PTO/AIA/15 (03-13)
Approved for use through 01/31/2014. 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.	38075-716.301
First Inventor	Emil Kartalov
Title	Signal Encoding and Decoding in Multiplexed Biochemical Assays
Express Mail Label No.	E-Filed August 5, 2014

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. ☒ **Specification** [*Total Pages* 78] Both the claims and abstract must start on a new page. *(See MPEP § 608.01(a) for information on the preferred arrangement)*
5. ☒ **Drawing(s).** (*35 U.S.C. 113*) [*Total Sheets* 08]
6. **Inventor's Oath or Declaration** [*Total Sheets* _____] *(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. ☒ **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. ☐ Computer Readable Form (CRF)
   b. ☐ Specification Sequence Listing on:
      i. ☐ CD-ROM or CD-R (2 copies); or
      ii. ☐ Paper
   c. ☐ Statements verifying identity of above copies

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

ACCOMPANYING APPLICATION PARTS

10. ☐ **Assignment Papers** (cover sheet & document(s))
    Name of Assignee _____
11. ☐ **37 CFR 3.73(c) Statement** ☐ **Power of Attorney** *(when there is an assignee)*
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. ☐ **Other:**

***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

☒ The address associated with Customer Number: 21971 **OR** ☐ Correspondence address below

Name					
Address					
City		State		Zip Code	
Country		Telephone		Email	

Signature	/Reena Bajpai/	Date	August 5, 2014
Name (Print/Type)	Reena Bajpai	Registration No. (Attorney/Agent)	69,996

This collection of information is required by 37 CFR 1.53(b). 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 12 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.

Electronic Acknowledgement Receipt	
EFS ID:	19779551
Application Number:	14451876
International Application Number:	
Confirmation Number:	1065
Title of Invention:	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS
First Named Inventor/Applicant Name:	Emil Kartalov
Customer Number:	21971
Filer:	Reena Bajpai/Mai Pham
Filer Authorized By:	Reena Bajpai
Attorney Docket Number:	38075-716.301
Receipt Date:	05-AUG-2014
Filing Date:	
Time Stamp:	15:32:28
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		38076716301_SPECIFICATION.pdf	460148 f5eb44cc5ee7bda67b4f720c40683a7c20eb9484	yes	78

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Specification	1	67
	Claims	68	77
	Abstract	78	78

Warnings:					
Information:					
2	Drawings-only black and white line drawings	38075716301_DRAWINGS.pdf	461550 7de515a902ceb8abb3604e0091fac7babce0ccc8	no	8
Warnings:					
Information:					
3	Application Data Sheet	38075716301_ADS.pdf	1565842 66511aa3c1c6ed94b26d2182b1d8cdd03e7c7c14	no	8
Warnings:					
Information:					
4	Transmittal of New Application	38075716301_TRANSMITTAL.pdf	215810 b20790e9d65a301c690772c5c2e1b1d631bd28ef	no	1
Warnings:					
Information:					
		Total Files Size (in bytes):	2703350		

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.

<u>New Applications Under 35 U.S.C. 111</u>
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.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
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.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
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.

PATENT APPLICATION

SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

Inventor(s): Emil P. Kartalov
Citizen of USA, Residing at
Pasadena, California

Aditya Rajagopal
Citizen of USA, Residing at
Irvine, California

Axel Scherer
Citizen of Germany, Residing at
Barnard, Vermont

Assignee: California Institute of Technology
1200 East California Boulevard
Pasadena, CA 91125

WSGR
Wilson Sonsini Goodrich & Rosati
PROFESSIONAL CORPORATION

650 Page Mill Road
Palo Alto, CA 94304
(650) 493-9300
(650) 493-6811

Electronically Filed on August 5, 2014

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 4

SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. provisional applications 61/594,480, filed February 3, 2012 and 61/703,093, filed September 19, 2012, each of which is incorporated herein by reference in its entirety.

BACKGROUND

[0002] Multiplexed reactions offer significant advantages over traditional uniplex reactions, including performance of parallel reactions on the same sample, use of the same chamber to perform multiple reactions, and the ability to extract rich information from a sample in a fast and efficient manner. However, to achieve these benefits, multiplexed assays generally require complex reporting mechanisms, namely spectrally resolved fluorescence or chemiluminescence (e.g., PCR, ELISA), spatially resolved signals (e.g., microarrays, gel electrophoresis), temporally resolved signals (e.g., capillary electrophoresis), or combinations thereof (e.g., Sanger sequencing). There is a need for multiplexed reactions that can be carried out in a single solution.

SUMMARY OF THE INVENTION

[0003] This disclosure provides methods, compositions, systems, and kits for the multiplexed detection of analytes. In some cases, this disclosure provides assays that are capable of unambiguously detecting the presence or absence of each of at least 7 analytes, in any combination of presence or absence, in a single sample volume without immobilization, separation, mass spectrometry, or melting curve analysis. In some examples, each of the analytes is encoded as a value of one (or at least one) component of a signal.

[0004] In some examples, an assay provided in this disclosure is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = \log_2 (F + 1)$ and F is the maximum cumulative value of one component of a signal (e.g., an intensity) when all the analytes are present.

[0005] In some cases, each analyte is encoded as at least one first value in a first component of a signal (e.g., at least one intensity or range of intensities) and at least one second value in a second component

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 5

of a signal (e.g., at least one wavelength or range of wavelengths). In some examples, each analyte is encoded as a first value in a first component of a signal at each of a plurality of second values in a second component of a signal (e.g., a signal intensity or range of signal intensities at each of a plurality of wavelengths or ranges of wavelengths).

[0006] In some examples, an assay provided herein is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = C * \log_2 (F + 1)$, where C is the number of the second values used to encode the analytes and F is the maximum cumulative value of the first component of the signal, for any second value, when all of the analytes are present.

[0007] In some cases, an assay provided in this disclosure is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = (P * T) + 1$, where P is the number of codes per tier in a coding scheme and T is the number of tiers. In some cases, the number of tiers $T = \log_4 (F + 1)$, and F is the maximum cumulative value of a first component of a signal, for any second value, when all of the analytes are present.

[0008] In some examples, the first value is an intensity or range of intensities. In some cases, a coding scheme comprises at least three intensities or ranges of intensities.

[0009] In some cases, the second value is a wavelength or range of wavelengths. In some examples, the coding scheme comprises at least five wavelengths or ranges of wavelengths.

[0010] In some examples, the signal is an electromagnetic signal. In some cases, the electromagnetic signal is a fluorescence emission signal. In some examples, the intensity of the fluorescence emission signal is measured at at least four wavelengths or ranges of wavelengths.

[0011] In some cases, an assay provided herein is performed with reagents that are lyophilized prior to use.

[0012] In some cases, a detecting step is performed with reagents comprising hybridization probes. In some examples, the number of the hybridization probes is greater than the number of analytes. In some cases, the set of hybridization probes comprise one or more hybridization probes specific for different analytes and comprising an identical fluorophore or combination of fluorophores. In some examples, a sample is contacted with at least 18 of said hybridization probes. In some examples, a detecting step is performed with reagents comprising at least one pair of primers. In some cases, at least one pair of primers are capable of amplifying a region complementary to at least three of said hybridization probes.

[0013] In some examples, the signal that is measured is generated during a polymerase chain reaction. In some cases, the polymerase chain reaction is selected from the group consisting of an end-point

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 6

polymerase chain reaction, a real-time polymerase chain reaction, a digital polymerase chain reaction, and combinations thereof.

[0014] In some examples, at least one cumulative measurement is performed on a solution.

[0015] In some examples, at least one analyte is encoded by at least one additional value (i.e., at least two values together), wherein the at least one additional value is selected from the group consisting of a value from at least one additional component of a signal, a value from at least one component of a different signal, and combinations thereof. For example, the at least one additional value may be selected from the group consisting of a fluorescence emission intensity, a fluorescence emission wavelength, a Förster resonance energy transfer (FRET) emission intensity, a FRET emission wavelength, an electrochemical signal, a chemiluminescence wavelength, a chemiluminescence intensity, a fluorescence bleaching rate, a chemiluminescence bleaching rate, and combinations thereof.

[0016] In some cases, a chromatogram is constructed. The chromatogram may be constructed by plotting all possible combinations of first values and second values for positive control samples for each analyte.

[0017] In some examples, at least one of the analytes comprises a polynucleotide. In some cases, the polynucleotide is from a source selected from the group consisting of an animal, a plant, a bacteria, a fungus, a parasite, and a virus. In some examples, the polynucleotide is from a source selected from the group consisting of human immunodeficiency virus, herpes simplex virus, human papilloma virus, Plasmodium, Mycobacterium, dengue virus, hepatitis virus, and influenza virus. In some cases, the polynucleotide is selected from the group consisting of human immunodeficiency virus polyprotease, human immunodeficiency virus p17, human papilloma virus E6, and human papilloma virus E7.

[0018] In some cases, a sample is selected from the group consisting of a clinical sample, a food sample, an environmental sample, a pharmaceutical sample, and a sample from a consumer product.

[0019] In some examples, information concerning the presence or absence of an analyte is transmitted through a computer network.

[0020] In some cases, information concerning the presence or absence of an analyte is provided to a physician. In some examples, a clinical decision is made based on such information.

[0021] In some examples, at least one step of a method provided herein is performed using instructions on a computer readable medium. In some cases, the instructions are located on a remote server. In some examples, the instructions are located on a thermal cycler. In some cases, the instructions are located on a computer in communication with a thermal cycler.

[0022] In some cases, at least one of the analytes is a positive control analyte.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 7

[0023] In some examples, the coding scheme is non-degenerate. In some cases, the coding scheme is designed to be non-degenerate. In some examples, the coding scheme is made non-degenerate by enumerating every legitimate result that can be obtained from the coding scheme, identifying each legitimate result that is degenerate, and eliminating at least one potential analyte code from the coding scheme to eliminate degeneracy.

[0024] In some cases, an assay provided herein is an end-point assay. In some examples an assay provided in this disclosure is ended at a threshold number of cycles set by the limit of detection of an instrument.

[0025] In certain examples, an assay provided herein is a liquid phase assay.

[0026] In some cases, an assay provided herein is quantitative.

[0027] In some cases, this disclosure provides a method of detecting the presence or absence of each analyte of a plurality of analytes, comprising: (a) encoding each of said analytes as a first value of a signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said first value, wherein said encoding is performed in a manner that eliminates degeneracy; (b) providing a sample comprising, or potentially comprising, at least one of said analytes; (c) contacting said sample with analyte-specific reagents that generate said first value, as represented in said coding scheme, when each of said analytes is present; (d) cumulatively measuring said first values of said signal within said sample, thereby providing a cumulative measurement; and (e) determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme.

[0028] In some examples, the method described in the preceding paragraph is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = \log_2 (F + 1)$ and F is the maximum cumulative value of the first values when all of the analytes are present. In some cases, the first value is an intensity or range of intensities. In some examples, the first value has a minimum value in the coding scheme that is selected from the group consisting of at least 1, at least 2, at least 4, at least 8, at least 16, at least 32, and at least 64. In some cases, the first value is incremented in the coding scheme by an amount equal to the cumulative maximum of said preceding first values plus one. In other cases, the first value is incremented in the coding scheme by an amount greater than the cumulative maximum of the preceding first values plus one.

[0029] In some cases, this disclosure provides a method of detecting the presence or absence of each analyte of a plurality of analytes, comprising: (a) encoding each of said analytes as at least one first value and at least one second value, wherein said first value is a value from a first component of a signal and said second value is a value from a second component of said signal, thereby generating a coding

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 8

scheme, wherein each of said analytes is represented in said coding scheme by said at least one first value and said at least one second value, wherein said encoding is performed in a manner that reduces or eliminates degeneracy; (b) providing a sample comprising, or potentially comprising, at least one of said analytes; (c) contacting said sample with analyte-specific reagents that generate said at least one first value and at least one said second value, as represented in said coding scheme, when each of said analytes is present; (d) cumulatively measuring said first values and said second values within said sample, thereby providing a cumulative measurement; and (e) determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme, wherein, when degeneracy is eliminated, said method is capable of unambiguously detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

[0030] In some cases, each of the analytes of the method described in the preceding paragraph is encoded as a first value in a first component of the signal at each of a plurality of second values in a second component of the signal. In some cases, when degeneracy is eliminated, this method is capable of unambiguously detecting the presence or absence of each of at least seven analytes in a single volume using four second values. In some examples, when degeneracy is eliminated, the coding scheme comprises T non-degenerate tiers, wherein $T = \log_4 (F + 1)$ and F is the maximum cumulative value of the first component of the signal, for any second value, when all of the analytes are present. In some cases, when degeneracy is eliminated, the method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = (P * T) + 1$ and P is the number of codes per tier. In some examples, when degeneracy is eliminated, the method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = C * \log_2 (F + 1)$, C is the number of second values in the coding scheme, and F is the maximum cumulative value of the first component of the signal, for any second value, when all of the analytes are present.

[0031] This disclosure also provides non-degenerate coding schemes capable of unambiguously encoding the presence or absence of each of at least 7 analytes, in any combination of presence or absence, in a single sample volume without immobilization, separation, mass spectrometry, or melting curve analysis.

[0032] In some examples, a non-degenerate coding scheme of this disclosure is generated by a method comprising: (a) generating a code for each potential analyte, wherein each potential analyte is encoded by at least one value of at least one component of a signal; (b) enumerating every legitimate cumulative

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 9

result for all possible combinations of presence or absence of each analyte; (c) identifying each legitimate result that is degenerate; and (d) eliminating at least one code to eliminate degeneracy. In some cases, the coding scheme may be expanded by adding additional codes that are at least one unit greater than the sum of all previous codes in the at least one component of said signal. In some cases a coding scheme is generated by a method of mathematical iteration that guarantees non-degeneracy by construction.

[0033] In some cases, this disclosure provides systems for detecting the presence or absence of each analyte of a plurality of analytes, comprising: (a) encoding each of said analytes as a first value of a signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said first value, wherein said encoding is performed in a manner that eliminates degeneracy; (b) providing a sample comprising, or potentially comprising, at least one of said analytes; (c) contacting said sample with analyte-specific reagents that generate said first value, as represented in said coding scheme, when each of said analytes is present; (d) cumulatively measuring said first values of said signal within said sample, thereby providing a cumulative measurement; and (e) determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme.

[0034] In some cases, this disclosure provides systems for detecting the presence or absence of each analyte of a plurality of analytes, comprising: (a) encoding each of said analytes as at least one first value and at least one second value, wherein said first value is a value from a first component of a signal and said second value is a value from a second component of said signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said at least one first value and said at least one second value, wherein said encoding is performed in a manner that reduces or eliminates degeneracy; (b) providing a sample comprising, or potentially comprising, at least one of said analytes; (c) contacting said sample with analyte-specific reagents that generate said at least one first value and said at least one second value, as represented in said coding scheme, when each of said analytes is present; (d) cumulatively measuring said first values and said second values within said sample, thereby providing a cumulative measurement; and (e) determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme, wherein, when degeneracy is eliminated, said system is capable of unambiguously detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

[0035] In some cases, this disclosure provides methods of detecting the presence or absence of each analyte of a plurality of analytes for a third party, comprising: (a) obtaining the identity of each of said

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 10

analytes from a party; (b) encoding each of said analytes as a first value of a signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said first value, wherein said encoding is performed in a manner that eliminates degeneracy; (c) providing said party with analyte-specific reagents that generate said first value, as represented in said coding scheme, when each of said analytes is present, said party: (i) contacting a sample comprising, or potentially comprising, at least one of said analytes with said reagents; and (ii) cumulatively measuring said first values within said sample, thereby providing a cumulative measurement; and (d) obtaining said cumulative measurement from said party; (e) determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme; and (f) providing said party with information about the presence or absence of each of said analytes.

