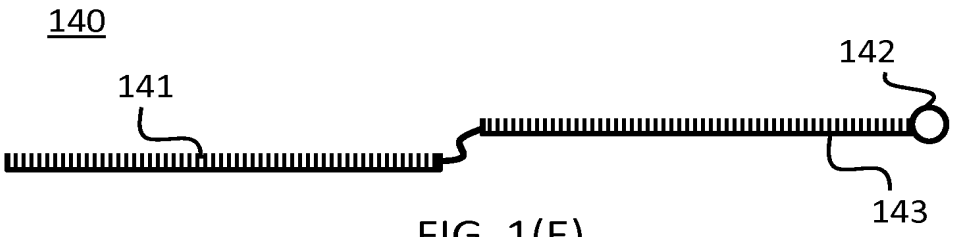


FIG. 1(E)

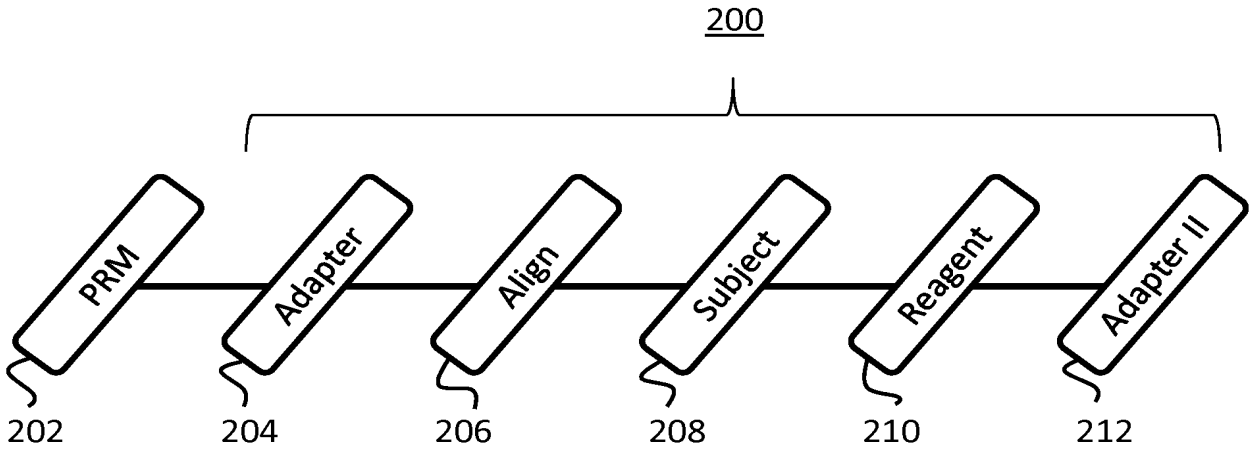

200
PRM
Adapter
Align
Subject
Reagent
Adapter II
202
204
206
208
210
212

FIG. 2(A)

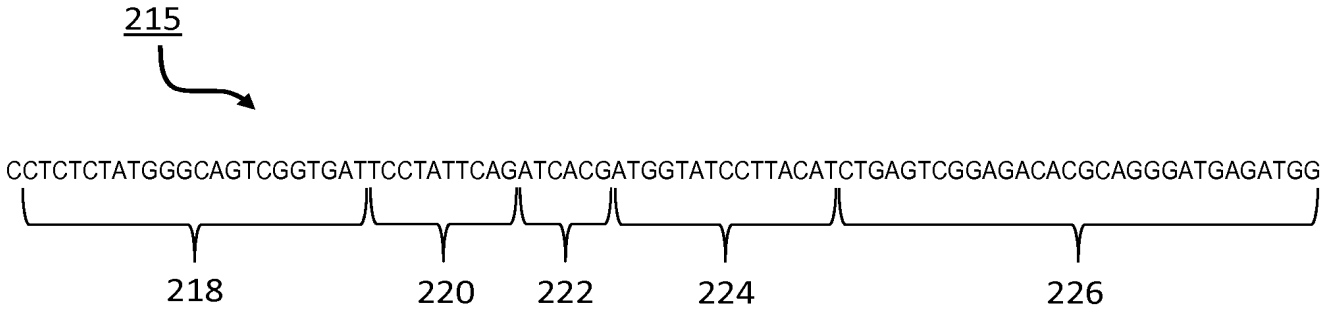

215
CCTCTCTATGGGCAGTCGGTGATTCCTATTCAGATCACGATGGTATCCTTACATCTGAGTCGGAGACACGCAGGGATGAGATGG
218
220
222
224
226

FIG. 2(B)

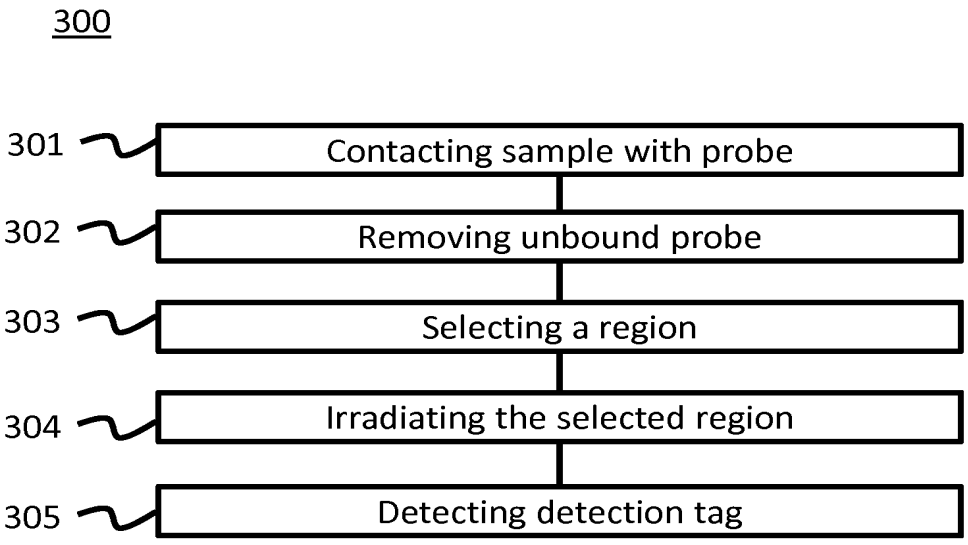

300
301
Contacting sample with probe
302
Removing unbound probe
303
Selecting a region
304
Irradiating the selected region
305
Detecting detection tag

FIG. 3

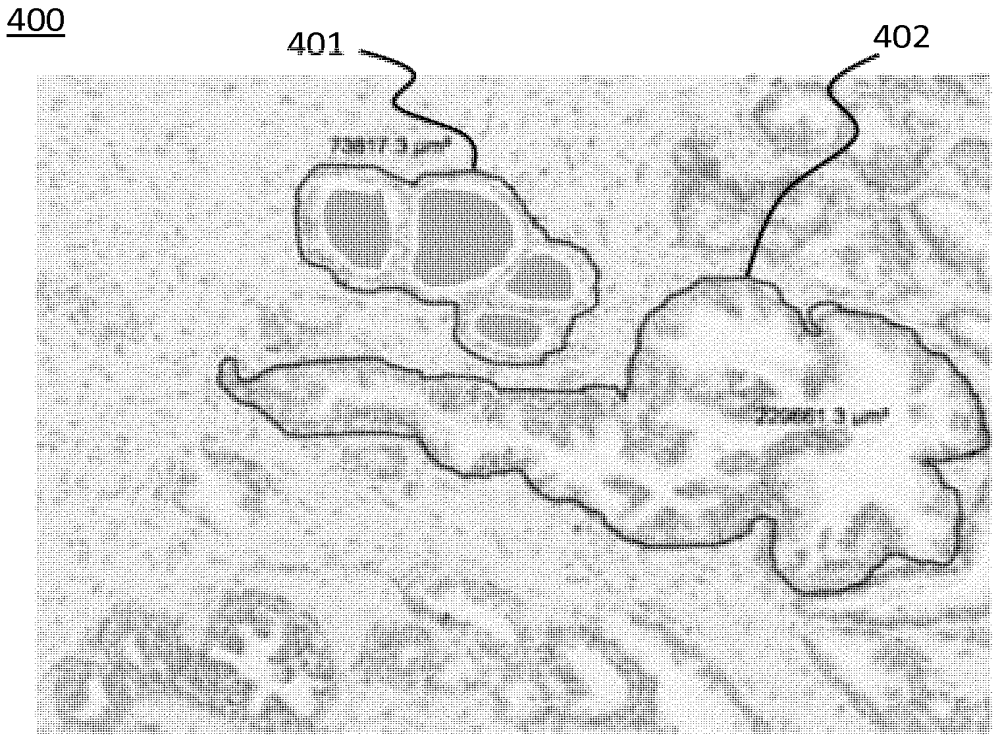

400
401
402

FIG. 4

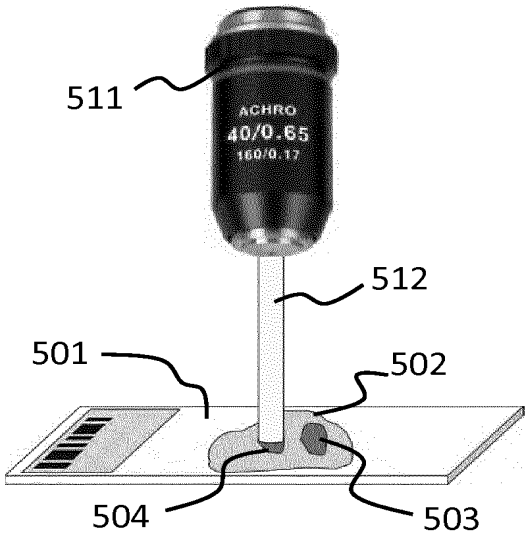


FIG. 5(A)

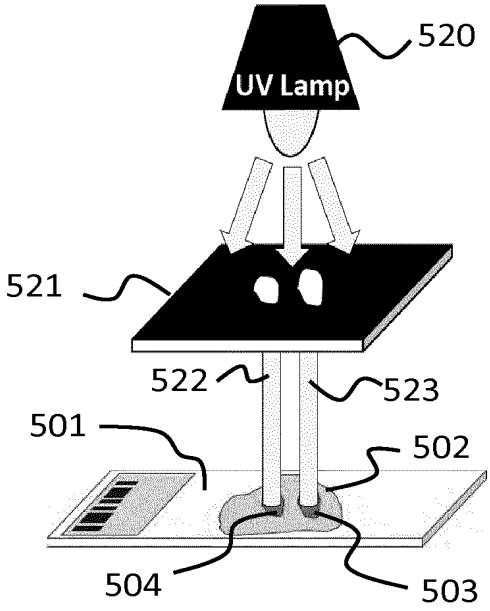


FIG. 5(B)

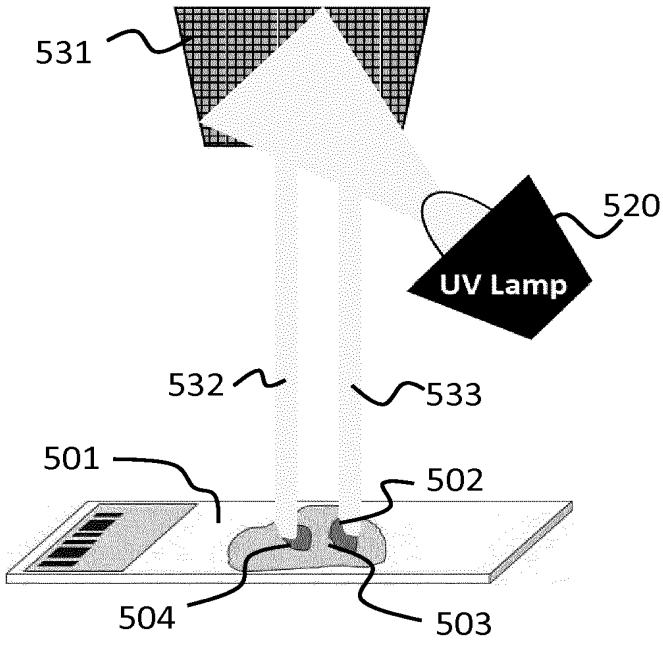


FIG. 5(C)

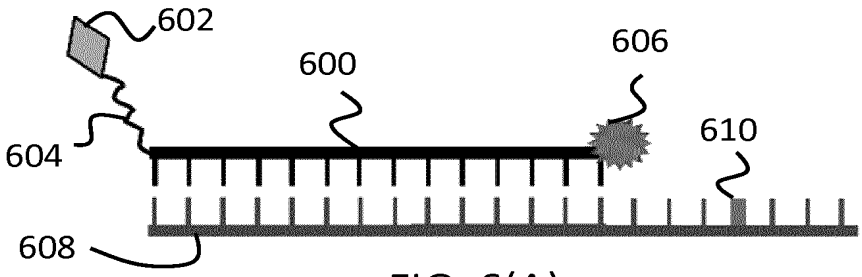


FIG. 6(A)

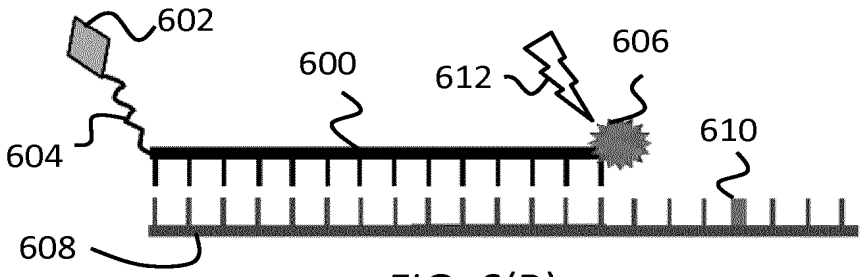


FIG. 6(B)

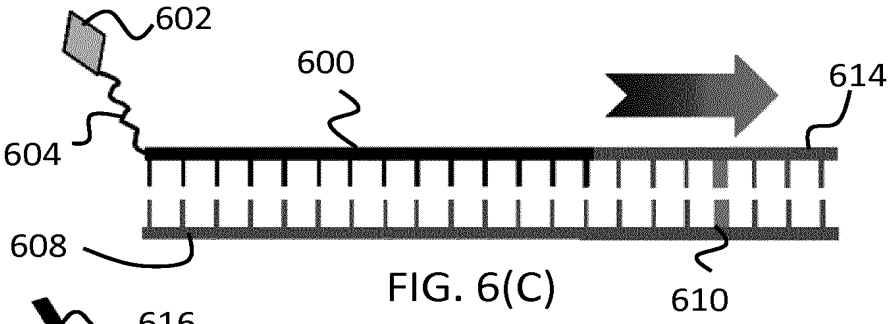


FIG. 6(C)

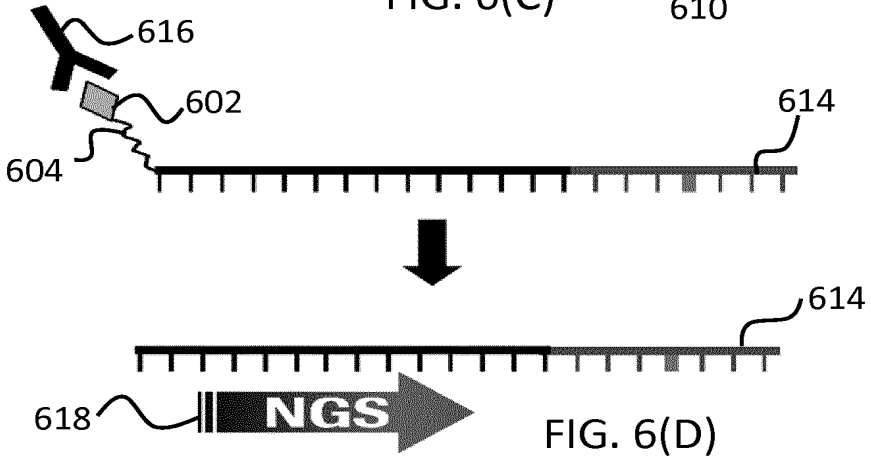


FIG. 6(D)