[0036] In some cases, this disclosure provides methods of detecting the presence or absence of each analyte of a plurality of analytes for a third party, comprising: (a) obtaining the identity of each of said analytes from a party; (b) encoding each of said analytes as at least one first value and at least one second value, wherein said first value is a value from a first component of a signal and said second value is a value from a second component of said signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said at least one first value and said at least one second value, wherein said encoding is performed in a manner that reduces or eliminates degeneracy; (c) providing said party with analyte-specific reagents that generate said at least one first value and said at least one second value, as represented in said coding scheme, when each of said analytes is present, said party: (i) contacting a sample comprising, or potentially comprising, at least one of said analytes with said reagents; and (ii) cumulatively measuring said first values and said second values within said sample, thereby providing a cumulative measurement; and (d) obtaining said cumulative measurement from said party; (e) determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme; and (f) providing said party with information about the presence or absence of each of said analytes.

[0037] In some cases, this disclosure provides compositions for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, each reagent generating a signal comprising a first value, wherein said reagents are capable of unambiguously detecting the presence or absence of each of at least seven analytes in a single sample volume, in any combination of presence or absence, without immobilization, separation, mass spectrometry, or melting curve analysis.

[0038] In some cases, this disclosure provides compositions for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, each reagent generating a

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 11

signal comprising at least one first value that is a value from a first component of said signal and at least one second value that is a value from a second component of said signal, wherein said reagents are capable of detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

[0039] In some cases, this disclosure provides kits for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, packaging, and instructions, each reagent generating a signal comprising a first value, wherein said kit is capable of unambiguously detecting the presence or absence of each of at least seven analytes in a single sample volume, in any combination of presence or absence, without immobilization, separation, mass spectrometry, or melting curve analysis.

[0040] In some examples, this disclosure provides kits for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, packaging, and instructions, each reagent generating a signal comprising at least one first value that is a value from a first component of said signal and at least one second value that is a value from a second component of said signal, wherein said kit is capable of detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

[0041] In some cases, this disclosure provides kits for detecting the presence or absence of each analyte of a plurality of analytes, comprising a kit body (601), clamping slots arranged in the kit body for placing bottles, a bottle comprising probes for the detection of an analyte (602), a bottle comprising primers for amplification (603), and a bottle comprising reagents for amplification (604), wherein said kit is capable of unambiguously detecting the presence or absence of each of at least seven analytes in a single sample volume, in any combination of presence or absence, without immobilization, separation, mass spectrometry, or melting curve analysis.

[0042] In some examples, this disclosure provides kits for detecting the presence or absence of each analyte of a plurality of analytes, a kit body (601), clamping slots arranged in the kit body for placing bottles, a bottle comprising probes for the detection of an analyte (602), a bottle comprising primers for amplification (603), and a bottle comprising reagents for amplification (604), each probe generating a signal comprising at least one first value that is a value from a first component of said signal and at least one second value that is a value from a second component of said signal, wherein said kit is capable of detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said probes generates only one second value.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 12

INCORPORATION BY REFERENCE

[0043] 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

[0044] 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:

[0045] **Fig. 1** shows a comparison between a traditional encoding method of detecting four analytes with four colors and an encoding method of the invention able to detect 16 or more sequences with four colors. Colors are indicated by B, G, Y, and R, which indicate blue, green, yellow, and red, respectively.

[0046] **Fig. 2** shows a schematic representation of detection of six analytes (including control) with four colors. The analytes are detected using hybridization probes attached to a fluorophore (B, G, Y, R; blue, green, yellow, and red, respectively) and a quencher (oval with "X"). The control sequence and the five other sequences (Dengue Fever = dengue virus; Tuberculosis = *Mycobacterium tuberculosis*; P17 = HIV p17; Malaria = *Plasmodium falciparum;* and Herpes = herpes simplex virus 2) are all detected using a probe labeled with a blue fluorophore. The non-control analytes are each detected using 1-3 additional probes. For example, the dengue virus analyte is detected using three additional probes with green, yellow, and red fluorophores; the herpes simplex virus 2 analyte is detected using one additional probe with a red fluorophore; and so on.

[0047] **Fig. 3** shows chromatograms of experimental results obtained as described in Example 2.

[0048] **Fig. 4** shows an exemplary embodiment of the invention in which a computer is used to perform one or more steps of the methods provided herein.

[0049] **Fig. 5** shows a schematic of two different methods of detecting wavelength and intensity of a light signal, such as a fluorescence emission signal.

[0050] **Fig. 6** shows a schematic of exemplary kits.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 13

[0051] **Fig. 7** shows chromatograms of experimental results obtained as described in Example 3.

[0052] **Fig. 8** shows a schematic of a legitimate result (top) and an illegitimate result (bottom).

DETAILED DESCRIPTION OF THE INVENTION

[0053] Fluorescence detection has been a preferred technique for multiplexed assays because of several desirable features, including compatibility with biochemical assays, the relatively small size of fluorescent labels, simple means of conjugation to molecules of interest, affordability, low toxicity, stability, robustness, detectability with inexpensive optics, and an ability to be combined with spatial arrays. However, standard fluorophores have wide emission spectra. Therefore, in order to avoid spectral overlap (i.e., to preserve spectral resolution) only a relatively small number of colors (e.g., 4 to 6) are typically used simultaneously in multiplexed fluorescent assays.

[0054] The traditional encoding method for multiplexed fluorescent assays has been to encode each analyte with a single color; i.e., M=N, where M is the number of analytes that can be detected and N is the number of spectrally resolved fluorescent probes. Whenever higher factors of multiplexing are required (i.e., M>N), fluorescence is generally combined with other techniques, such as aliquoting, spatial arraying, or sequential processing. These additional processing steps are labor-intensive and frequently require relatively expensive and complex optical and mechanical systems (e.g., spectrometers, mechanized microscopy stages, microfluidics, droplet generators, scanners, and the like). Such systems are often impractical to deploy in certain settings, particularly point-of care and low-resource settings. Thus, there is a significant need for multiplexed encoding and decoding methods that can provide an inexpensive means of multiplexing while avoiding the use of expensive additional processing steps.

[0055] This disclosure provides methods, systems, compositions, and kits for the detection of multiple analytes in a sample. Analytes are detected based on the encoding, analysis, and decoding methods presented herein. In some examples, each analyte to be detected is encoded as a value of a signal (e.g., intensity), where the values are assigned so that the results of the assay unambiguously indicate the presence or absence of the analytes being assayed. In other examples, each analyte to be detected is encoded as a value in each of at least two components of a signal (e.g., intensity and wavelength). The at least two components of a signal may be orthogonal. Similarly, as described more fully elsewhere in this disclosure, multiple orthogonal signals may be used, such as a combination of a fluorescent signal and an electrochemical signal. The analyte may be any suitable analyte, such as a polynucleotide, a

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 14

protein, a small molecule, a lipid, a carbohydrate, or mixtures thereof. The signal may be any suitable signal such as an electromagnetic signal, a light signal, a fluorescence emission signal, an electrochemical signal, a chemiluminescent signal, and combinations thereof. The at least two components of the signal may be any suitable two components, such as an amplitude and a frequency or an intensity and a wavelength.

[0056] After encoding of the analytes, a sample is provided wherein the sample comprises or may comprise at least one of the encoded analytes. The sample is contacted with an analyte-specific reagent or reagents that generate a particular signal, as specified for each analyte in the coding scheme, in the presence of an analyte. A reagent may be any suitable reagent that is capable of generating such a signal in the presence of its corresponding analyte, for example, an oligonucleotide probe attached to a fluorophore and a quencher (e.g., a TAQMAN probe). If the reagent is an oligonucleotide probe attached to a fluorophore and a quencher, a nucleic acid amplification may be performed to generate the signal.

[0057] After addition of the reagent(s), the signal is quantified. In some cases, this quantification is performed by measuring one component of the signal (e.g., fluorescence intensity) and determining the presence and absence of certain analytes based on the values used to encode the presence of each analyte and the cumulative value of the signal.

[0058] In some cases, at least two components of a signal (e.g., intensity and wavelength) are cumulatively measured for the sample. This measurement can be performed, for example, by measuring the intensity at a particular wavelength or the intensity within a particular range of wavelengths. The presence or absence of an analyte may then be determined based on the values of each of the at least two components of the signal and the values used to encode the presence of the analyte (i.e., those values in the coding scheme).

[0059] The encoding may be performed in a manner that reduces or eliminates the number of possible degenerate (e.g., ambiguous) results that can be obtained by the method. As described elsewhere in this specification, the full coding capability of a particular coding scheme may be enumerated, and certain potential analyte codes may be eliminated from the coding scheme to reduce or eliminate degeneracy. Similarly, a coding scheme may be designed to be non-degenerate, so that a reduction or elimination of degeneracy is not necessary. A decoding matrix may be constructed to translate cumulative signal measurements (e.g., intensities or intensities at particular wavelengths) into the presence or absence of certain analytes, corresponding to the constituent signals of the cumulative signal measurement.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 15

I. Definitions

[0060] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

[0061] As used herein, the singular forms "a," "an," and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. Furthermore, to the extent that the terms "including," "includes," "having," "has," "with," "such as," or variants thereof, are used in either the specification and/or the claims, such terms are not limiting and are intended to be inclusive in a manner similar to the term "comprising".

[0062] The term "about," as used herein, generally refers to a range that is 15% greater than or less than a stated numerical value within the context of the particular usage. For example, "about 10" would include a range from 8.5 to 11.5.

[0063] The terms "dimension" and "component," as used herein when referring to a signal, generally refer to an aspect of the signal that may be quantified. For example, if a signal is generated by a fluorescent molecule, it may be quantified in terms of its wavelength (e.g., a first dimension or component) and its intensity (e.g., a second dimension or component).

[0064] The term "decoding," as used herein, generally refers to a method of determining which analytes are present based on the cumulative signal and a coding scheme or decoding matrix that enables the conversion of a cumulative signal to information concerning the presence or absence of one or more analytes.

[0065] The term "encoding," as used herein, generally refers to the process of representing an analyte using a code comprising values of a signal, such as intensity, or values in each of at least two components of a signal or signals, such as wavelength and intensity.

[0066] The terms "oligonucleotide probe attached to a fluorophore and a quencher" and "TAQMAN probe" generally refer to hydrolysis probes used to detect the presence of an analyte in a polynucleotide amplification assay. These probes comprise an oligonucleotide probe attached to a fluorophore and a quencher. So long as the quencher and the fluorophore are in proximity, the quencher quenches the fluorescence emitted by the fluorophore upon excitation by a light source. The sequence of the oligonucleotide probe is designed to be complementary to a polynucleotide sequence present in an analyte, and therefore capable of hybridizing to the polynucleotide sequence present in the analyte. Hybridization of the oligonucleotide probe is performed in a nucleic acid amplification reaction comprising primers (e.g., a polymerase chain reaction). Upon extension of the primers by a DNA polymerase, the 5' to 3' exonuclease activity of the polymerase degrades the probe, releasing the

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 16

fluorophore and the quencher into the medium. The proximity between the fluorophore and the quencher is broken and the signal from the fluorophore is no longer quenched. Thus, the amount of fluorescence detected is a function of the amount of analyte present. If no analyte is present, the probe will not hybridize to an analyte, and the fluorophore and quencher will remain in close proximity. Little or no signal will be produced.

[0067] The term "orthogonal," as used herein, generally refers to at least two components of a signal (e.g., wavelength and intensity), or at least two different signals (e.g., fluorescence emission and electrochemical signal), that can be varied independently or approximately independently. For example, wavelength and intensity are considered orthogonal or approximately orthogonal when fluorescent molecules are used. Among other factors, the wavelength of fluorescence emission will depend on the composition of the fluorescent molecule and the intensity of the fluorescence will depend on the amount of molecule present. Although wavelength and intensity are examples of two components of a signal that can be varied approximately independently, the methods described herein are not limited to components that can be varied independently or approximately independently. Components of a signal, or signals, that vary non-independently may also be used, so long as the components or signals are characterized well enough to enable encoding, measurement, and decoding. For example, if the variance in one component or signal affects the variance in another component or signal, the two components or signals may still be used so long as the relationship between the variances is understood.

[0068] The terms "polynucleotide," "oligonucleotide," or "nucleic acid," as used herein, are used herein to refer to biological molecules comprising a plurality of nucleotides. Exemplary polynucleotides include deoxyribonucleic acids, ribonucleic acids, and synthetic analogues thereof, including peptide nucleic acids.

[0069] The term "probe," as used herein, generally refers to a reagent capable of generating a signal in the presence of a particular analyte. A probe generally has at least two portions: a portion capable of specifically recognizing an analyte, or a portion thereof, and a portion capable of generating a signal in the presence of an analyte, or a portion thereof. A probe may be an oligonucleotide probe attached to a fluorophore and a quencher, as described above and elsewhere in this disclosure. A probe may also be any reagent that generates a signal in the presence of an analyte, such as an antibody that detects an analyte, with a fluorescent label that emits or is quenched upon binding of the antibody to an analyte. Any suitable probe may be used with the methods presented in this disclosure, so long as the probe generates a quantifiable signal in the presence of an analyte. For example, the analyte-specific portion of a probe may be coupled to an enzyme that, in the presence of an analyte, converts an uncharged

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 17

substrate into a charged product, thereby increasing the electrical conductivity in the medium over time. In this case, different analytes may be encoded by coupling the analyte-specific portion of the probe (e.g., hybridization probe or antibody) to an enzyme at different ratios. The resulting rate of increased conductivity in the medium will be cumulative for all analytes present in the medium. Encoding analytes according to the methods provided herein enables conversion of the conductivity measurements into unambiguous (i.e., non-degenerate) results providing information about the presence or absence of particular analytes. Similarly, a probe may comprise an enzyme producing a chemiluminescent product from a substrate. The amount of chemiluminescence may then be used to encode the presence of particular analytes.

II. Encoding and Decoding Methods

A. Traditional Fluorescent Encoding and Decoding Method

[0070] A commonly used method of determining the presence of an analyte uses four spectrally resolved fluorescent molecules to indicate the presence or absence of four analytes. An example of this method is presented on the left-hand side of **Fig. 1**. The left-hand side of **Fig. 1** shows an encoding method where four analytes (Seq 1, Seq 2, Seq 3, and Seq 4) are each encoded by a single color (blue, green, yellow, and red, respectively). The color represents a fluorophore attached to an oligonucleotide probe that also comprises a quencher. In the system shown on the left-hand side of **Fig. 1**, there are four different oligonucleotide probes, each comprising a single fluorophore (blue, green, yellow, or red) and a quencher. The presence or absence of an analyte is determined based on the presence or absence of a signal in a particular color.

[0071] The chart on the left-hand side of **Fig. 1** shows intensity versus color for a hypothetical sample containing two analytes: Seq 1 and Seq 3. The presence of these analytes is determined based on the measurement of a blue signal (corresponding to Seq 1) and a yellow signal (corresponding to Seq 3). The absence of Seq 2 and Seq 4 is indicated by the absence of a blue and red signal.