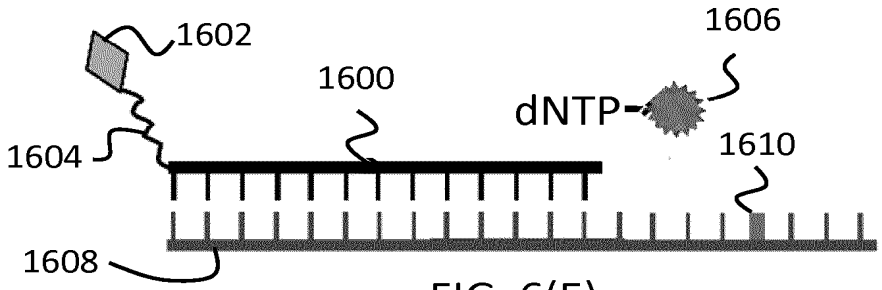

1602
1606
1600
dNTP
1604
1610
1608

FIG. 6(E)

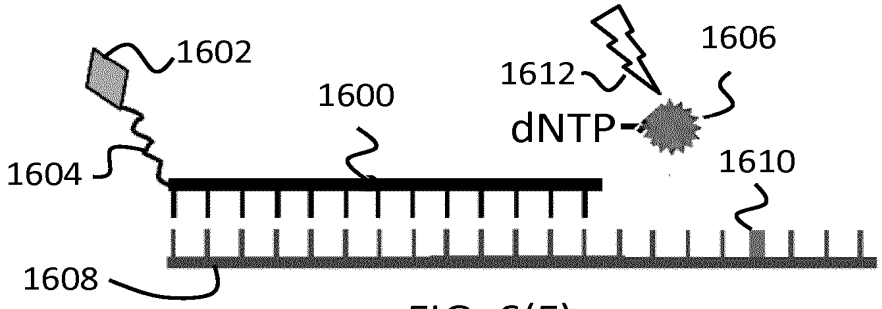

1602
1606
1612
1600
dNTP
1604
1610
1608

FIG. 6(F)

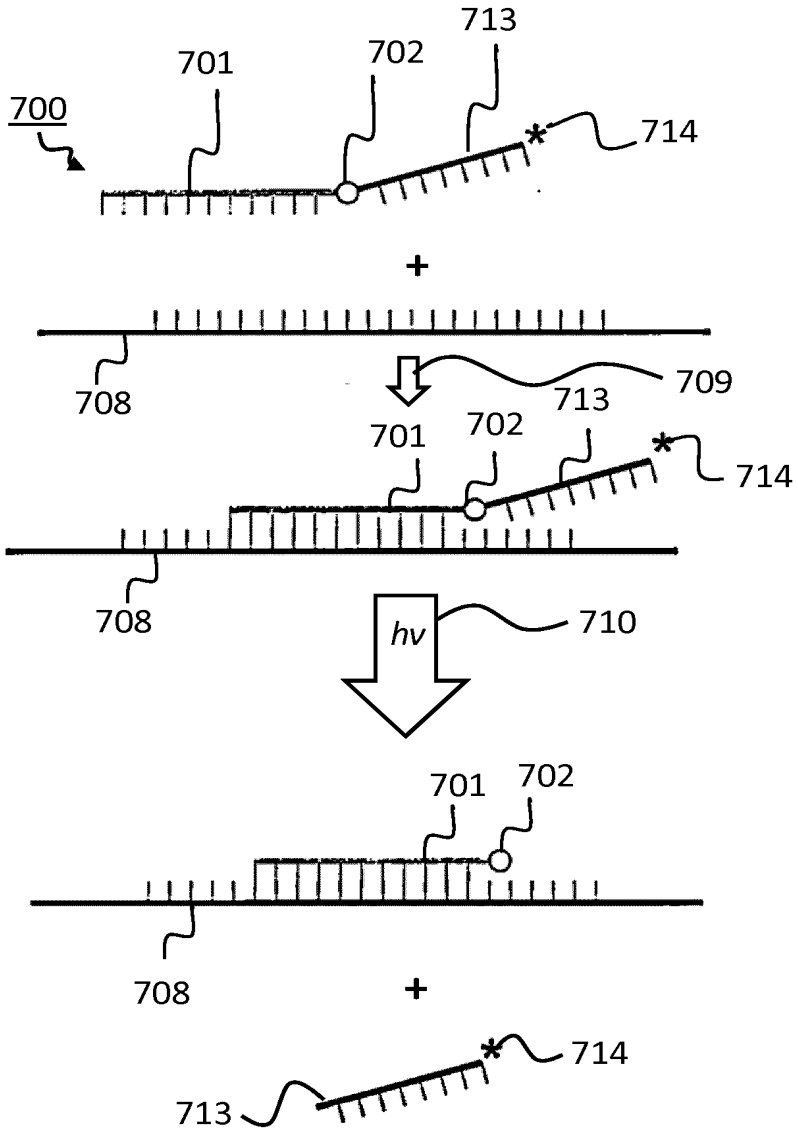

713
701
702
700
714
+
708
709
713
701
702
714
708
hv
710
701
702
708
+
714
713

FIG. 7

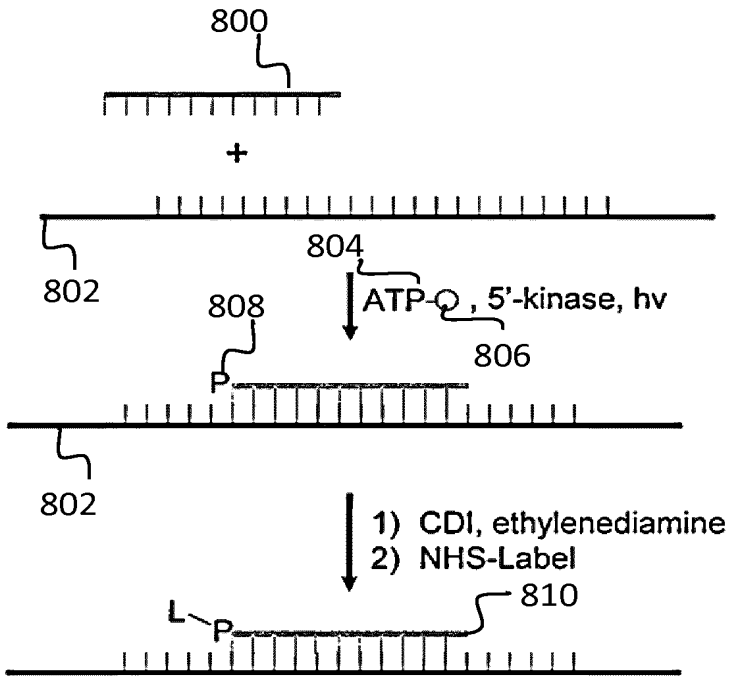

800
+
802
804
ATP-Q , 5'-kinase, hv
808
806
P
802
1) CDI, ethylenediamine
2) NHS-Label
810
L
P

FIG. 8

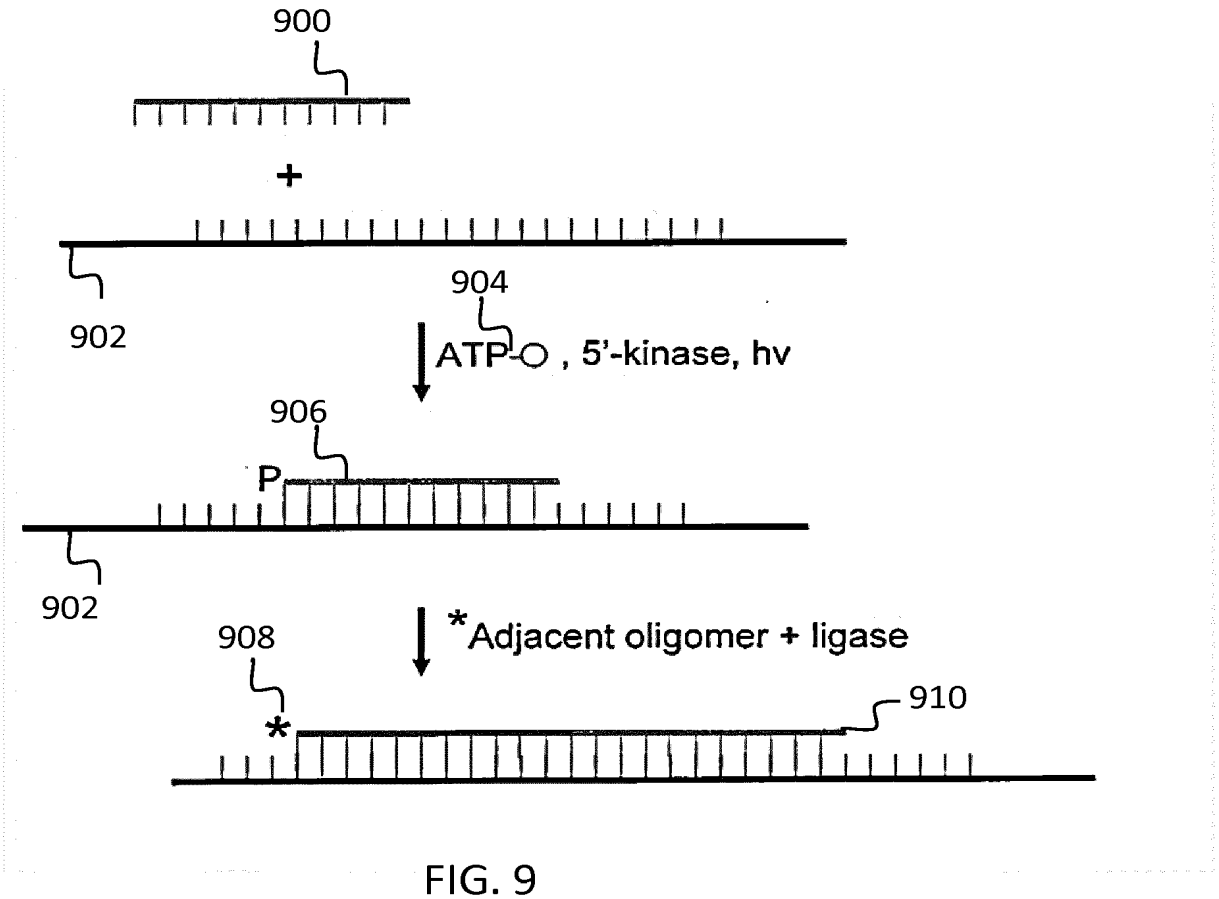

900
+
902
904
ATP-O , 5'-kinase, hv
906
P
902
908
*Adjacent oligomer + ligase
910
*