[0072] Table 1 shows a translation of this coding scheme into a binary format. Each analyte is encoded as a value in each of two components of a fluorescent signal: (1) color (also known as wavelength; or range of wavelengths) and (2) intensity (indicated by the numbers within the table). For example, A (e.g., Seq 1) has a color of blue and an intensity of 1; B (e.g., Seq 2) has a color of green and an intensity of 1; and so on. The intensity of the signal within each color range may be quantified as described herein, for example by measuring the intensity of the signal within a particular wavelength range

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 18

determined by a band pass filter. A result of 1000 indicates that only analyte A is present; a result of 1100 indicates that analytes A and B are present; and so on.

Table 1. Traditional encoding of four analytes with four probes, each probe having a single color, and one probe per analyte.

	Blue	Green	Yellow	Red
A	1	0	0	0
B	0	1	0	0
C	0	0	1	0
D	0	0	0	1

B. Encoding Methods Using More Than One Color Per Analyte

[0073] The traditional method described above suffers from the fact that it is limited by the number of spectrally resolvable fluorophores. More specifically, the number of detectable analytes is equal to the number of spectrally resolvable fluorophores. Therefore, the number of analytes may only be increased by increasing the number of spectrally resolvable fluorophores.

[0074] This disclosure provides methods that overcome this limitation. More specifically, in some cases, by utilizing at least two components of a signal during encoding, the methods described herein may be used to detect more than one analyte per fluorophore. For example, using the method provided herein 1.5, 1.6, 1.7, 1.8, 1.9, 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, 49, or 50 analytes may be detected per fluorophore. In some cases, the methods provided herein may be used to detect at least 1.5, 1.6, 1.7, 1.8, 1.9, 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, 49, or 50 analytes may be detected per fluorophore. In some cases, the methods provided herein may be used to detect 1.5-2, 2-4, 1.5-4, 4-6, 2-6, 6-10 analytes per fluorophore.

[0075] In some cases, the methods provided in this disclosure may include the use of a control color. The control color may be attached to one or more probes binding a positive control analyte, and each analyte to be detected, in a sample. If the same sequence occurs in the positive control analyte and each analyte to be detected, a single control probe may be used. If the same sequence does not occur in the positive control analyte and each analyte to be detected, different probes may be used, but each probe may still be attached to the control color.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 19

[0076] For example, building on the traditional methods described above, one color (e.g., blue) may be used to encode the presence of a control analyte that is always present in the sample. The control analyte may be added to the sample, or may be inherently present in the sample. The remaining colors (e.g., green, yellow, red) may be used to encode the presence of additional analytes. Table 2 shows an example of one such method.

Table 2. Encoding of four analytes, including control, with four colors and up to two colors per analyte.

	Blue	Green	Yellow	Red
Ctrl. (p)	1	0	0	0
A	1	0	0	1
B	1	0	1	0
C	1	1	0	0

[0077] In Table 2, the presence of the control (p) is indicated by a result of 1000. The presence of the control and analyte A is indicated by a result of 2001. The presence of the control and all three other analytes is indicated by a result of 4111. The control color (blue) provides an indication that the assay is functioning properly. The intensity of the blue color reports the number of analytes that are present in a test sample. Of course, any color may be used as the control color. One of skill in the art will recognize that certain practical considerations might make it preferable to use one color over another as the control color. For example, if one color is better detected in a particular optical system (or system of fluorophores), it might be practically preferable to use that color as a control color.

[0078] The coding scheme shown in Table 2 encodes each analyte using one control color and one additional color (up to 2 probes per analyte). However, as shown below, the number of analytes that can be encoded increases when up to 4 colors are used per analyte. Table 3 shows an exemplary coding scheme where each analyte is encoded by up to 4 colors. In the scheme shown in Table 3, the control color (blue) is generated in the presence of the control sequence and each of the other seven analytes (A-G). The other seven analytes are each encoded by the presence of one to three additional colors. The colors may be contained on different probes (e.g., oligonucleotide probes attached to a fluorophore and a quencher) or on the same probe (which can have multiple fluorophores and multiple quenchers).

Table 3. Encoding of eight analytes (including control) with four colors and up to four colors per analyte.

	Blue	Green	Yellow	Red
Ctrl. (p)	1	0	0	0

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 20

A	1	0	0	1
B	1	0	1	0
C	1	1	0	0
D	1	0	1	1
E	1	1	0	1
F	1	1	1	0
G	1	1	1	1

[0079] In Table 3, the presence of the control and all seven other analytes is indicated by a result of 8444. Three analytes (A, B, and C) are encoded by two colors. Three analytes (D, E, and F) are encoded by three colors, and one analyte (G) is encoded by four colors. The presence of the control (p) is indicated by a result of 1000. The presence of the control and analyte A is indicated by a result of 2001. The presence of the control and analyte G is indicated by a result of 2111. These results express the cumulative intensity of the signal in each color. For example, the result 2111 has a 2X signal intensity in the blue channel, while the result 1000 has only a 1X signal intensity in the blue channel.

[0080] Using Table 3, each of the possible cumulative assay results can be enumerated, in terms of color (blue, green, yellow, red) and intensity (0-8). Table 4 shows a "decoding matrix" generated by enumerating each of the possible cumulative assay results based on the encoding method presented in Table 3 and providing the corresponding decoded result of each assay in terms of the analyte(s) present in the sample. A similar decoding matrix may be generated for any coding scheme described herein, by enumerating each of the possible cumulative assay results based on the coding scheme and providing the corresponding decoded result of each assay in terms of the analyte(s) present in (or absent from) the sample. In some cases, as described below, one or more analytes may be removed from the coding scheme in order to reduce or eliminate degeneracy.

Table 4. Decoding matrix for encoding method presented in Table 3.

Cumulative Assay Result				
Blue	**Green**	**Yellow**	**Red**	**Analyte(s) Present**
1	0	0	0	p
2	0	0	1	pA
2	0	1	0	pB
2	1	0	0	pC
2	0	1	1	pD
2	1	0	1	pE
2	1	1	0	pF
2	1	1	1	pG
3	0	1	1	pAB

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 21

3	1	0	1	pAC
3	0	1	2	pAD
3	1	0	2	pAE
3	1	1	1	pAF, pCD, pBE
3	1	1	2	pAG, pDE
3	1	1	0	pBC
3	0	2	1	pBD
3	1	2	0	pBF
3	1	2	1	pBG, pDF
3	2	0	1	pCE
3	2	1	0	pCF
3	2	1	1	pCG, pEF
3	1	2	2	pDG
3	2	1	2	pEG
3	2	2	1	pFG
4	1	1	1	pACB
4	0	2	2	pABD
4	2	2	0	pBCF
4	2	0	2	pACE
4	1	1	2	pABE, pACD
4	1	2	1	pABF, pBCD
4	2	1	1	pACF, pBCE
4	1	2	2	pADF, pABG, pBDE
4	2	1	2	pACG, pCDE, pAEF
4	2	2	1	pBCG, pBEF, pCDF
4	2	2	2	pCDG, pAFG, pBEG, pDEF
4	1	1	3	pADE
4	1	3	1	pBDF
4	3	1	1	pCEF
4	1	2	3	pADG
4	3	2	1	pCFG
4	2	1	3	pAEG
4	2	3	1	pBFG
4	1	3	2	pBDG
4	3	1	2	pCEG
4	2	2	3	pDEG
4	3	2	2	pEFG
5	1	2	2	pABCD
5	2	1	2	pABCE
5	2	2	1	pABCF
5	2	2	2	pABCG, pABEF, pBCDE, pACDF
5	1	2	3	pABDE
5	2	1	3	pACDE
5	1	3	2	pABDF
5	2	1	3	pBCDF
5	3	1	2	pACEF

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 22

5	3	2	1	pBCEF
5	2	2	3	pABEG, pACDG, pADEF
5	2	3	2	pABFG, pBCDG, pBDEF
5	3	2	2	pACFG, pBCEG, pCDEF
5	1	3	3	pABDG
5	3	1	3	pACEG
5	3	3	1	pBCFG
5	2	3	3	pADFG, pBDEG
5	3	3	2	pBEFG, pDEFG
5	3	2	3	pCDEG, pAEFG
5	2	2	4	pADEG
5	2	4	2	pBDFG
5	4	2	2	pCEFG
5	3	3	3	pDEFG
6	2	2	3	pABCDE
6	2	3	2	pABCDF
6	3	2	2	pABCEF
6	2	3	3	pABCDG, pABDEF
6	3	2	3	pABCEG, pACDEF
6	3	3	2	pABCFG, pBCDEF
6	4	3	2	pBCEFG
6	4	2	3	pACEFG
6	3	3	3	pABEFG, pACDFG, pBCDEG
6	3	4	2	pBCDFG
6	2	4	3	pABDFG
6	3	2	4	pACDEG
6	2	3	4	pABDEG
6	4	3	3	pCDEFG
6	3	4	3	pBDEFG
6	3	3	4	pADEFG
7	4	4	3	pBCDEFG
7	4	3	4	pACDEFG
7	3	4	4	pABDEFG
7	4	3	3	pABCEFG
7	3	4	3	pABCDFG
7	3	3	4	pABCDEG
7	3	3	3	pABCDEF
8	4	4	4	pABCDEFG

[0081] The decoding matrix provided in Table 4 is constructed using two assumptions. First, the decoding matrix assumes that the positive control (p) always produces a positive outcome. Second, the decoding matrix assumes that, within each color, the intensity is additive and scales in the same way with changing probe concentration, regardless of which probe the signal may come from. This

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 23

essentially means that the signals are additive and digital. The conditions underlying these assumptions may be met by properly preparing the assay. If a fluorescent signal is used, the intensity need only be approximately additive and digital, as demonstrated in the Examples provided herein.

[0082] Table 4 allows the conversion of a cumulative measurement of intensity in four ranges of fluorescent wavelengths (i.e., signal intensity within each color range), into the corresponding analytes present in the sample. For example, a result of 4321 indicates that p, C, F, and G are present and the other analytes are not present. This result is referred to as a "legitimate" result, because it is present in the decoding matrix. By contrast, a result of 4000, while possible to measure, does not occur in the decoding matrix. More specifically, the result of 4000 cannot be achieved by adding any combination of control and analyte codes from Table 3. This result is referred to as an "illegitimate" result. An illegitimate result may indicate that the assay malfunctioned. Thus, the control (p) and the decoding matrix provide a means of verifying that the assay is functioning properly.

[0083] Table 4 is exhaustive for any combination of fluorophores generating four resolvable emission spectra (e.g., colors). The term "rank" is used to describe the number of detected analytes, including the control. In the example provided above, if the assay functions properly, the rank is equal to the value of the blue signal. For example, a rank of 8 indicates that the control and all seven other analytes (A-G) are present. A rank of 2 indicates that the control and only one analyte are present. The lowest rank is a rank of 1, which indicates that only the control is present. The number of possibilities at each rank can be enumerated. For example, continuing to refer to the encoding and decoding method described in Tables 3-4, there are 7 possibilities for rank 2. The number of possibilities at rank 3 can be calculated as a combination of 7 take 2, or $7!/(5! * 2!) = 21$ possibilities. More generally, the number of possibilities for a combination of N take K is $N!/((N-K)! * K!)$. Analogously, at ranks 4, 5, 6, 7, and 8, the number of possibilities is 35, 35, 21, 7, and 1, respectively. Referring to Table 4 shows that the table agrees with the theoretical prediction. Thus, Table 4 is an exhaustive decoding matrix for the encoding method provided in Table 3.

C. Reducing or Eliminating Degeneracy

[0084] The terms "degenerate" and "degeneracy," as used herein, generally describe a situation where a legitimate result is not definitive, because it can indicate more than one possibility in terms of the presence or absence of an analyte. For example, with reference to Table 4, result 5233 is degenerate because it can be decoded as either pADFG or pBDEG. Similarly, result 4222 is degenerate because it

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 24

can be decoded as any of the following: pCDG, pAFG, pBEG, pDEF. By contrast, result 3110 can only indicate pBC and thus is not degenerate.

[0085] This disclosure provides methods of reducing or eliminating degeneracy, thereby increasing the confidence with which an analyte is detected. In one embodiment, degeneracy is eliminated by a method comprising (i) encoding each potential analyte to be detected as a value of a signal and, optionally, as a value in each of at least two components of a signal; (ii) enumerating every legitimate result that can be obtained from the coding scheme; (iii) identifying each legitimate result that is degenerate; and (iv) eliminating at least one potential analyte (or potential analyte code) from the coding scheme, wherein eliminating the at least one potential analyte reduces or eliminates degeneracy. For example, with reference to the coding scheme described in Table 3 and the decoding matrix described in Table 4 (enumerating every legitimate result), eliminating any two of analytes D, E, and F eliminates the degeneracy. Eliminating any one of analytes D, E, and F would not eliminate the degeneracy, but would reduce it.

[0086] With continued reference to the coding scheme described in Table 3, eliminating any two of analytes D, E, and F from the coding scheme results in a scheme where six analytes (including control) can be analyzed, with no degeneracy, using only 4 colors. By contrast, conventional methods of multiplexing would allow for only the reporting of 3 analytes and 1 control using 4 colors. Therefore, the number of analytes that can be analyzed is nearly doubled by using the methods provided herein.

[0087] **Fig. 2** shows an exemplary embodiment of the invention in which 4 colors are used to detect five analytes and a control. The analytes are nucleic acids from dengue virus ("Dengue Fever"), *Mycobacterium tuberculosis* ("Tuberculosis"), human immunodeficiency virus (HIV) p17 (P17), *Plasmodium* ("Malaria"), and herpes simplex ("Herpes"). The "colors" in this example are fluorophores attached to oligonucleotide probes that also comprise a quencher. The oligonucleotide probes will generally be different for different analytes, whereas the color is the same. For example, all probes designated as Probe 1 have a blue color but generally will have a different oligonucleotide sequence. Of course, probes designated as Probe 1 could also have the same oligonucleotides sequence, if the complementary sequence was present in each of the analytes. The mechanism of detection with these probes is described elsewhere in this disclosure. Probe 1 (blue) hybridizes to all six analytes, including control. Probe 2 (green) hybridizes to analytes from dengue virus, *Mycobacterium*, and HIV p17. Probe 3 (yellow) hybridizes to analytes from dengue fever and *Plasmodium*. Probe 4 hybridizes to analytes from dengue virus, *Mycobacterium*, and herpes simplex. The coding scheme defined by these probes is illustrated in Table 5. With reference to Table 3, potential analytes D (1011) and F (1110) have been

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 25

eliminated from the coding scheme. Therefore, the coding scheme presented in Table 5 is non-degenerate and each legitimate result from the assay corresponds to the presence or absence of a unique combination of analytes in a sample.

Table 5. Coding scheme for detection of dengue virus, *Mycobacterium*, HIV, *Plasmodium*, and herpes simplex, as exemplified in Fig. 2.

	Blue	Green	Yellow	Red
Ctrl. (p)	1	0	0	0
Herpes Simplex	1	0	0	1
Plasmodium	1	0	1	0
HIV	1	1	0	0
Mycobacterium	1	1	0	1
Dengue Virus	1	1	1	1

[0088] The methods for encoding and decoding presented above, including the methods for eliminating degeneracy, are all equally applicable to encoding methods using additional colors. For example, Table 3 could be extended by including additional colors and additional intensities (described further below). The decoding matrix is then generated as described above, enumerating every legitimate result that can be obtained from the coding scheme. A decoding matrix analogous to the decoding matrix provided in Table 4 may be constructed for any coding scheme described in this disclosure. The legitimate results that are degenerate are then identified and at least one potential analyte code is eliminated from the coding scheme to reduce or eliminate degeneracy. The method of eliminating degeneracy may be carried out using software on a computer readable medium, or hardware configured to carry out the method (e.g., a microchip).

[0089] Although degeneracy can be reduced or eliminated by the methods described above, and elsewhere in this disclosure, this disclosure also provides coding schemes that are non-degenerate by design. For example, after elimination of degeneracy in the coding scheme described in Table 3, the coding scheme may be extended indefinitely in a non-degenerate manner where the non-degeneracy is by design (see, e.g., the coding scheme exemplified in Table 6, described more fully below). Similarly, this disclosure provides coding schemes that are completely non-degenerate by design and therefore do not require any reduction or elimination of degeneracy (see, e.g., the coding scheme exemplified in

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 26

Table 8, described more fully below). Thus, the coding schemes provided in this disclosure may have reduced or eliminated degeneracy, or be non-degenerate by design.

D. Encoding Methods Using More Than One Color and More Than One Intensity

[0090] The encoding method described above (e.g., Tables 3-5) may be further extended by allowing analytes to be encoded by an intensity greater than 1. For example, each of the analytes encoded in the coding scheme provided in Table 5 is encoded by a fluorescence intensity of either 1 or 0. Allowing higher values for the signal intensity in at least one color further increases the number of analytes that can be encoded by any of the methods provided in this disclosure. In some examples, these higher intensity values may be assigned any color except for the control color, in order to maintain the analyte counting capability of the control color.

[0091] Table 6 shows the first three tiers of an exemplary coding scheme that utilizes four colors and multiple intensities. Tier 1 of Table 6 is a reproduction of Table 3, showing the encoding of seven analytes and a control with four colors. As described above, any two of potential analytes D, E, and F may be eliminated from the coding scheme in order to produce a non-degenerate coding scheme. Tier 1 of Table 6 indicates that potential analytes D (1011) and F (1110) have been eliminated from the coding scheme to eliminate degeneracy. Therefore, the coding scheme presented in Tier 1 of Table 6 is capable of determining the presence of five analytes and one control using four colors, as described above.

[0092] The coding scheme of Tier 1 of Table 6 may be expanded to a second tier (Tier 2) by allowing the intensity in any of the colors to increase. As described above, the intensity of the control color may be maintained at 1, in order to preserve the sequence counting capability of the control. Increasing the intensity of any of the remaining three colors will yield codes 100Y, 10Y0, 1Y00, 10YY, 1Y0Y, 1YY0, and 1YYY, where $Y > 1$. The minimal value of Y for a new tier of encoding is equal to the cumulative maximum value from the prior tier(s) plus 1. As described below, a value greater than 1 could also be used, to maximize the differences between the intensities.

[0093] Thus, in this context, the term "tier" is generally used to describe a set of codes that fully utilize the coding capability provided by a particular number of first values (e.g., intensities) and second values (e.g., colors), without degeneracy. For example, Tier 1 of Table 6 fully utilizes the coding capability provided by four colors with up to one intensity in each color, without degeneracy. As shown in Table 6, Tier 1, this results in six encoded analytes, including the control. To introduce a second tier, the minimum value of Y (described above) may be incremented to equal the cumulative maximum result from the prior tier(s) plus one (or more than one). In the example provided in Table 6, the intensity of

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 27

the blue (control) color is maintained as one, to preserve the sequence counting capability in this color. Thus, Tier 2 consists of five non-degenerate encoding possibilities obtained by incrementing the intensities in the green, yellow, and red channels to equal the cumulative maximum results in each of these channels from Tier 1, plus one. All possibilities of these codes may then be enumerated for Tier 2, and codes resulting in degeneracy may be eliminated, or Tier 2 may be made non-degenerate by design, using the information used to eliminate the corresponding codes from Tier 1. Further coding capacity may then be achieved by adding a third tier, or further tiers, which are constructed according to analogous methods. A coding scheme may have an infinite number of tiers, for example 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100 or more tiers.

[0094] More specifically, with reference to Table 6, Tier 2, analyte H is encoded by 1004. The value of the control color is maintained as 1. The value of the red color is equal to the cumulative maximum result from the prior tier (3) plus 1, or 4. Similarly, analytes I and J are encoded by 1030 and 1400, respectively. Combinations of these codes are used to encode the remaining four analytes (K-N), as was done for analytes D-G in Tier 1. This completes Tier 2. In Tier 2, the inclusion of potential analytes K (1034) and M (1430) results in degeneracy. Therefore, these analytes have been eliminated from the coding scheme, to eliminate degeneracy.

[0095] Continuing to refer to Table 6, a third tier (Tier 3) is constructed using the same principles described above. Analyte O is encoded by 1-0-0-16. The value of the control color is still maintained as 1. The value of the red color is equal to the cumulative maximum result from the prior tiers (15) plus 1, or 16. Similarly, analytes P and Q are encoded by 1-0-9-0 and 1-16-0-0, respectively. Combinations of these codes are used to encode the remaining four analytes (R-U), as was done for analytes D-G in Tier 1 and analytes K-N in Tier 2. In Tier 3, the inclusion of potential analytes R (1-0-9-16) and T (1-16-9-0) results in degeneracy. Therefore, these analytes have been eliminated from the coding scheme, to eliminate degeneracy.

[0096] The three-tier coding scheme shown in Table 6 shows the encoding of 15 analytes and one control using four colors. This coding scheme may be indefinitely extended, by adding more intensities to generate additional tiers and/or adding more colors, to generate additional coding capability within the tiers. The methods of reducing or eliminating degeneracy, as described in this disclosure, may be used with the addition of each intensity and/or color, to reduce or eliminate degenerate results.

[0097] More generally, the coding scheme depicted in Table 6 is a non-degenerate, infinite extension of the coding scheme depicted in Tables 3-5. The maximum intensity of the cumulative measurement at

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 28

the first tier is 6. The maximum intensity of the cumulative measurement at the second tier is 15. The maximum intensity of the cumulative measurement at the third tier is 63, and so on. Given a maximum cumulative intensity value (F), the maximum number of tiers (T) available in this coding scheme is $T = \log_4 (F + 1)$. The coding scheme depicted in Table 6 provides five non-degenerate codes per tier (P). Thus, the maximum number of codes $M = 5 * \log_4 (F + 1)$, or $M = P * T$, where P is the number of non-degenerate codes per tier and T is the number of tiers. For example, given $F = 63$, the maximum number of codes (i.e., analytes) is 15 for 5 non-degenerate codes per tier. This formula does not include the 1000 code, which is reserved for a positive control in Table 6. To include the control in the total number of analytes, one would simply add one, to provide the formula $M = (P * T) + 1$.

[0098] The methods provided in this disclosure may be used to expand this coding scheme infinitely. For example, by utilizing combinations of different intensities (i.e., first values) and colors (i.e., second values) one can encode any number of analytes (M) by varying the number of non-degenerate codes per tier (P) and the number of tiers (T). For example, the number of non-degenerate codes per tier is 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, or more. Similarly, the number of tiers may be 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, or more.

Table 6. Encoding of 15 analytes and one control using four colors and multiple intensities.

Tier	Analyte	B	G	Y	R	Comments
1	**Ctrl. (p)**	1	0	0	0	
	A	1	0	0	1	
	B	1	0	1	0	
	C	1	1	0	0	
	D	1	0	1	1	Eliminated from coding scheme to eliminate degeneracy.
	E	1	1	0	1	
	F	1	1	1	0	Eliminated from coding scheme to eliminate degeneracy.
	G	1	1	1	1	
Cumulative Maximum Result		*6*	*3*	*2*	*3*	
2	**H**	1	0	0	4	
	I	1	0	3	0	

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 29

	J	1	4	0	0	
	K	1	0	3	4	Eliminated from coding scheme to eliminate degeneracy.
	L	1	4	0	4	
	M	1	4	3	0	Eliminated from coding scheme to eliminate degeneracy.
	N	1	4	3	4	
Cumulative Maximum Result		*11*	*15*	*8*	*15*	
3	**O**	1	0	0	16	
	P	1	0	9	0	
	Q	1	16	0	0	
	R	1	0	9	16	Eliminated from coding scheme to eliminate degeneracy.
	S	1	16	0	16	
	T	1	16	9	0	Eliminated from coding scheme to eliminate degeneracy.
	U	1	16	9	16	
Cumulative Maximum Result						
Cumulative Maximum Result for Three Non-Degenerate Tiers		***16***	***63***	***26***	***63***	

[0099] The right-hand side of **Fig. 1** shows one exemplary embodiment of the method described above. More specifically, the right-hand side of **Fig. 1** shows an encoding method where nine or more analytes (Seq 1 - Seq 9, etc.) are each encoded by at least two colors, with varying intensities within each color. The color represents a fluorophore attached to an oligonucleotide probe that also comprises a quencher. The system can be designed so that each probe is labeled with single fluorophore, or each probe is labeled with more than one fluorophore. For example, the code for analyte H, in Table 6, is 1004. The intensity of 4 in the red channel may be achieved by either using an H-specific probe comprising 4 red fluorophores, or by using 4 H-specific probes each comprising a single red fluorophore. Of course, any combination of probes and fluorophores producing a result of 4 in the red channel would be equally appropriate, such as 2 probes with 2 red fluorophores each, and 1 probe with 1 red fluorophore and 1

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 30

probe with 3 red fluorophores, or simply one probe with one red fluorophore but present at 4x amount in the reaction mixture.

[00100] The coding scheme depicted on the right-hand side of **Fig. 1** may be represented as in Table 7. The result of the analysis shown on the right-hand side of **Fig. 1** is 4112, or an intensity of 4 in the blue channel, 1 in the green channel, 1 in the yellow channel, and 2 in the red channel. Using a decoding matrix constructed as described herein, this result is decoded to indicate the presence of Seq 1, Seq 3, Seq 4, and Seq 5.

Table 7. Coding scheme illustrated on the right-hand side of Fig. 1.

Tier	Analyte	B	G	Y	R	Comments
1	**Seq 1**	1	0	0	0	
	Seq 2	1	1	0	0	
	Seq 3	1	0	1	0	
	Seq 4	1	0	0	1	
	Seq 5	1	1	0	1	
	Seq 6	1	1	1	1	
	~~Eliminated Potential Seq~~	~~1~~	~~0~~	~~1~~	~~1~~	Eliminated from coding scheme to eliminate degeneracy.
	~~Eliminated Potential Seq~~	~~1~~	~~1~~	~~1~~	~~0~~	Eliminated from coding scheme to eliminate degeneracy.
Cumulative Maximum Result		*6*	*3*	*2*	*3*	
2	**Seq 7**	1	4	0	0	
	Seq 8	1	0	3	0	
	Seq 9	1	0	0	4	
Etc. - i.e. Continue as provided in Table 6.						

E. Encoding Methods Using One Color and One Intensity Per Analyte but Different Intensities Among Analytes

[00101] In some methods provided herein, each analyte is encoded by a single color and intensity combination. For example, in a four color system, the first four analytes may be encoded by 1000, 2000, 4000, and 8000. The next four analytes may be encoded by 0100, 0200, 0400, and 0800. Analytes 9-12 and 13-16 would be assigned analogously, as shown in Table 8.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 31

[00102] Like the encoding method described in Table 6, this coding scheme is theoretically infinite, non-degenerate by construction, and only limited by the bandwidth of the instrument used to measure the signal. However, this coding scheme enables more analytes to be quantified per unit of bandwidth than the encoding method described in Table 6. The reason is that the available multiplicity of signal is used with maximal efficiency, as each level of intensity is utilized in the coding (i.e. there are no gaps in the chromatogram; see below for description of chromatograms). Table 8 shows one embodiment of this method, illustrating four tiers of encoding based on four colors and intensities 1, 2, 4, and 8. The coding scheme is non-degenerate, and the result if all 16 analytes are present is 15-15-15-15. This encoding method is more efficient, in terms of bandwidth utilization, than the encoding methods presented above. However, this method does not have the proofreading capability of the first scheme, as all the results decode to legitimate outcomes in the absence of gaps in the chromatogram.

Table 8. Example of encoding method using one color and one intensity per analyte, but different intensities among analytes.

Tier	Analyte	B	G	Y	R
1	**A**	1	0	0	0
	B	2	0	0	0
	C	4	0	0	0
	D	8	0	0	0
2	**E**	0	1	0	0
	F	0	2	0	0
	G	0	4	0	0
	H	0	8	0	0
3	**I**	0	0	1	0
	J	0	0	2	0
	K	0	0	4	0
	L	0	0	8	0
4	**M**	0	0	0	1
	N	0	0	0	2
	O	0	0	0	4
	P	0	0	0	8
		15	***15***	***15***	***15***

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 32

[00103] The method presented above, and exemplified in Table 8, may be extended by introducing additional colors and/or intensities. For example, 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, 49, 50, 60, 70, 80, 90, 100, or more colors may be used. Additional intensity levels, such as 16, 32, 64, 128, 255, 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536, and so on may also be used. The method is generalizable and may be used to encode an infinite number of analytes. Each analyte is represented by a code in a single color, wherein the value of the code in that single color equal to the sum of all previous values plus one. For example, if the first code contains a 1 in a particular color, the next codes are 2, 4, 8, 16, 32, 64, 128 and so on. Other progressions are possible, but will not be as efficient in terms of usage of the bandwidth. However, one of ordinary skill in the art will recognize that less efficient use of bandwidth may be desirable when one wishes to maximize separation between values. Similarly, the value 1 could be excluded from the coding scheme, for example to maximize the difference in the intensity between the first encoded value and instrumental noise. For example, in some cases a coding scheme may begin with a value of 2, which would provide a progression of 2, 4, 8, 16, 32, 64, 128, and so on for the analyte codes, while the progression of cumulative results would be 2, 4, 6, 8, 10, 12, 14, 16, and so on. This approach generates uniform gaps in the possible cumulative results, allowing for higher tolerance to noise in comparison to an analyte code progression of 1, 2, 4, 8, 16, 32, and so on, and its corresponding cumulative result progression of 1, 2, 3, 4, 5, 6, 7, 8, and so on. In this example, the increased robustness in the measurement comes at a cost of a decrease in the number of available codes within a fixed bandwidth of multiplicity. For example, if only 63 multiplicity states can be measured reliably, the coding scheme starting with 1 offers 7 codes per color whereas the coding scheme starting with 2 offers 6 codes per color. If 4 colors are available, the former will offer 28 codes, while the latter will offer 24 codes. In summary, both the starting signal intensity and the progression (i.e., difference between intensity values) may be scaled in order to maximize the ability to distinguish over instrumental noise and maximize the differences between the intensity values, thereby enhancing the ability to distinguish between distinct experimental outcomes.

[00104] The coding scheme illustrated in Table 8 is non-degenerate by design. Although the coding scheme illustrated in Table 8 uses both intensity and color to encode each of the 16 analytes, each of the analytes could also be encoded by simply utilizing intensity. For example, given the 16 analytes provided in Table 8 (A-P), a single color coding scheme encoding all 16 analytes may assign the values 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, 8192, and 16384 to analytes A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, and P, respectively. Such a coding scheme is non-degenerate by

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 33

design and the cumulative intensity result can be unambiguously decoded to indicate the presence or absence of analytes A-P. This coding scheme is capable of detecting the presence or absence of M analytes, where $M = \log_2 (F + 1)$ and F is the maximum cumulative value of a signal (e.g., signal intensity). By adding a second component to this signal (e.g., color, as depicted in Table 8), the capacity of this coding scheme can be increased to $M = C * \log_2 (F + 1)$, where C is the number of colors used in said coding scheme.

[00105] As described throughout this specification, any suitable value may be used for C or F. For example, C may be at least about 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, 49, 50, 60, 70, 80, 90, 100, or more. F may be at least about 2, 4, 8, 16, 32, 64, 128, 255, 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536, 131072, 262144, 524288, 1048576, and so on. As described elsewhere in this disclosure, both the initial starting value of F and the steps in its progression may also be varied.