FIG. 9

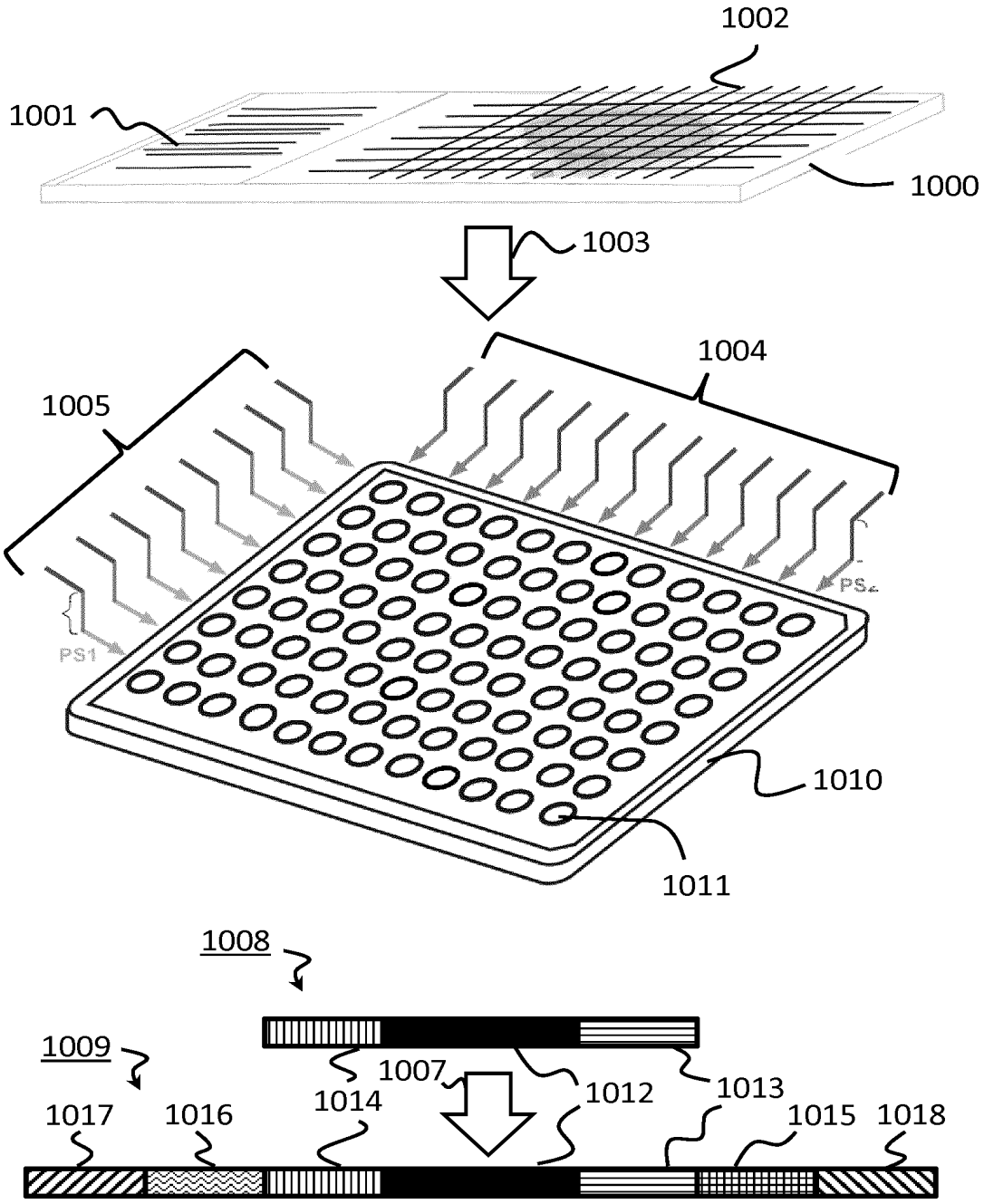

1002
1001
1000
1003
1004
1005
PS1
1010
1011
1008
1009
1007
1012
1013
1014
1017
1016
1015
1018

FIG. 10

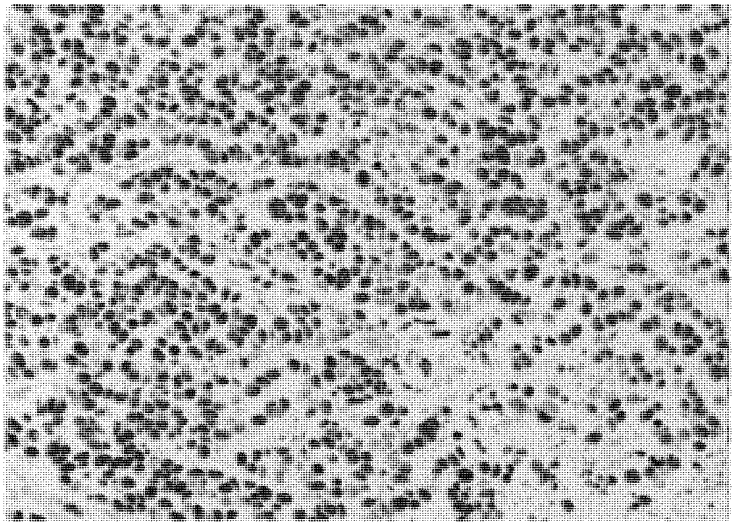


FIG. 11

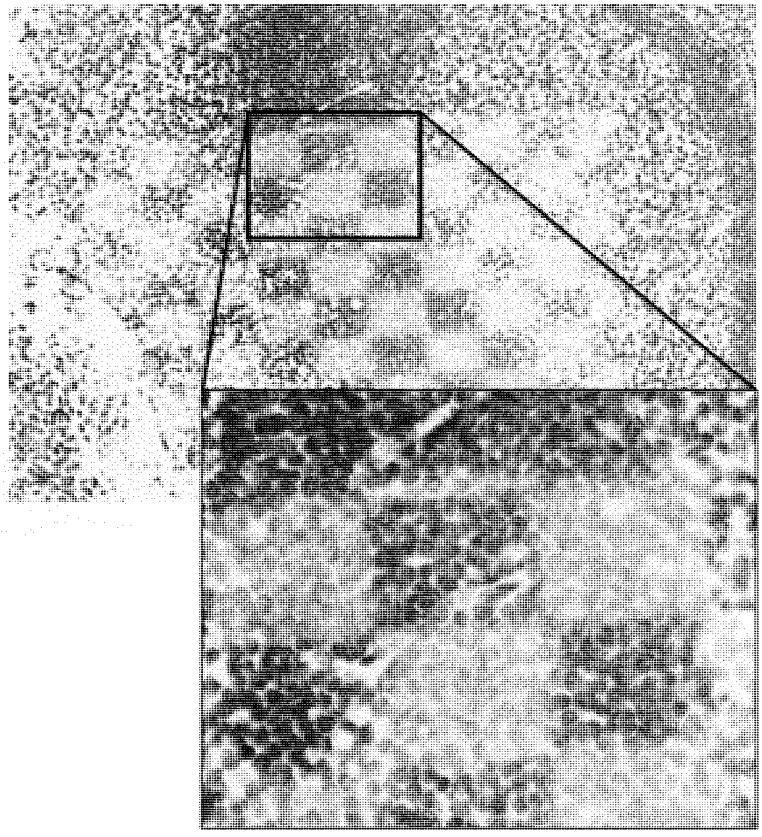


FIG. 12

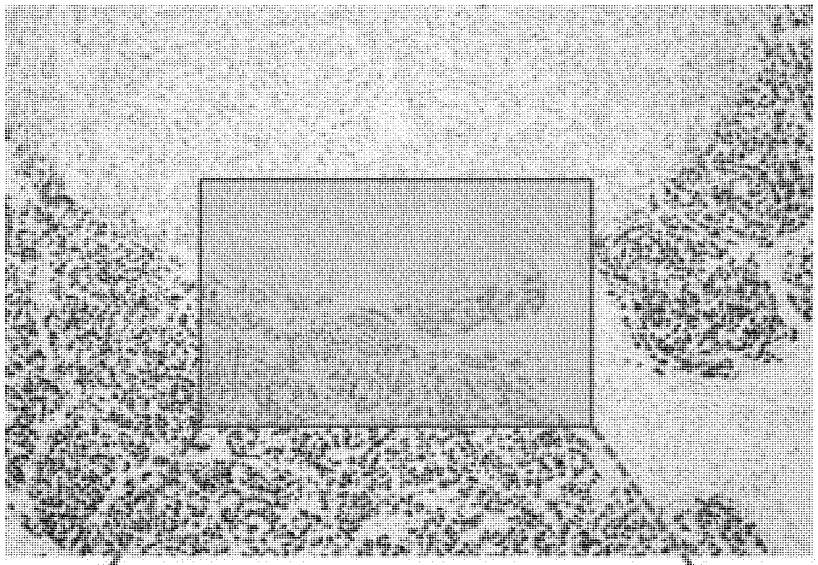


FIG. 13

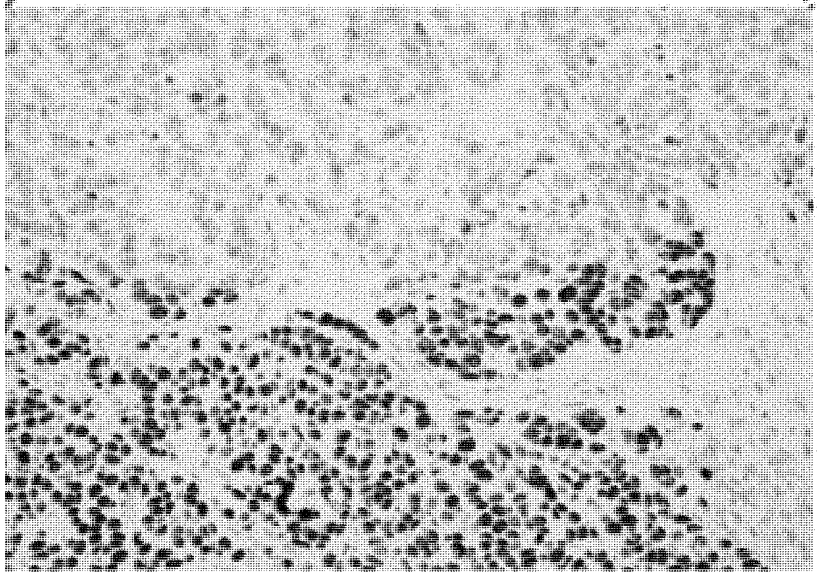


FIG. 14