III. Analytes

[00106] An analyte may be any suitable analyte that may be analyzed using the methods and compositions of the invention, where the analyte is capable of interacting with a reagent in order to generate a signal with at least two components that can be measured. An analyte may be naturally-occurring or synthetic. An analyte may be present in a sample obtained using any methods known in the art. In some cases, a sample may be processed before analyzing it for an analyte. The methods and compositions presented in this disclosure may be used in solution phase assays, without the need for particles (such as beads) or a solid support.

[00107] In some cases, an analyte may be a polynucleotide, such as DNA, RNA, peptide nucleic acids, and any hybrid thereof, where the polynucleotide contains any combination of deoxyribo- and/or ribo-nucleotides. Polynucleotides may be single stranded or double stranded, or contain portions of both double stranded or single stranded sequence. Polynucleotides may contain any combination of nucleotides or bases, including, for example, uracil, adenine, thymine, cytosine, guanine, inosine, xanthine, hypoxanthine, isocytosine, isoguanine and any nucleotide derivative thereof. As used herein, the term "nucleotide" may include nucleotides and nucleosides, as well as nucleoside and nucleotide analogs, and modified nucleotides, including both synthetic and naturally occurring species. Polynucleotides may be any suitable polynucleotide for which one or more reagents (or probes) as described herein may be produced, including but not limited to cDNA, mitochondrial DNA (mtDNA),

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 34

messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), nuclear RNA (nRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), small Cajal body-specific RNA (scaRNA), microRNA (miRNA), double stranded (dsRNA), ribozyme, riboswitch or viral RNA. Polynucleotides may be contained within any suitable vector, such as a plasmid, cosmid, fragment, chromosome, or genome.

[00108] Genomic DNA may be obtained from naturally occurring or genetically modified organisms or from artificially or synthetically created genomes. Analytes comprising genomic DNA may be obtained from any source and using any methods known in the art. For example, genomic DNA may be isolated with or without amplification. Amplification may include PCR amplification, multiple displacement amplification (MDA), rolling circle amplification and other amplification methods. Genomic DNA may also be obtained by cloning or recombinant methods, such as those involving plasmids and artificial chromosomes or other conventional methods (see Sambrook and Russell, *Molecular Cloning: A Laboratory Manual*., cited *supra*.) Polynucleotides may be isolated using other methods known in the art, for example as disclosed in *Genome Analysis: A Laboratory Manual Series* (Vols. I-IV) or *Molecular Cloning: A Laboratory Manual*. If the isolated polynucleotide is an mRNA, it may be reverse transcribed into cDNA using conventional techniques, as described in Sambrook and Russell, *Molecular Cloning: A Laboratory Manual*., cited *supra*.

[00109] An analyte may be a protein, polypeptide, lipid, carbohydrate, sugar, small molecule, or any other suitable molecule that can be detected with the methods and compositions provided herein. An analyte may be an enzyme or other protein. An analyte may be a drug or metabolite (e.g., anti-cancer drug, chemotherapeutic drug, anti-viral drug, antibiotic drug, or biologic). An analyte may be any molecule, such as a co-factor, receptor, receptor ligand, hormone, cytokine, blood factor, antigen, steroid, or antibody.

[00110] An analyte may be any molecule from any pathogen, such as a virus, bacteria, parasite, fungus, or prion (e.g., PrP^{Sc}). Examples of viruses include those from the families Adenoviridae, Flaviviridae, Hepadnaviridae, Herpesviridae, Orthomyxoviridae, Papovaviridae, Paramyxoviridae, Picornaviridae, Polyomavirus, Retroviridae, Rhabdoviridae, and Togaviridae. Specific examples of viruses include adenovirus, astrovirus, bocavirus, BK virus, coxsackievirus, cytomegalovirus, dengue virus, Ebola virus, enterovirus, Epstein-Barr virus, feline leukemia virus, hepatitis virus, hepatitis A virus, hepatitis B virus, hepatitis C virus, hepatitis D virus, hepatitis E virus, herpes simplex virus (HSV), HSV type 1, HSV type 2, human immunodeficiency virus (HIV), HIV type 1, HIV type 2, human papilloma virus (HPV), HPV type 1, HPV type 2, HPV type 3, HPV type 4, HPV type 6, HPV

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 35

type 10, HPV type 11, HPV type 16, HPV type 18, HPV type 26, HPV type 27, HPV type 28, HPV type 29, HPV type 30, HPV type 31, HPV type 33, HPV type 34, HPV type 35, HPV type 39, HPV type 40, HPV type 41, HPV type 42, HPV type 43, HPV type 44, HPV type 45, HPV type 49, HPV type 51, HPV type 52, HPV type 54, HPV type 55, HPV type 56, HPV type 57, HPV type 58, HPV type 59, HPV type 68, HPV type 69, influenza virus, JC virus, Marburg virus, measles virus, mumps virus, Norwalk virus, parovirus, polio virus, rabies virus, respiratory syncytial virus, retrovirus, rhinovirus, rotavirus, Rubella virus, smallpox virus, vaccinia virus, West Nile virus, and yellow fever virus.

[00111] Examples of bacteria include those from the genuses *Bordetella, Borrelia, Brucella, Campylobacter*, *Chlamydia, Clostridium, Corynebacterium, Enterococcus, Escherichia, Francisella, Haemophilus, Helicobacter, Legionella, Leptospira, Listeria, Mycobacterium*, *Mycoplasma, Neisseria, Pseudomonas*, *Rickettsia, Salmonella*, *Shigella*, *Staphylococcus, Streptococcus*, *Treponema, Vibrio, and Yersinia.* Specific examples of bacteria include *Bordetella par apertussis, Bordetella pertussis, Borrelia burgdorferi, Brucella abortus, Brucella canis, Brucella melitensis, Brucella suis, Campylobacter jejuni, Chlamydia pneumoniae, Chlamydia psittaci, Chlamydia trachomatix, Clostridium botulinum, Clostridium difficile, Clostridium perfringens, Clostridium tetani, Corynebacterium diphtheriae, Enterococcus faecalis, Enterococcus faecium, Escherichia coli*, *Francisella tularensis, Haemophilus influenzae, Helicobacter pylori, Legionella pneumophila, Leptospira interrogans, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Mycobacterium ulcerans, Mycoplasma pneumoniae, Neisseria gonorrhoeae, Neisseria meningitidis, Pseudomonas aeruginosa, Rickettsia rickettsii, Salmonella choleraesuis, Salmonella dublin, Salmonella enteritidis, Salmonella typhi, Salmonella typhimurium, Shigella sonnei, Staphylococcus aureus, Staphylococcus epidermidis, Staphylococcus saprophyticus, Streptococcus agalactiae, Streptococcus pneumoniae, Streptococcus pyogenes, Treponema pallidum, Vibrio cholerae, Yersinia pestis, and Yersinia enterocolitica.*

[00112] Examples of parasites include those from the genuses *Acanthamoeba, Babesia, Balamuthia, Balantidium, Blasocystis, Cryptosporidium, Dientamoeba, Entamoeba, Giardia, Isospora, Leishmania, Naegleria, Pediculus, Plasmodium, Rhinosporidium, Sarcocystis, Schistosoma, Toxoplasma, Trichomonas,* and *Trypanosoma.* Specific examples of parasites include *Babesia divergens, Babesia bigemina, Babesia equi, Babesia microfti, Babesia duncani, Balamuthia mandrillaris, Balantidium coli, Dientamoeba fragilis, Entamoeba histolytica, Giardia lamblia, Isospora belli, Naegleria fowleri, Pediculus humanus, Plasmodium falciparum, Plasmodium knowlesi, Plasmodium malariae, Plasmodium ovale, Plasmodium vivax, Rhinosporidium seeberi, Sarcocystis*

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 36

bovihominis, Sarcocystis suihominis, Schistosoma mansoni, Toxoplasma gondii, Trichomonas vaginalis, Trypanosoma brucei, and *Trypansoma cruzi.*

[00113] Examples of fungi include those from the genuses *Apophysomyces, Aspergillus, Blastomyces, Candida, Cladosporium, Coddidioides, Cryptococcos, Exserohilum, Fusarium, Histoplasma, Pichia, Pneumocystis, Saccharomyces, Sporothrix, Stachybotrys,* and *Trichophyton.* Specific examples of fungi include *Aspergillus fumigatus, Aspergillus flavus, Aspergillus clavatus, Blastomyces dermatitidis, Candida albicans, Coccidioides immitis, Crytptococcus neoformans, Exserohilum rostratum, Fusarium verticillioides, Histoplasma capsulatum, Pneumocystis jirovecii, Sporothrix schenckii, Stachybotrys chartarum,* and *Trichophyton mentagrophytes.*

[00114] In some cases, the methods provided in this disclosure may be used to detect any one of the analytes described above, or elsewhere in the specification. In some cases the methods provided in this disclosure may be used to detect panels of the analytes described above, or elsewhere in the specification. For example, a panel may comprise an analyte selected from the group consisting of any 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, 49, 50, 60, 70, 80, 90, 100 or more analytes described above or elsewhere in the specification.

[00115] An analyte may be obtained from any suitable location, including from organisms, whole cells, cell preparations and cell-free compositions from any organism, tissue, cell, or environment. Analytes may be obtained from environmental samples, biopsies, aspirates, formalin fixed embedded tissues, air, agricultural samples, soil samples, petroleum samples, water samples, or dust samples. In some instances, an analyte may be obtained from bodily fluids which may include blood, urine, feces, serum, lymph, saliva, mucosal secretions, perspiration, central nervous system fluid, vaginal fluid, or semen. Analytes may also be obtained from manufactured products, such as cosmetics, foods, personal care products, and the like. Analytes may be the products of experimental manipulation including, recombinant cloning, polynucleotide amplification, polymerase chain reaction (PCR) amplification, purification methods (such as purification of genomic DNA or RNA), and synthesis reactions.

[00116] More than one type of analyte may be detected in each multiplexed assay. For example, a polynucleotide, a protein, a polypeptide, a lipid, a carbohydrate, a sugar, a small molecule, or any other suitable molecule may be detected simultaneously in the same multiplexed assay with the use of suitable reagents. Any combination of analytes may be detected at the same time.

[00117] Detection of an analyte may be useful for any suitable application, including research, clinical, diagnostic, prognostic, forensic, and monitoring applications. Exemplary applications include

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 37

detection of hereditary diseases, identification of genetic fingerprints, diagnosis of infectious diseases, cloning of genes, paternity testing, criminal identification, phylogeny, anti-bioterrorism, environmental surveillance, and DNA computing. For example, an analyte may be indicative of a disease or condition. An analyte may be used to make a treatment decision, or to assess the state of a disease. The presence of an analyte may indicate an infection with a particular pathogen, or any other disease, such as cancer, autoimmune disease, cardiorespiratory disease, liver disease, digestive disease, and so on. The methods provided herein may thus be used to make a diagnosis and to make a clinical decision based on that diagnosis. For example, a result that indicates the presence of a bacterial polynucleotide in a sample taken from a subject may lead to the treatment of the subject with an antibiotic.

[00118] In some cases the methods and compositions of the invention may be used to detect at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, or 1000 analytes. In some cases the methods and compositions of the invention may be used to detect 7-50, 8-40, 9-30, 10-20, 10-15, 8-12, or 7-12 analytes.

[00119] In some cases, this disclosure provides assays that are capable of unambiguously detecting the presence or absence of each of 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, or 1000 analytes, in any combination of presence or absence, in a single sample volume without immobilization, separation, mass spectrometry, or melting curve analysis. In some cases, this disclosure provides assays that are capable of unambiguously detecting the presence or absence of each of at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, or 1000 analytes, in any combination of presence or absence, in a single sample volume without immobilization, separation, mass spectrometry, or melting curve analysis. In some cases, this disclosure provides assays that are capable of unambiguously detecting the presence or absence of less than 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, or 1000 analytes, in any combination of

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 38

presence or absence, in a single sample volume without immobilization, separation, mass spectrometry, or melting curve analysis.

IV. Signals

[00120] The methods presented in this disclosure may be used with any quantifiable signal. In some cases, this disclosure provides coding schemes and methods to encode an infinite number of targets without degeneracy, using a single component of a signal (e.g., intensity). For example, as described above and in Example 3, a coding scheme may rely on a multiplicity of signal intensity without consideration of color. Although fluorescent probes have been used to illustrate this principle, the coding scheme is equally applicable to any other method providing a quantifiable signal, including an electrochemical signal and a chemiluminescent signal, as described elsewhere in this disclosure.

[00121] The methods presented in this disclosure may also utilize the measurement of a signal in at least two dimensions, also referred to as the measurement of at least two components of a signal. In comparison to the coding scheme described in the paragraph above, which relies on, for example, signal intensity to differentiate between analytes, utilization of at least two components of a signal (e.g., color and intensity) allows the generation of more unique codes per unit of signal intensity bandwidth. When at least two components of a signal are utilized, a cumulative measurement of the at least two components may be obtained for a single sample volume. For example in the coding scheme described in Table 5, the presence of each analyte results in a particular intensity (one component of the signal) in each of the four colors (a second component of a signal). The combination of these constituent signals leads to a cumulative signal that may be measured by measuring an intensity at each wavelength or range of wavelengths. The corresponding coding scheme or decoding matrix may then be used to convert the cumulative measurement into a determination of the presence or absence of an analyte.

[00122] In some cases, a quantifiable signal comprises a waveform that has both a frequency (wavelength) and an amplitude (intensity). A signal may be an electromagnetic signal. An electromagnetic signal may be a sound, a radio signal, a microwave signal, an infrared signal, a visible light signal, an ultraviolet light signal, an x-ray signal, or a gamma-ray signal. In some cases, an electromagnetic signal may be a fluorescent signal, for example a fluorescence emission spectrum that may be characterized in terms of wavelength and intensity.

[00123] In certain portions of this disclosure, the signal is described and exemplified in terms of a fluorescent signal. This is not meant to be limiting, and one of ordinary skill in the art will readily recognize that the principles applicable to the measurement of a fluorescent signal are also applicable to

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 39

other signals. For example, like fluorescent signals, any of the electromagnetic signals described above may also be characterized in terms of a wavelength and an intensity. The wavelength of a fluorescent signal may also be described in terms of color. The color may be determined based on measuring intensity at a particular wavelength or range of wavelengths, for example by determining a distribution of fluorescent intensity at different wavelengths and/or by utilizing a band pass filter to determine the fluorescence intensity within a particular range of wavelengths. Such band pass filters are commonly employed in a variety of laboratory instrumentation, including quantitative PCR machines. Intensity may be measured with a photodetector. A range of wavelengths may be referred to as a "channel."

[00124] In some cases, the methods provided in this disclosure may be used with any signal where the cumulative signal scales with the constituent signals of the same color, frequency, absorption band, and so on. However, the cumulative signal need not be digital or scale linearly with the number and intensity of the constituent signals. For example, if the physical principle of measurement is absorption, the cumulative attenuation is a product of constituent attenuations while the constituent concentrations are additive, due to the exponential nature of the Beer-Lambert law. The logarithm of the cumulative attenuation will then scale linearly with constituent concentrations in each absorption band (the equivalent of color, if fluorescent detection is used). The methods of the invention are therefore applicable. The methods of the invention may also be used with chemiluminescent signals and electrochemical signals.

[00125] The number of signals, and the number of dimensions or components measured, may also be expanded beyond the numbers shown in the exemplary embodiments of the invention, leading to an expansion in multiplexing capability. The exemplary embodiments provided in this disclosure utilize coding schemes constructed utilizing a fluorescent signal with one or two components measured: wavelength and/or intensity. The number of analytes that can be encoded can be increased by increasing the number of wavelengths and/or intensities. The number of analytes that can be encoded can also be increased by increasing the number of signals, for example by combining a fluorescent signal with an electrochemical signal or a FRET signal (fluorescence resonance energy transfer).

[00126] In some cases, more than two components of a signal may be measured. For example, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 18, 20 or more components of a signal may be measured. At least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 18, or 20 components of a signal may be measured. At least 2, but fewer than 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 18, or 20 components of a signal may be measured. In some cases, 2-3, 2-4, 2-5, 2-6, 3-5, 3-6, 3-8, or 5-

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 40

10 components of a signal may be measured. These additional components may include kinetic components, such as a rate of signal decay and rate of photobleaching.

[00127] If a fluorescent signal is employed, the number of analytes that can be encoded may be further expanded by utilizing additional fluorophores. For example, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more fluorophores may be used. In some cases, at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 fluorophores may be used. In some cases, fewer than 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, 49, or 50 fluorophores may be used. In some cases, 4-6, 4-7, 4-8, 4-9, 4-10, 5-6, 5-7, 5-8, 5-9, 5-10, 10-15, or 10-20 fluorophores may be used.

[00128] Generally, the number of analytes that can be encoded may be further expanded by utilizing additional first values (e.g., intensities) that are values or ranges of values from a first component of a signal. For example, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more first values that are values or ranges of values from a first component of a signal may be used. In some cases, at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more first values that are values or ranges of values from a first component of a signal may be used. In some cases, fewer than 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 first values that are values or ranges of values from a first component of a signal may be used. In some cases, 4-20, 4-6, 4-7, 4-8, 4-9, 4-10, 5-6, 5-7, 5-8, 5-9, 5-10, 10-15, or 10-20 first values that are values or ranges of values from a first component of a signal may be used.

[00129] In instances in which intensity is a component of a signal that is quantified, such as when a fluorescent signal is utilized, the presence of an analyte may be encoded using a variety of intensities or ranges of intensities. For example, a coding scheme may utilize 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more intensities or ranges of intensities. A coding scheme may utilize at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more intensities

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 41

or ranges of intensities. A coding scheme may utilize fewer than 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, 49, or 50 intensities or ranges of intensities. In some cases, a coding scheme may utilize 2-20, 2-3, 2-4, 2-5, 2-6, 2-7, 2-8, 2-9, 2-10, 3-4, 3-5, 3-6, 3-7, 3-8, 3-9, 3-10, 4-5, 4-6, 4-7, 4-8, 4-9, 4-10, 5-6, 5-7, 5-8, 5-9, or 5-10 intensities or ranges of intensities.

[00130] The number of analytes that can be encoded may be further expanded by utilizing additional second values (e.g., wavelengths) that are values or ranges of values from a second component of a signal. For example, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more second values that are values or ranges of values from a second component of a signal may be used. In some cases, at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more second values that are values or ranges of values from a second component of a signal may be used. In some cases, fewer than 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 second values that are values or ranges of values from a second component of a signal may be used. In some cases, 4-20, 4-6, 4-7, 4-8, 4-9, 4-10, 5-6, 5-7, 5-8, 5-9, 5-10, 10-15, or 10-20 second values that are values or ranges of values from a second component of a signal may be used.

[00131] In instances in which wavelength is a component of a signal that is quantified, such as when a fluorescent signal is utilized, the presence of an analyte may be encoded using a variety of wavelengths or ranges of wavelengths. For example, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more wavelengths or ranges of wavelengths may be used. In some cases, at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more wavelengths or ranges of wavelengths may be used. In some cases, fewer than 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 wavelengths or ranges of wavelengths may be used. In some cases, 4-20, 4-6, 4-7, 4-8, 4-9, 4-10, 5-6, 5-7, 5-8, 5-9, 5-10, 10-15, or 10-20 wavelengths or ranges of wavelengths may be used.

[00132] In some cases, when degeneracy is eliminated, the methods of the invention are capable of detecting the presence or absence of at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 42

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, 49, 50, or more analytes in a single volume when each reagent used to generate a signal in the volume generates only one second value (e.g., each reagent, emits light at only one wavelength).

[00133] In other cases, when degeneracy is eliminated, the methods of the invention are capable of detecting the presence or absence of at least 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, 49, 50, or more analytes in a single volume using a total of four second values in that volume (e.g., a total of four wavelengths or ranges or wavelengths, which might be implemented by using four spectrally resolvable fluorophores).

[00134] As described throughout the specification, the assay provided herein utilizes cumulative measurements on a sample. A cumulative measurement may be, for example, a single measurement of intensity values, or a measurement of intensity values at one or more wavelengths or ranges of wavelengths. A plurality of cumulative measurements may be obtained. For example, an intensity may be measured at 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more wavelengths or ranges of wavelengths. An intensity may be measured at at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more wavelengths or ranges of wavelengths. An intensity may be measured at less than 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 wavelengths or ranges of wavelengths.

[00135] More generally a cumulative measurement may be obtained for any quantifiable component of a signal, and for any quantifiable component of a signal at another quantifiable component of a signal. For example, at least a first component of a signal may be measured at 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more second components of a signal. At lest a first component of a signal may be measured at at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more second components of a signal. At least a first component of a signal may be measured at less than 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 second components of a signal.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 43

[00136] As is apparent from this disclosure, each analyte to be detected can be encoded as a code utilizing any number of suitable components of a signal or any number of signals. For example, each analyte to be detected can be encoded in 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more components of a signal or signals. In some cases, each analyte to be detected can be encoded in at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more components of a signal or signals. In some cases, each analyte to be detected can be encoded in fewer than 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 components of a signal or signals. In some cases, 1-4, 4-20, 4-6, 4-7, 4-8, 4-9, 4-10, 5-6, 5-7, 5-8, 5-9, 5-10, 10-15, or 10-20 components of signal or signals may be used to encode each analyte. Each analyte in a coding scheme may be encoded by the same number of components of a signal, or signals, or different numbers of components of a signal, or signals.

A. Probes, Primers, Fluorophores, and Quenchers

[00137] Some of the methods provided in this disclosure utilize a reagent that generates a signal in the presence of an analyte. Any suitable reagent may be used with the invention. Generally, a reagent will have an analyte-specific component and a component that generates a signal in the presence of the analyte. In some cases, these reagents are referred to as probes. The probes may be hybridization probes. The hybridization probes may be an oligonucleotide probe attached to a fluorophore and a quencher (e.g., a TAQMAN probe).

[00138] The methods of the invention may use one or more reagents or probes to detect the presence or absence of each analyte. For example, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more or probes may be used to detect the presence or absence of each analyte. In some cases, at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 probes may be used to detect the presence or absence of each analyte. In some cases, fewer than 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, 49, 50 probes may be used to detect the presence or

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 44

absence of each analyte. In some cases, the number of probes used to detect the presence or absence of each analyte is 1-10, 2-9, 2-8, 2-7, 2-6, 2-5, 2-4, 1-3, or 1-4.

[00139] In some cases, a sample is contacted with 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, 49, 50, or more probes to detect the presence or absence of all analytes. In some cases, a sample is contacted with at least 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, 49, 50 or more probes to detect the presence or absence of all analytes. In some cases, a sample is contacted with fewer than 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, 49, or 50 probes to detect the presence or absence of all analytes. In some cases, the number of probes that a sample is contacted with to detect the presence or absence of all analytes is 4-50, 4-40, 4-30, 4-20, 5-15, 5-10, or 3-10.

[00140] As described above, oligonucleotide probes attached to a fluorophore and a quencher may be used to detect the presence of an analyte in a polynucleotide amplification assay. So long as the quencher and the fluorophore are in proximity, the quencher quenches the fluorescence emitted by the fluorophore upon excitation by a light source. The sequence of the oligonucleotide probe is designed to be complementary to a polynucleotide sequence present in an analyte, and therefore capable of hybridizing to the polynucleotide sequence present in the analyte. Hybridization of the oligonucleotide probe may be performed in a nucleic acid amplification reaction comprising primers (e.g., a polymerase chain reaction). Upon extension of the primers by a DNA polymerase, the 5' to 3' exonuclease activity of the polymerase degrades the probe, releasing the fluorophore and the quencher into the medium. The proximity between the fluorophore and the quencher is broken and the signal from the fluorophore is no longer quenched. Thus, the amount of fluorescence detected is a function of the amount of analyte present. If no analyte is present, the probe will not hybridize to an analyte, and the fluorophore and quencher will remain in close proximity. Little or no signal will be produced.

[00141] Oligonucleotide probes may have one or a plurality of fluorophores and quenchers per probe. For example, in some embodiments an oligonucleotide probe may comprise 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more fluorophores. An oligonucleotide probe may comprise at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 fluorophores. An oligonucleotide probe may comprise fewer than 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 45

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, 49, or 50 fluorophores.

[00142] An oligonucleotide probe may comprise 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more quenchers. An oligonucleotide probe may comprise at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 quenchers. An oligonucleotide probe may comprise fewer than 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, 49, or 50 quenchers.

[00143] Attachment of probes and quenchers to a probe may be performed in the same reaction or in serial reactions. A series of reactions may be performed to label probes with at least one fluorophore and the reaction products may be mixed to generate a mixture of probes with different fluorophores.

[00144] If an oligonucleotide probe comprises two or more fluorophores, these fluorophores may be arranged to allow Förster (Fluorescence) resonance energy transfer (FRET) to occur between the fluorophores. Briefly, FRET is a mechanism of energy transfer between fluorophores. Using FRET-based probes allows one excitation source to generate two different fluorescence emission signals by excitation of a single fluorophore. For example, the methods described herein may used paired probes, where one probe in the pair is attached to a fluorophore and a quencher and a second probe is attached to two fluorophores (in close enough proximity for FRET to occur) and a quencher. Any combination of fluorophores and quenchers that provide for FRET may be used. Such an approach doubles the number of fluorescent probes that may be used with an excitation source, because a single excitation source can be used to produce two signals: one from the probe with one fluorophore and one from the probe with two fluorophores in close enough proximity to provide for FRET. For example, using the encoding method described in Table 8, the number of unambiguously detectable analytes could be increased from 16 to 32 by pairing each of the probes in Table 8 with a corresponding FRET probe.

[00145] The primers may be specific for a particular analyte and capable of amplifying a region complementary to a probe. In some cases, the number of pairs of primers used is equivalent to the number of probes. In other cases, the number of probes used may exceed the number of primer pairs used. In still other cases, the number of primer pairs used may exceed the number of probes used. In some cases a sample to be analyzed is contacted with 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more pairs of primers. In some cases a sample to be analyzed is contacted with

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 46

at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more pairs of primers. In some cases a sample to be analyzed is contacted with fewer than 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, 49, or 50 pairs of primers. In some cases, the number of pairs of primers is 2-10, 3-15, 4-20, 3-10, 4-10, 5-10, 6-8, or 6-10.

[00146] The primers may amplify regions of a polynucleotide in which different numbers of hybridization probes hybridize. For example, at least one pair of primers may amplify a region complementary to at least 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, 49, or 50 hybridization probes. In some cases, all of said pairs of primers may amplify a region complementary to at least 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, 49, or 50 hybridization probes.

[00147] In one example, a first probe may be labeled with the fluorophore 5'-fluorescein amidite (5'-FAM) and the quencher black hole quencher 1 (BHQ-1). A second probe may be labeled with 5'-FAM in close proximity (e.g., attached to) to cyanine 5.5 (Cy5.5) and the quencher black hole quencher 3 (BHQ-3). Upon digestion of the probe, via the nuclease activity of the polymerase, two fluorescent signals are generated from a single excitation wavelength (e.g., 470 nm). The fluorophore from the first probe will fluoresce at about 520 nm. The fluorophores on the second probe undergo FRET. The donor fluorophore, FAM is excited by the excitation light source (e.g., 470 nm) and transfers its energy to the acceptor fluorophore, Cy5.5. The acceptor fluorophore emits at about 705 nm. These fluorophores and quenchers are merely exemplary. Any fluorophores that can undergo FRET and any quenchers that can quench fluorescence are suitable for use with the invention. Methods for producing FRET-based probes are described in Jothikumar et al., BioTechniques, 2009, 46(7):519-524.

[00148] In some examples, a single hybridization probe may be used for each analyte. In order to utilize multi-color encoding (e.g., as shown in Table 3), each probe may be labeled with a plurality of fluorophores at pre-determined ratios. The ratio may be determined so that a positive control signal provides the same intensity as other positive control signals of the same intensity, in the same color. This approach reduces the number of probes that must be synthesized and may be less expensive to deploy than methods utilizing multiple probes. Moreover, in the case of hybridization probes, it is easier to fit one hybridization probe to an analyte sequence than multiple probes.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 47

[00149] Although many aspects of the invention are exemplified using nucleic acid-based probes, one of ordinary skill in the art will readily recognize that other forms of probes would work equally well with the invention described in this disclosure. For example, a binding molecule specific to an analyte could be used as a probe. Non-limiting exemplary binding molecules include an antibody recognizing an analyte, and generating a signal in the presence of an analyte.

[00150] In embodiments of the invention that utilize fluorescent labels, any suitable fluorescent label may be used. Exemplary fluorescent labels suitable for use with the invention include rhodamine, rhodol, fluorescein, thiofluorescein, aminofluorescein, carboxyfluorescein, chlorofluorescein, methylfluorescein, sulfofluorescein, aminorhodol, carboxyrhodol, chlororhodol, methylrhodol, sulforhodol; aminorhodamine, carboxyrhodamine, chlororhodamine, methylrhodamine, sulforhodamine, and thiorhodamine, cyanine, indocarbocyanine, oxacarbocyanine, thiacarbocyanine, merocyanine, cyanine 2, cyanine 3, cyanine 3.5, cyanine 5, cyanine 5.5, cyanine 7, oxadiazole derivatives, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, pyrene derivatives, cascade blue, oxazine derivatives, Nile red, Nile blue, cresyl violet, oxazine 170, acridine derivatives, proflavin, acridine orange, acridine yellow, arylmethine derivatives, auramine, crystal violet, malachite green, tetrapyrrole derivatives, porphin, phtalocyanine, and bilirubin. Exemplary quenchers include black hole quenchers, such as BHQ-0, BHQ-1, BHQ-2, BHQ-3; ATTO quenchers, such as ATTO 540Q, ATTO580Q, and ATTO612Q; and the like.

[00151] The fluorophores that may be used with the invention are not limited to any of the fluorophores described herein. For example, fluorophores with improved properties are continually developed, and these fluorophores could readily be used with the methods provided in this disclosure. Such improved fluorophores include quantum dots, which may emit energy at different wavelengths after being excited at a single wavelength. The advantage of using such fluorophores is that only a single excitation source is needed, but many different signals may be quantified, for example in terms of color and intensity. Moreover, fluorophores with narrow emission spectra would be particularly useful with the method described herein, as such fluorophores could be included in multiplex assays with minimal or no overlap between their emission spectra, thereby offering many more "colors" and boosting significantly the overall number of coded analytes.

[00152] In some cases, one or more reagents are lyophilized. Any suitable reagent may be lyophilized. For example, probes, primers, enzymes, antibodies, or any other reagent used for detection may be lyophilized. In some cases, a sample comprising an analyte may also be lyophilized. Lyophilization may be useful, for example, when distributing reagents and/or samples in developing

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 48

regions where access to cold storage is expensive, not readily available, or not reliably available. In one example, lyophilized reagents comprise any probe or analyte-specific reagent described in this disclosure or otherwise suitable for use with the invention. In another example, lyophilized reagents comprise PCR primers. In yet another example, lyophilized reagents comprise reagents suitable for performing a PCR reaction.