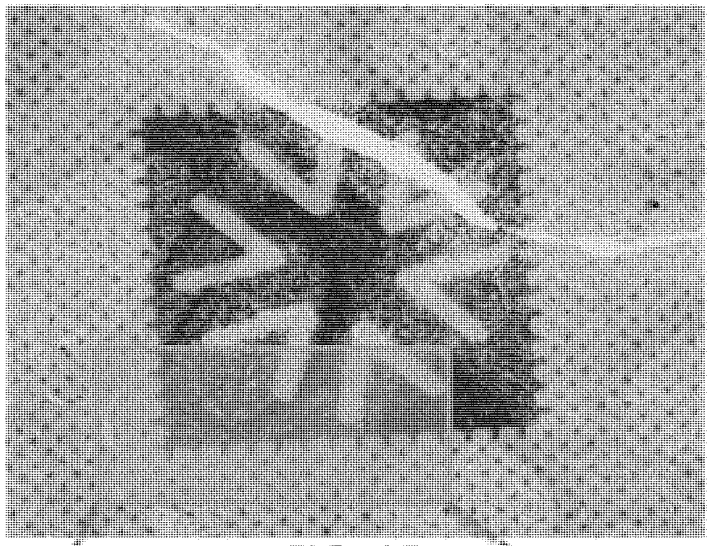


FIG. 15

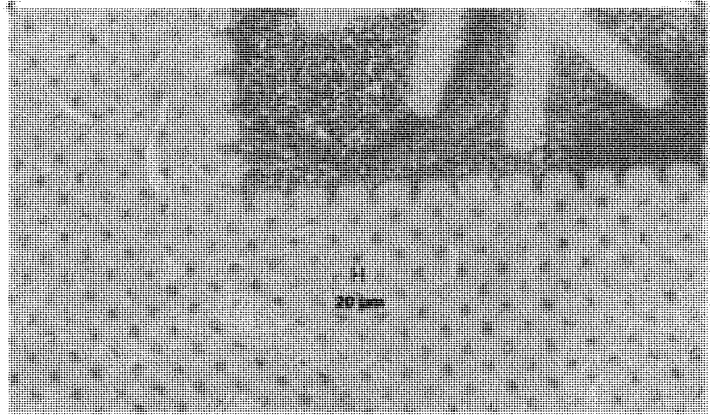


FIG. 16

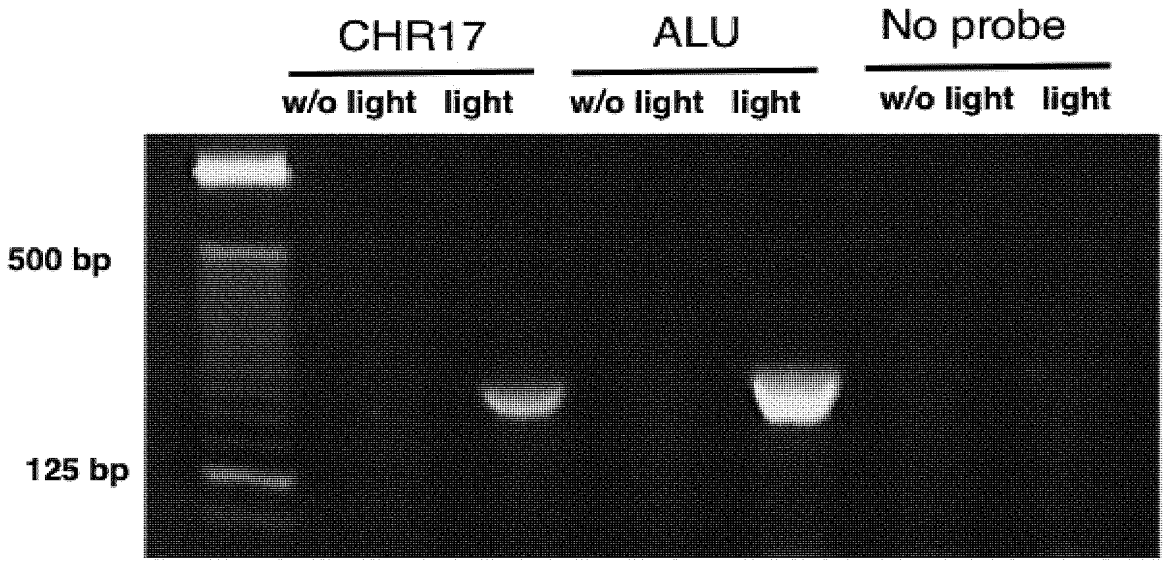


FIG. 17

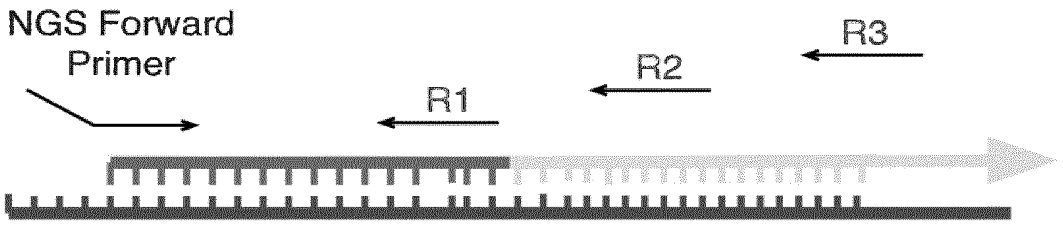


FIG. 18

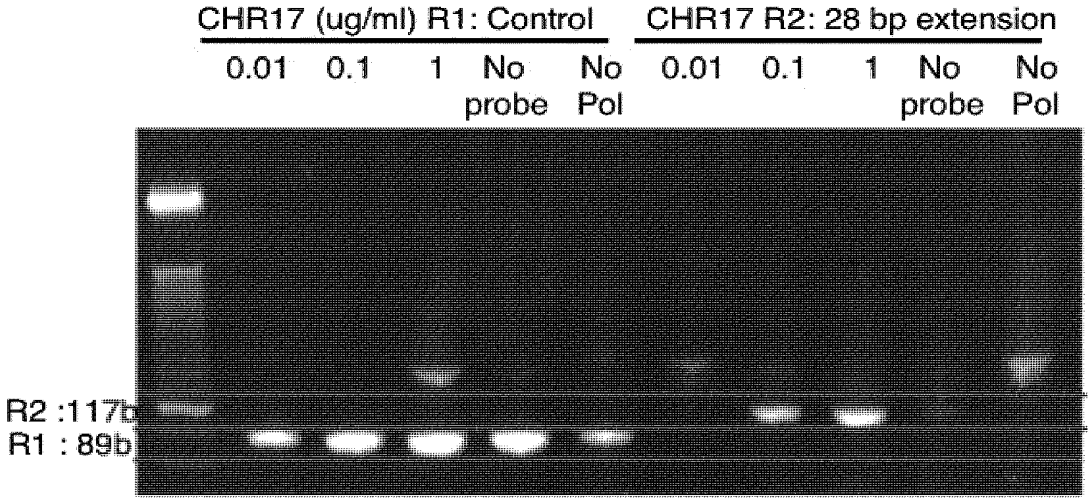


FIG. 19

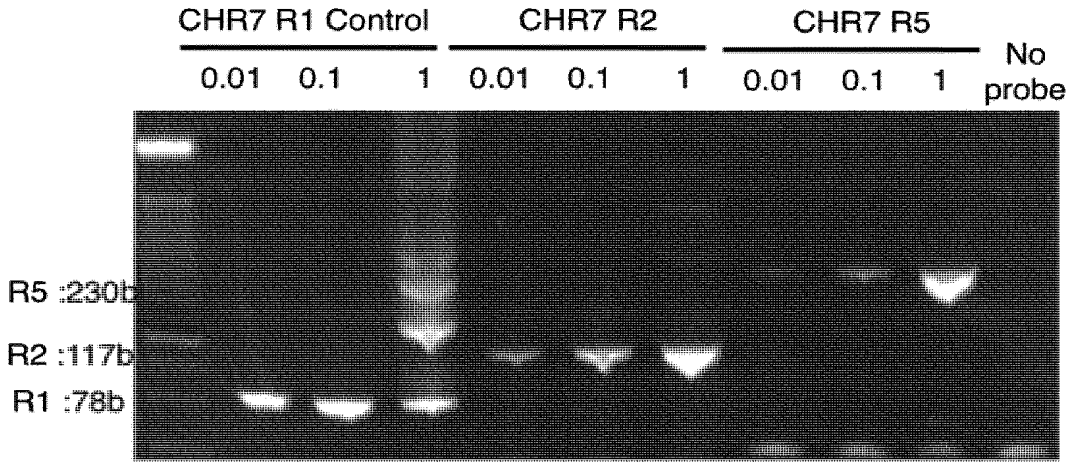

CHR7 R1 Control
CHR7 R2
CHR7 R5
0.01
0.1
1
0.01
0.1
1
0.01
0.1
1
No probe
R5 :230b
R2 :117b
R1 :78b