V. Analytical Techniques and Instrumentation

[00153] The methods provided herein are suitable for use with a variety of detection methods. For example, the methods may be applied using an analytical technique that measures the wavelength and intensity of a fluorescent signal. This may be accomplished by measuring the intensity of a signal across a spectrum of wavelengths, or by using band pass filters that restrict the passage of certain wavelengths of light, thereby allowing only light of certain wavelengths to reach a photodetector. Many real-time PCR and quantitative PCR instruments comprise an excitation light source and band pass filters that enable the detection of fluorescent signals in four colors (e.g., blue, green, yellow, and red). Therefore, the methods of the invention can be readily applied using instruments widely used in the art. Importantly, the methods and compositions provided herein may be used to detect multiple analytes by obtaining a cumulative measurement on a single solution. No separation is necessary. The invention does not require the use of beads or a solid phase. Of course, one of ordinary skill in the art would understand that the invention could be used with separation, beads, or a solid phase, if desired.

[00154] **Fig. 5** shows two examples of instrumental configurations that may be used to obtain cumulative measurements useful for the methods of the invention. With reference to **Fig. 5**, an instrument with a detector configured to detect both wavelength and intensity (e.g., a fluorometer) is shown in **501.** With reference to **501**, at least one excitation light source **502** is directed into a chamber **503** containing analytes and reagents that generate a signal in the presence of the analytes **504.** As described elsewhere, the analyte may be a nucleic acid and the reagents generating the signal may be hybridization probes comprising a fluorophore and a quencher. If the analyte is present, an emission signal **505** is generated. In the configuration shown in **501**, the wavelength and intensity of this emission signal are measured across a spectrum by a detector capable of generating a fluorescence emission spectrum **506.**

[00155] The wavelength and intensity may also be determined using a combination of a photodetector and band pass filters. This configuration is used in several thermal cyclers known in the art. With reference to **Fig. 5**, an instrument with band pass filters and a photodetector is depicted in **507.**

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 49

With reference to **507**, at least one excitation light source **508** is directed into a chamber **509** containing analytes and reagents that generate a signal in the presence of the analytes **510.** As described elsewhere, the analyte may be a nucleic acid and the reagents generating the signal may be hybridization probes comprising a fluorophore and a quencher. If the analyte is present, an emission signal **511** is generated. In the configuration shown in **507**, band pass filters **512, 513,** and **514** are used to restrict the passage of light to light within certain ranges of wavelengths. For example band pass filter **512** may restrict the passage of light to light within wavelength range 1. Band pass filter **513** may restrict the passage of light to light within wavelength range 2. Band pass filter **514** may restrict the passage of light to light within wavelength range 3. Any number of band pass filters (e.g., 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, or more) may be used with the invention. In addition, the positions of the band pass filters in **507** are not meant to be limiting. In some cases, emitted light may be passed through more than one band pass filter at any time. In other cases, the band pass filters may be configured so that all emitted light passes through only one band pass filter at a time. After passage through a band pass filter, filtered light **515, 516,** and **517** is detected by a photodetector. For example, filtered light **515** may be light within wavelength range 1, filtered light **516** may be light within wavelength range 2, and filtered light **517** may be light within wavelength range 3. The intensity of each of the filtered lights **515, 516,** and **517** may then be quantified by a photodetector **518.**

[00156] The methods described in this disclosure are compatible with a variety of amplification methods, including quantitative PCR (qPCR) methods, end point PCR methods, reverse transcriptase PCR, and digital PCR methods. Digital PCR methods (e.g., DROPLET DIGITAL (BIORAD) and DYNAMIC ARRAY (FLUIDIGM)) produce highly sensitive quantification of polynucleotide copy numbers. The methods provided herein can be easily integrated into these systems to significantly expand their throughput by allowing multiplexing in a droplet or a dynamic array.

[00157] Similarly, the methods described in this disclosure may be applied in a real-time PCR assay. For example, real-time data may be fully recorded as the PCR runs to completion. The end-point values may then be decoded to indicate the presence or absence of analytes. The individual cycle thresholds (Ct) values may be analyzed for each color by detecting the maxima of the second derivative of the real-time curve in that color. As a particular analyte is amplified, fluorophores are released from its corresponding probes in the same ratios as the color multiplicities in the coding of that sequence. Hence, e.g. a code of 1100 will have the same Ct in blue and green, while a code of 1400 will have its green Ct precede its blue Ct by 2 cycles. This additional data enables determination of the identity and starting quantity of each analyte.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 50

[00158] The methods described herein may also be used directly on a tissue sample. For example, a tissue sample may be obtained and photolithographic methods may be used to build wells directly onto the tissue sample. The tissue sample may be fixed prior to building the wells. The tissue sample may then be analyzed by dispensing appropriate reagents into the wells in the tissue sample. The encoding and decoding methods described in this disclosure may be used for the multiplexed detection of analytes within the wells etched into the tissue. Each well may correspond to a single cell or a few cells, providing excellent spatial resolution when analyzing different portions of a tissue for an analyte. This method may be used to detect analytes in different areas of the same tissue.

[00159] In some cases, instruments may be modified or constructed, for example, to provide additional excitation light sources, at multiple wavelength, and/or to provide additional band pass filters or a capability of determining a complete spectrum. Including additional excitation sources would allow for the excitation of a larger variety of fluorophores. Including additional band pass filters, or modifying or constructing an instrument capable of determining an entire spectrum allows detection of a wider variety of emissions from fluorophores. These techniques can be used to increase the number of fluorophores that can be used with the methods described in this disclosure and, accordingly, to increase the number of analytes that be simultaneously detected.

[00160] In some cases, the methods described in this disclosure utilize end-point PCR methods. However, in some cases a result may be obtained before the end-point. This may be advantageous when, for example, results are needed as soon as possible. For example, in one case calibration experiments are performed in which different starting concentrations of an analyte are analyzed to determine the cycle number at which the signal becomes saturated, as a function of the starting amount of the analyte. Using this data, a computer monitoring a PCR reaction in real-time can be programmed to search for saturation up to the maximal cycle number according to the limit of detection (LOD) of the particular system. If there is no amplification of an analyte (e.g., other than a positive control, if present), by the maximal cycle number, the result for that analyte is negative. If there is amplification of the analyte by the maximal cycle number, the result for that analyte is positive and the saturation intensity in each color may be used to decode the result using the coding scheme. In both cases, there would be no need to run the PCR reaction beyond the number of cycles set by the LOD for that instrument. Although PCR has been used to illustrate this quantitative method, one of ordinary skill in the art will readily recognize that similar principles could be applied in any catalytic system where the starting amount of an analyte limits the rate of reaction, and the rate of reaction can be measured and reported by an evolution in a signal intensity over time.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 51

VI. Compositions and Kits

[00161] This disclosure also provides compositions and kits for use with the methods described herein. The compositions may comprise any component, reaction mixture and/or intermediate described herein, as well as any combination. For example, the disclosure provides detection reagents for use with the methods provided herein. Any suitable detection reagents may be provided, including hybridization probes labeled with a fluorophore and a quencher and primers, as described elsewhere in the specification.

[00162] In some cases, compositions comprise reagents for the detection of at least seven analytes using four fluorophores. In some cases, compositions comprise reagents for the detection 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, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 analytes using four fluorophores. In some cases, compositions comprise reagents for the detection of at least 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, 49, or 50 analytes using five fluorophores. In some cases, compositions comprise reagents for the detection of at least 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, 49, or 50 analytes using six fluorophores. In some cases, compositions comprise reagents for the detection of at least 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, 49, or 50 analytes using seven fluorophores.

[00163] In some cases the compositions comprise primers. The compositions may comprise at least 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 pairs of primers.

[00164] The invention also provides kits for carrying out the methods of the invention. Accordingly, a variety of kits are provided in suitable packaging. The kits may be used for any one or more of the uses described herein, and, accordingly, may contain instructions for detecting the presence or absence of each analyte of a plurality of analytes. A kit may comprise a coding scheme, or a decoding matrix, to assist the user in converting a cumulative measurement to a result indicating the presence or absence of each of a plurality of analytes. A kit may be a diagnostic kit, for example, a diagnostic kit suitable for the detection of any analyte, including the analytes recited herein. A kit may contain any of the compositions provided in this disclosure, including those recited above.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 52

[00165] **Fig. 6** shows schematics of exemplary kits of the invention. With reference to **Fig. 6**, a kit comprising a kit body **601** is shown that contains reagents for the detection of analytes (e.g., polynucleotide probes comprising a fluorophore and a quencher) **602** and primers for the amplification of a nucleic acid **603.** The kit may also contain any other reagents, such as reagents suitable for performing an amplification reaction **604.** In the embodiment depicted in **601**, the reagents for the detection of analytes and the primers are each contained in a single volume (e.g., tube or a bottle). However, these reagents may also be provided in separate volumes, as depicted the kit body shown in **605.** The kit body **605** contains seven separate volumes. Three volumes (**606, 607,** and **608**) contain reagents for the detection of analytes (or mixtures of such reagents) and three volumes (**609, 610,** and **611**) contain primers (or mixtures of primers). The kit may also contain reagents suitable for performing an amplification reaction **604.**

[00166] **Fig. 6** is provided for illustrative purposes only. Any number of reagents for the detection of analytes (e.g., probes) and any number of primers may be included in a single tube or bottle, as appropriate for the application. For example, in some embodiments, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more reagents for the detection of analytes may be included in a single bottle. In other cases, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more primers for amplification may be included in a single bottle or tube. In some cases, primers and probes may be provided in the same bottle or tube. For example, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more probes may be provided per pair of primers. In other cases, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more primers may be provided per pair of probes. The kits may contain 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more bottles or tubes. Control analytes, and their associated detection reagents and primers, may also be included in the kit. These may be included in a separate bottle or tube or, for example, included within the bottle or tube providing reagents for an amplification reaction **604**. As described elsewhere in the specification, any reagent in the kits may be lyophilized.

VII. Systems and Software

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 53

[00167] The methods of this disclosure may also be performed as part of a system. For example, one or more steps of the methods described herein may be performed by software contained on a computer readable medium. The software may be used to encode analytes, check for degeneracy, reduce or eliminate degeneracy, and/or produce a decoding matrix. The software may also be used to convert cumulative measurements made on an instrument (e.g., a PCR machine) into information concerning the presence or absence of an analyte. For example, with reference to **Fig. 4**, a computer **401** may be connected to an input device **403**, a display **402**, and a printer **404**. The computer may execute one or more of the steps of any of the methods described in this disclosure, including, for example, encoding methods **410**, analysis methods **420**, and decoding methods **430.** A computer may reduce or eliminate degeneracy, for example as part of the encoding methods **410.** Similarly, a computer may generate a coding schme that is non-degenerate by design using, for example, mathematical iteration. A computer may also be connected to a network **405**. The network may be a local network and/or the Internet. The network may be used to perform the methods of the invention on multiple computers in different locations. For example, one computer may perform the encoding, another computer may perform the measuring, and yet another computer may perform the decoding. A computer may perform any or all of these methods. The network may be used to transmit information obtained using the methods described in this disclosure throughout the world, for example across a border.

[00168] In one example, the software is embedded in the instrument making the measurement, such as a thermal cycler. In another example, the software is installed on a computer attached to the instrument making the measurement, such as a computer attached to a thermal cycler. In another example, the software is in the "cloud" -- i.e., on a computer that another computer may communicate with through a network.

[00169] Each of the elements described above may be connected to a controller that communicates with each of the elements and coordinates their actions. For example, a controller may initiate the encoding of a series of analytes into values of color and intensity. As described throughout this disclosure, the controller can execute software to reduce or eliminate degeneracy by removing one or more potential analyte codes from the encoding matrix. The controller may then automatically order probes corresponding to this code from a manufacturer of probes. Alternatively, the controller may operate an oligonucleotide synthesis instrument that can synthesize the appropriate probes. After analysis of the sample on the instrument (which may also be automated), the controller may process the results of the analysis using the decoding matrix. The controller may then provide these results to a user.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 54

VIII. Services

[00170] The methods provided herein may also be performed as a service. For example, a service provider may obtain the identity of a plurality of analytes that a customer wishes to analyze. The service provider may then encode each analyte to be detected by any of the methods described herein and provide appropriate reagents to the customer for the assay. The customer may perform the assay and provide the results to the service provider for decoding. The service provider may then provide the decoded results to the customer. The customer may also encode analytes, generate probes, and/or decode results by interacting with software installed locally (at the customer's location) or remotely (e.g., on a server reachable through a network). Exemplary customers include clinical laboratories, physicians, manufacturers of food and consumer products, industrial manufacturers (e.g., petroleum companies) and the like. A customer or party may be any suitable customer or party with a need or desire to use the methods, systems, compositions, and kits of the invention.

EXAMPLES

Example 1: General Materials & Methods

1. General Reagents

[00171] TE Buffer, pH7 (LIFE TECHNOLOGIES, Carlsbad, CA); UltraPure RNAse-free Water (LIFE TECHNOLOGIES, Carlsbad, CA); Taq 5x Master Mix (FISHER SCIENTIFIC COMPANY, Tustin, CA).

2. DNA Sequences, Primers, and Probes

[00172] Five nucleic acids from organisms of clinical relevance were chosen for the exemplary study: (1) human immunodeficiency virus 1 (HIV-1); (2) *Plasmodium falciparum* (Malaria); (3) herpes simplex virus-2 (HSV-2); (4) *Mycobacterium tuberculosis* (TB); (5) and dengue virus type 3 (dengue fever).

[00173] Two regions of diagnostic relevance from the HIV-1 genome, p17 and poly-protease, were selected from the Los Alamos National Laboratory HIV-1 reference sequence. A diagnostic sequence from the *Plasmodium falciparum* ChR7 gene was obtained from the UCSC *Plasmodium falciparum* Genome Browser. A sequence for Herpes Simplex Virus-2 was synthesized from the

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 55

sequence obtained from the European Molecular Biology Library. Similarly a diagnostic sequence for the rpoB gene in *Mycobacterium tuberculosis* was synthesized from a sequence obtained from the European Molecular Biology Library. A PCR diagnostic sequence for Dengue Virus Type 3 was obtained from the National Institute of Health genetic sequence database.

[00174] Oligonucleotides were synthesized by INTEGRATED DNA TECHNOLOGIES (IDT). Probes and primer pairs were chosen for each analyte using OLIGOANALYZER, from IDT. Sense probes containing a fluorophore at the 5' end and a quencher at the 3' end were synthesized for all analytes. All oligonucleotides were lyophilized and reconstituted in TE buffer before use. Tables 9 - 14 show sequence information, target sequence, primers, and probes for each of the six analytes.

Table 9. HIV-1 polyprotease sequence information.

Sequence Information	HIV-1 Poly protease 198mer synthesized from bases 2253 – 2550
Source	HIV-1 Reference Sequence, Los Alamos National Laboratory
Target sequence (3' to 5')	GGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGA GTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTT ATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAA AGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGG
Forward Primer (3' to 5')	GGAAGCTCTATTAGATACAGGAGCAG
Reverse Primer (3' to 5')	CCAATTATGTTGACAGGTGTAGGTCC
Probe 1 (3' to 5')	/**56-FAM**/TGAGTTTGCCAGGAAGATGGAAACCA/3BHQ_1/

Table 10. HIV p17 sequence information.