FIG. 20

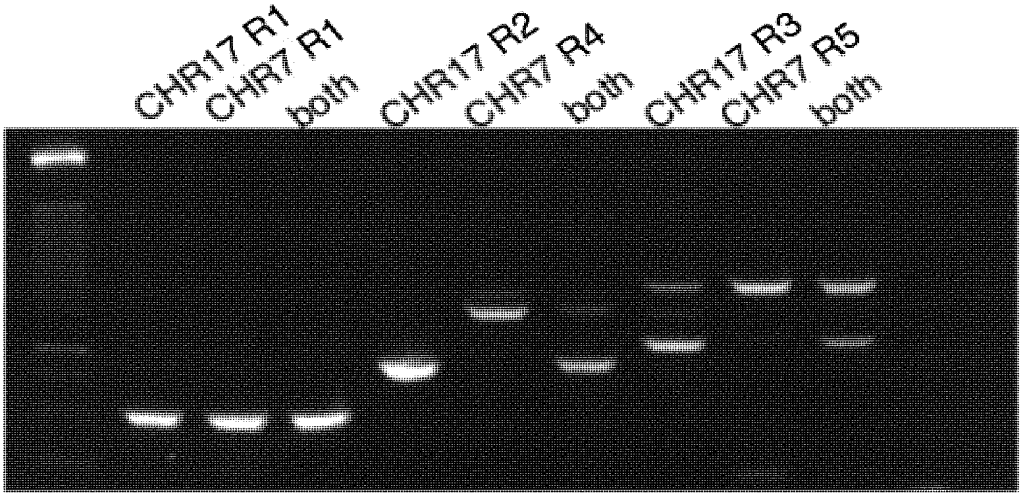

CHR17 R1
CHR7 R1
both
CHR17 R2
CHR7 R4
both
CHR17 R3
CHR7 R5
both

FIG. 21

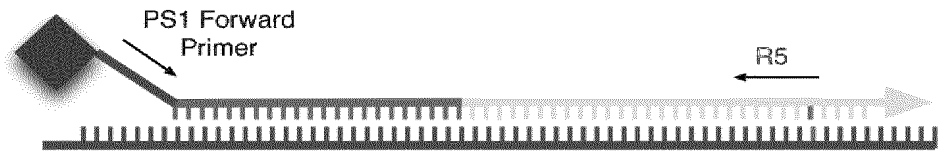


FIG. 22(A)

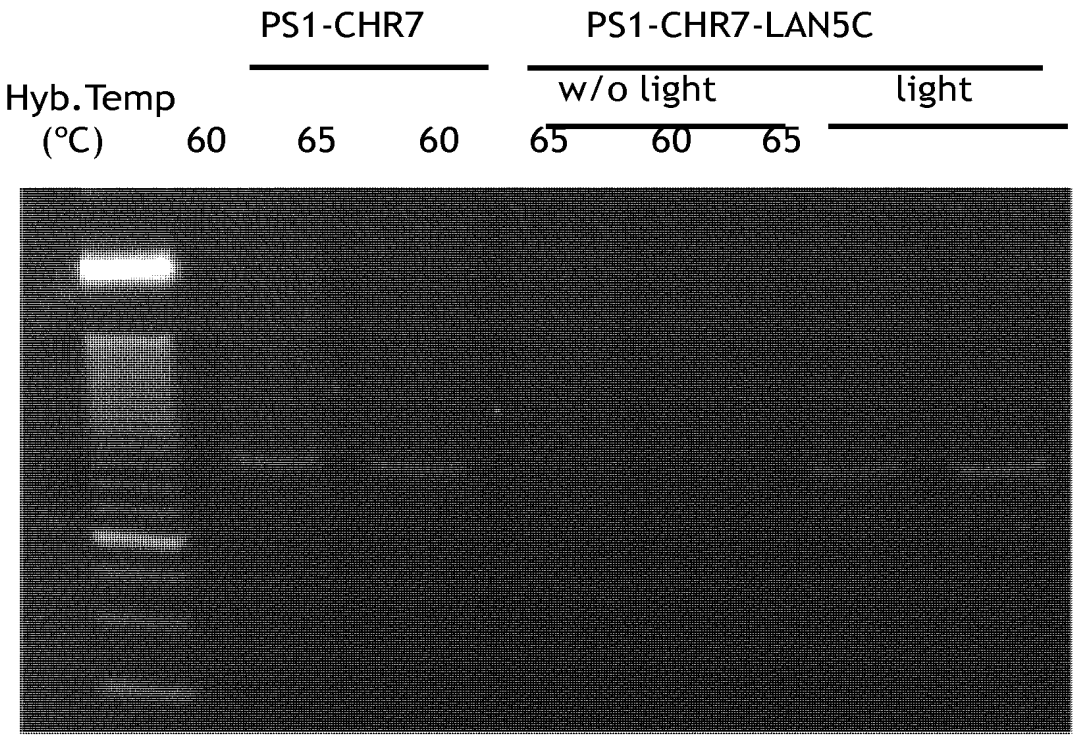


FIG. 22(B)

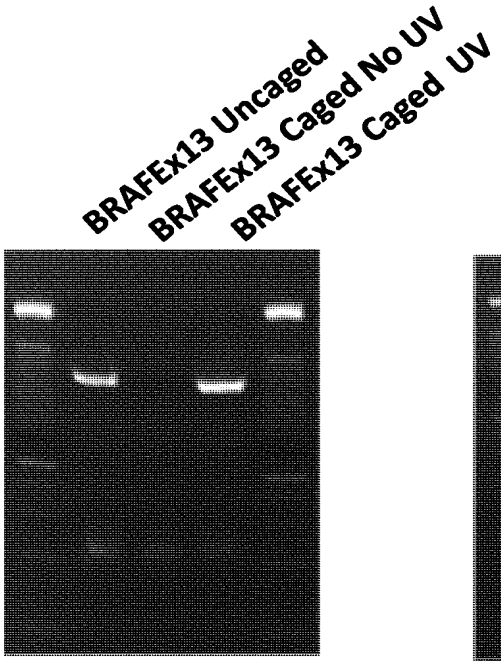


FIG. 23(A)

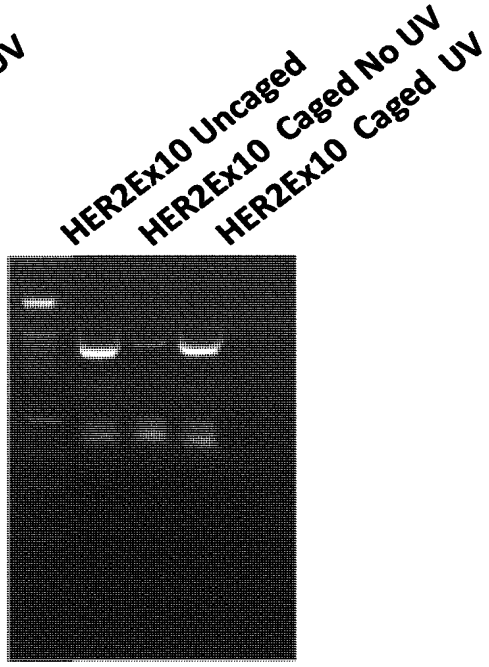


FIG. 23(B)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 January 2017 (26.01.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/015099 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2016/042460

(22) International Filing Date:
15 July 2016 (15.07.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

62/193,809	17 July 2015 (17.07.2015)	US
62/261,657	1 December 2015 (01.12.2015)	US
62/277,289	11 January 2016 (11.01.2016)	US
62/323,023	15 April 2016 (15.04.2016)	US

(71) Applicant: **NANOSTRING TECHNOLOGIES, INC.** [US/US]; 530 Fairview Avenue North, Suite 2000, Seattle, Washington 98109 (US).

(72) Inventors: **BEECHEM, Joseph M.**; 46 Ridgeline Drive, Eugene, Oregon 97405 (US). **WARREN, Charles**; 2280 Enetai Beach Drive, Bremerton, Washington 98310 (US). **MERRITT, Chris**; 1021 NE 125th Street, Unit A, Seattle, Washington 98125 (US). **JUNG, Jaemyeong**; 14010 SE 51st Place, Bellevue, Washington 98006 (US). **DUNAWAY, Dwayne L.**; 2751 NE 94th Street, Seattle, Washington 98115 (US). **CROWDER, Scott**; c/o NanoString Technologies, Inc., 530 Fairview Avenue North, Suite 2000, Seattle, Washington 98109 (US). **SORG, Kristina**; 300 N 42nd Street, Apt. 1, Seattle, Washington 98103 (US).

(74) Agents: **PAVAO, Matthew** et al.; Cooley LLP, 1299 Pennsylvania Avenue, NW, Suite 700, Washington, District of Columbia 20004 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

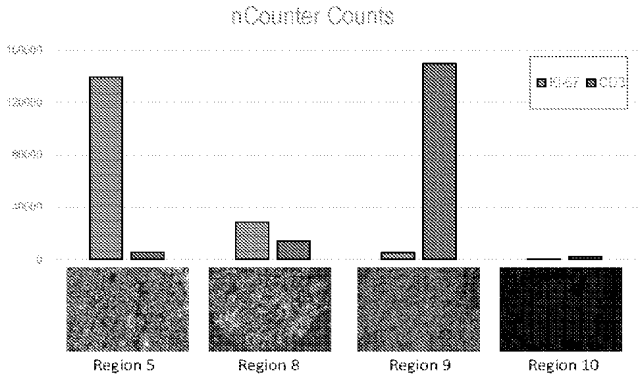
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

(54) Title: SIMULTANEOUS QUANTIFICATION OF GENE EXPRESSION IN A USER-DEFINED REGION OF A CROSS-SECTIONED TISSUE

Figure 19



(57) Abstract: The present invention relates to, among other things, probes, compositions, methods, and kits for simultaneous, multiplexed detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell.

WO 2017/015099 A1

SIMULTANEOUS QUANTIFICATION OF GENE EXPRESSION IN A USER-DEFINED REGION OF A CROSS-SECTIONED TISSUE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is claims priority to and the benefit of U.S. Provisional Application No. 62/193,809, filed July 17, 2015; U.S. Provisional Application No. 62/261,657, filed December 1, 2015; U.S. Provisional Application No. 62/277,289, filed January 11, 2016; and U.S. Provisional Application No. 62/323,023, filed April 15, 2016. Each of the above-mentioned applications is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

[0002] Standard immunohistochemical and *in situ* hybridization methods allow for simultaneous detection of, at most, six to ten protein or nucleic acid targets, with three to four targets being typical. There exists a need for probes, compositions, methods, and kits for simultaneous, multiplexed detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell.

SUMMARY OF THE INVENTION

[0003] The present invention relates to probes, compositions, methods, and kits for simultaneous, multiplexed detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell.

[0004] An aspect of the present invention relates to a method including steps of (1) contacting at least one target in or from at least one cell in a tissue sample with at least one probe comprising a target-binding domain and a signal oligonucleotide; (2) providing a force to a location of the tissue sample sufficient to release the signal oligonucleotide; and (3) collecting and identifying the released signal oligonucleotide, thereby detecting the at least one target in or from a specific location of the tissue sample that was provided the force. The specific location is a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell. The

target may be a nucleic acid, (*e.g.,* mRNA and miRNA) and/or a protein. The target-binding domain may be a single-stranded nucleic acid, a partially double-stranded nucleic acid, or a protein-binding molecule, *e.g.*, an antibody, a peptide, an aptamer, and a peptoid.
In embodiments, two or more targets (*i.e.*, proteins, nucleic acids, and a combination thereof) are detected. In embodiments, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more targets, and any number therebetween, are detected; for example, 800 or more different targets can be detected. In embodiments, detecting includes quantifying the abundance of each target.

[0005] In embodiments, the method further includes repeating at least steps (2) and (3) on at least a second specific location of the tissue sample, the second specific location comprising at least a second cell. In embodiments, detecting includes comparing the abundance of the at least one target in or from the first specific location and in or from the at least second specific location. The at least one cell and at least second cell may be the same cell type or distinct cell types. In some embodiments, detecting includes quantifying the abundance of the at least one target in or from a first cell type and in or from the at least a second cell type. In embodiments, first and second cell types are independently selected from a normal cell and an abnormal cell, *e.g.*, a diseased and cancerous cell.

[0006] In embodiments, the at least one cell is directly immobilized to a surface or is indirectly immobilized to the surface *via* at least one other cell. A tissue sample may be a 2 to 1000 μm thick tissue section, *e.g.*, obtained from a formalin-fixed paraffin embedded (FFPE) sample or from an unfixed sample. The at least one cell may be fixed or unfixed. The at least one cell may be stained or labeled prior to step (2) allowing visualization of a subcellular or cellular structure in the stained or labeled cell. Alternately, for tissue sections, a section adjacent to the section that is contacted with the probes may be stained or labeled prior to step (2), thereby allowing estimation of a subcellular, cellular, or tissue-related structure in the corresponding cell or nearby cell in the section that is contacted with the probes. Such staining or labeling techniques are well known in the art.

[0007] In the above aspect, at least one probe further includes a linker (*e.g.*, a cleavable linker) located between the target-binding domain and the signal oligonucleotide. The cleavable linker may be photo-cleavable, which is cleaved by light provided by a suitable coherent light source (*e.g.*, a laser and a UV light source) or a suitable incoherent light source (*e.g.*, an arc-lamp and a

light-emitting diode (LED)). The light source may irradiate at least one subcellular structure of the at least one cell and the abundance of the at least one nucleic acid target in or from the at least one subcellular structure of the at least one cell can be detected. Also, the light source may first irradiate at least one subcellular structure in the at least one cell and later irradiate at least one subcellular structure in the at least second cell, allowing a comparison of the abundance of the at least one target in or from the at least one subcellular structure in the at least one cell and the at least one subcellular structure in the at least second cell.

[0008] In embodiments, the signal oligonucleotide is a single-stranded nucleic acid or a partially double-stranded nucleic acid.

[0009] In embodiments, the sample may be cultured cells or dissociated cells (fixed or unfixed) that have been immobilized onto a slide. The sample may comprise cells (including both primary cells and cultured cell lines) and/or tissues (including cultured or explanted). The sample may comprise a cultured cell, a primary cell, or a dissociated cell from an explant.

[0010] In embodiments, the illumination of a region of interest smaller that a field of view (for example a single cell or a subcellular structure within a cell) comprises use of a laser scanning device (e.g., confocal) or a digital mirror device (DMD) to direct the light.

[0011] In embodiments, a probe is prepared by a cysteine bioconjugation method that is stable, site-specific to, preferably, the antibody's hinge-region heavy-chain. In embodiments, a probe can comprise a plurality (i.e., more than one, e.g., 2, 3, 4, 5, or more) labeled oligonucleotides per antibody.

[0012] Detecting comprises a polymerase reaction, a reverse transcriptase reaction, hybridization to an oligonucleotide microarray, mass spectrometry, hybridization to a fluorescent molecular beacon, a sequencing reaction, or nCounter® Molecular Barcodes. In preferred embodiments, nCounter® systems and methods from NanoString Technologies® are used.

[0013] In embodiments, the signal oligonucleotide is collected from a tissue *via* liquid laminar, turbulent, or transitional flow. The flow may be *via* a channel, *e.g.*, having 25 to 500 μm depth between the tissue and a fluidic device or impermeable barrier placed over the tissue.

[0014] In embodiments, the signal oligonucleotide is collected from a solution proximal to, e.g., at least immediately above, the at least one cell. The proximal solution may be collected by aspirating, e.g., via a pipette, a capillary tube, a microarray pin, a flow cell comprising holes, or another suitable aspirating system known in the art or any combination thereof. The capillary