Sequence Name	HIV-1 P17 199mer synthesized from bases 790 – 1186
Source	HIV-1 Reference Sequence, Los Alamos National Laboratory
Target sequence (3' to 5')	CAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATA TAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAG ACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAG AAAAAAGCACAGCAAGCAGCAGCTGACACAGGACACAGCAATCAGGTCA
Forward Primer (3' to 5')	CAGCTACAACCATCCCTTCAGACA
Reverse Primer (3' to 5')	TGACCTGATTGCTGTGTCCTGTGT
Probe 1 (3' to 5')	/**56-FAM**/AGCAACCCTCTATTGTGTGCATCAAAGG/3BHQ_1

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 56

Probe 2 (3' to 5')	/**5Cy3**/AAAGCACAGCAAGCAGCAGCTGA/3BHQ_2/

Table 11. *Plasmodium* ChR7 sequence information.

Sequence Name	*Plasmodium* ChR7 199mer synthesized from bases 1139138 - 1141223
Source	UCSC Plasmodium falciparum Genome Browser Gateway
Target sequence (3' to 5')	GCCTAACATGGCTATGACGGGTAACGGGGAATTAGAGTTCGATTCCGGAG AGGGAGCCTGAGAAATAGCTACCACATCTAAGGAAGGCAGCAGGCGCGTA AATTACCCAATTCTAAAGAAGAGAGGTAGTGACAAGAAATAACAATGCAA GGCCAATTTAAAACCTTCCCAGAGTAACAATTGGAGGGCAAGTCTGGTG
Forward Primer (3' to 5')	GCCTAACATGGCTATGACGGGTAA
Reverse Primer (3' to 5')	CACCAGACTTGCCCTCCAATTGTT
Probe 1 (3' to 5')	/**56-FAM**/ATTCCGGAGAGGGAGCCTGAGAAATA/3BHQ_1/
Probe 2 (3' to 5')	/**56-ROXN**/AAGGAAGGCAGCAGGCGCGTAAATTA/3BHQ_2/

Table 12. Herpes simplex virus 2 sequence information.

Sequence Name	HSV-2 193mer synthesized from HSV-2 genome
Source	EMBL Bank AJ303204
Target sequence (3' to 5')	TCAGCCCATCCTCCTTCGGCAGTATGGAGGGTGTCGCGGCGGCGAGCCGC CGTCCCCAAAGACGTGCGGGTCGTACACGTACACGTACCAGGGCGGCGGG CCTCCGACCCGGTACGCTCTCGTAAATGCTTCCCTGCTGGTGCCGATCTG GGACCGCGCCGCGGAGACATTCGAGTACCAGATCGAACTCGG
Forward Primer (3' to 5')	TCAGCCCATCCTCCTTCGGCAGTAT
Reverse Primer (3' to 5')	CCGAGTTCGATCTGGTACTCGAATGT
Probe 1 (3' to 5')	/**56-FAM**/AAAGACGTGCGGGTCGTACACGTACA/3BHQ_1/
Probe 2 (3' to 5')	/**5Cy5**/TAAATGCTTCCCTGCTGGTGCCGAT/3IAbRQSp/

Table 13. *Mycobacterium tuberculosis* rpoB sequence information

Sequence Name	Mycobacterium Tuberculosis rpoB 200mer synthesized from Mycobacterium tuberculosis rpoB genome
Source	EMBL Bank GQ395623

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 57

Target sequence (3' to 5')	GAGTGCAAAGACAAGGACATGACGTACGCGGCCCCGCTGTTCGTCACGGC CGAGTTCATCAACAACAACACCGGCGAGATCAAGAGCCAGACGGTGTTCA TGGGTGACTTCCCGATGATGACCGAGAAGGGCACCTTCATCATCAACGGC ACCGAGCGCGTCGTGGTCAGCCAGCTGGTCCGCTCGCCCGGTGTGTACTT
Forward Primer (3' to 5')	GAGTGCAAAGACAAGGACATGACG
Reverse Primer (3' to 5')	AAGTACACACCGGGCGAGC
Probe 1 (3' to 5')	/**56-FAM**/CGCTGTTCGTCACGGCCGAGTTCAT/3BHQ_1/
Probe 2 (3' to 5')	/**5Cy3**/AGATCAAGAGCCAGACGGTGTTCATG/3BHQ_2/
Probe 3 (3' to 5')	/**5Cy5**/AAGGGCACCTTCATCATCAACGGCA/3IAbRQSp/

Table 14. Dengue virus type 3 sequence information.

Sequence Name	Dengue Virus Type 3 200mer synthesized from Dengue Virus Type 3 genome
Source	GenBank M93130
Target sequence (3' to 5')	ATGCCAACTGTGATTGAGCACTTAGAAAGACTACAAAGGAAACATGGAGG AATGCTTGTGAGAAATCCACTCTCACGAAACTCCACGCACGAAATGTATT GGATATCCAATGGTACAGGCAACATCGTCTCATTCACAATGACACACAGG AGACCCACCATAGAGAAAGATGTGGATTTAGGAGCAGGAACCCGACATGT
Forward Primer (3' to 5')	ATGCCAACTGTGATTGAGCACT
Reverse Primer (3' to 5')	ACATGTCGGGTTCCTGCTCCTAAA
Probe 1 (3' to 5')	/**56-FAM**/ACAAAGGAAACATGGAGGAATGCTTGTGA/3BHQ_1/
Probe 2 (3' to 5')	/**5Cy3**/ACTCTCACGAAACTCCACGCACGAAA/3BHQ_2/
Probe 3 (3' to 5')	/**56-ROXN**/ACAATGACACACAGGAGACCCACCAT/3BHQ_2/
Probe 4 (3' to 5')	/**5Cy5**/GGATATCCAATGGTACAGGCAACATCGT/3IAbRQSp/

3. Polymerase Chain Reactions

[00175] All PCR reactions were performed in a ROCHE 480 LIGHTCYCLER instrument. Forty-five thermal cycles were performed, with a 60 second hot-start at 95°C. The cycling conditions for denaturation, annealing, and extension were 45 seconds, 50 seconds, and 60 seconds at 95°C, 65°C, and 70°C, respectively. Each experiment was run in quintuplicate, with a reaction volume of 15 μL. Fluorescence measurements were obtained in the following channels after every annealing step: 483 nm to 533 nm (FAM), 523 nm to 568 nm (Cy3), 558 nm to 610 nm (ROX), and 615 nm to 670 nm (Cy5).

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 58

Fluorescence measurements were also obtained after the hot-start and the end of the 45 thermal cycles. The change in the fluorescence intensity between these two measurements (after hot-start and at the end of the 45 cycles) determined the end-point fluorescent signal.

Example 2: Simultaneous Detection of Polynucleotide Markers from Pathogens Based on an Encoding Method Using More Than One Color Per Analyte

[00176] This example describes the simultaneous detection of up to four polynucleotides selected from HIV (2 polynucleotides), HSV-2, *Mycobacterium*, *Plasmodium*, and dengue virus. Fourteen PCR reactions were assembled, using the reagents described in Table 15.

Table 15. Reagents for 14 PCR reactions.

Reaction No.	Type of Reagent	Reagent	Concentration	Volume
1		UltraPure Water	-	34 µL
		Taq 5x Master Mix	-	42 µL
	Template	Poly Protease Template	17 nM	2 µL
	Primers	Poly Protease FWD Primer	1 nM	2 µL
		Protease RVS Primer	1 nM	2 µL
	Probe	Poly Protease Probe 1	1 nM	2 µL
2		UltraPure Water	-	28 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Malaria Template	15 nM	2 µL
		Herpes Template	17 nM	2 µL
		TB Template	16 nM	2 µL
		Dengue Virus Template	17 nM	2 µL
	Primers	Poly Protease FWD Primer	1 nM	2 µL
		Protease RVS Primer	1 nM	2 µL
	Probe	Poly Protease Probe 1	1 nM	2 µL
3		UltraPure Water	-	32 µL
		Taq 5x Master Mix	-	42 µL
	Template	Malaria Template	15 nM	2 µL
	Primers	Malaria FWD Primer	1 µM	2 µL
		Malaria RVS Primer	1 µM	2 µL
	Probes	Malaria Probe 1	1 µM	2 µL
		Malaria Probe 2	1 µM	2 µL
4		UltraPure Water	-	26 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Herpes Template	17 nM	2 µL
		TB Template	16 nM	2 µL
		Dengue Virus Template	17 nM	2 µL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 59

	Primers	Malaria FWD Primer	1 µM	2 µL
		Malaria RVS Primer	1 µM	2 µL
	Probes	Malaria Probe 1	1 µM	2 µL
		Malaria Probe 2	1 µM	2 µL
5		UltraPure Water	-	32 µL
		Taq 5x Master Mix	-	42 µL
	Template	Herpes Template	17 nM	2 µL
	Primers	Herpes FWD Primer	1 µM	2 µL
		Herpes RVS Primer	1 µM	2 µL
	Probes	Herpes Probe 1	1 µM	2 µL
		Herpes Probe 2	1 µM	2 µL
6		UltraPure Water	-	26 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Malaria Template	15 nM	2 µL
		TB Template	16 nM	2 µL
		Dengue Virus Template	17 nM	2 µL
	Primers	Herpes FWD Primer	1 µM	2 µL
		Herpes RVS Primer	1 µM	2 µL
	Probes	Herpes Probe 1	1 µM	2 µL
		Herpes Probe 2	1 µM	2 µL
7		UltraPure Water	-	30 µL
		Taq 5x Master Mix	-	42 µL
	Template	TB Template	16 nM	2 µL
	Primers	TB FWD Primer	1 µM	2 µL
		TB RVS Primer	1 µM	2 µL
	Probes	TB Probe 1	1 µM	2 µL
		TB Probe 2	1 µM	2 µL
		TB Probe 3	1 µM	2 µL
8		UltraPure Water	-	30 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Malaria Template	15 nM	2 µL
		Herpes Template	17 nM	2 µL
		Dengue Virus Template	17 nM	2 µL
	Primers	TB FWD Primer	1 µM	2 µL
		TB RVS Primer	1 µM	2 µL
	Probes	TB Probe 1	1 µM	2 µL
		TB Probe 2	1 µM	2 µL
		TB Probe 3	1 µM	2 µL
9		UltraPure Water	-	28 µL
		Taq 5x Master Mix	-	42 µL
	Template	Dengue Virus Template	17 nM	2 µL
	Primers	Dengue Virus FWD Primer	1 µM	2 µL
		Dengue Virus RVS Primer	1 µM	2 µL
	Probes	Dengue Virus Probe 1	1 µM	2 µL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 60

		Dengue Virus Probe 2	1 µM	2 µL
		Dengue Virus Probe 3	1 µM	2 µL
		Dengue Virus Probe 4	1 µM	2 µL
10		UltraPure Water	-	22 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Malaria Template	15 nM	2 µL
		Herpes Template	17 nM	2 µL
		TB Template	16 nM	2 µL
	Primers	Dengue Virus FWD Primer	1 µM	2 µL
		Dengue Virus RVS Primer	1 µM	2 µL
	Probes	Dengue Virus Probe 1	1 µM	2 µL
		Dengue Virus Probe 2	1 µM	2 µL
		Dengue Virus Probe 3	1 µM	2 µL
		Dengue Virus Probe 4	1 µM	2 µL
11		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Malaria Template	15 nM	2 µL
		Herpes Template	17 nM	2 µL
		Dengue Virus Template	17 nM	2 µL
	Primers	Poly Protease FWD Primer	1 µM	2 µL
		Poly Protease RVS Primer	1 µM	2 µL
		Malaria FWD Primer	1 µM	2 µL
		Malaria RVS Primer	1 µM	2 µL
		Herpes FWD Primer	1 µM	2 µL
		Herpes RVS Primer	1 µM	2 µL
		Dengue Virus FWD Primer	1 µM	2 µL
		Dengue Virus RVS Primer	1 µM	2 µL
	Probes	Poly Protease Probe 1	1 µM	2 µL
		Malaria Probe 1	1 µM	2 µL
		Malaria Probe 2	1 µM	2 µL
		Herpes Probe 1	1 µM	2 µL
		Herpes Probe 2	1 µM	2 µL
		Dengue Virus Probe 1	1 µM	2 µL
		Dengue Virus Probe 2	1 µM	2 µL
		Dengue Virus Probe 3	1 µM	2 µL
		Dengue Virus Probe 4	1 µM	2 µL
12		UltraPure Water	-	2 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Malaria Template	15 nM	2 µL
		Herpes Template	17 nM	2 µL
		TB Template	16 nM	2 µL
	Primers	Poly Protease FWD Primer	1 µM	2 µL
		Poly Protease FWD Primer	1 µM	2 µL
		Malaria FWD Primer	1 µM	2 µL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 61

		Malaria RVS Primer	1 µM	2 µL
		Herpes FWD Primer	1 µM	2 µL
		Herpes RVS Primer	1 µM	2 µL
		TB FWD Primer	1 µM	2 µL
		TB RVS Primer	1 µM	2 µL
	Probes	Poly Protease Probe 1	1 µM	2 µL
		Malaria Probe 1	1 µM	2 µL
		Malaria Probe 2	1 µM	2 µL
		Herpes Probe 1	1 µM	2 µL
		Herpes Probe 2	1 µM	2 µL
		TB Probe 1	1 µM	2 µL
		TB Probe 2	1 µM	2 µL
		TB Probe 3	1 µM	2 µL
13		UltraPure Water	-	12 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Herpes Template	17 nM	2 µL
		TB Template	16 nM	2 µL
	Primers	Poly Protease FWD Primer	1 µM	2 µL
		Poly Protease FWD Primer	1 µM	2 µL
		Herpes FWD Primer	1 µM	2 µL
		Herpes RVS Primer	1 µM	2 µL
		TB FWD Primer	1 µM	2 µL
		TB RVS Primer	1 µM	2 µL
	Probes	Poly Protease Probe 1	1 µM	2 µL
		Herpes Probe 1	1 µM	2 µL
		Herpes Probe 2	1 µM	2 µL
		TB Probe 1	1 µM	2 µL
		TB Probe 2	1 µM	2 µL
		TB Probe 3	1 µM	2 µL
14		UltraPure Water	-	10 µL
		Taq 5x Master Mix	-	42 µL
	Templates	Poly Protease Template	17 nM	2 µL
		Malaria Template	15 nM	2 µL
		Dengue Virus Template	17 nM	2 µL
	Primers	Poly Protease FWD Primer	1 µM	2 µL
		Poly Protease FWD Primer	1 µM	2 µL
		Malaria FWD Primer	1 µM	2 µL
		Malaria RVS Primer	1 µM	2 µL
		Dengue Virus FWD Primer	1 µM	2 µL
		Dengue Virus RVS Primer	1 µM	2 µL
	Probes	Poly Protease Probe 1	1 µM	2 µL
		Malaria Probe 1	1 µM	2 µL
		Malaria Probe 2	1 µM	2 µL
		Dengue Virus Probe 1	1 µM	2 µL
		Dengue Virus Probe 2	1 µM	2 µL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 62

		Dengue Virus Probe 3	1 μM	2 μL
		Dengue Virus Probe 4	1 μM	2 μL

[00177] Referring to Table 15, reactions 1, 3, 5, 7, and 9 (containing HIV polyprotease, *Plasmodium*, HSV, *Mycobacterium*, and dengue analytes, respectively) were positive controls that provided baseline fluorescence intensity for each polynucleotide analytes alone, in the absence of other analytes. The change in fluorescence intensity in each color was used to provide the expected cumulative intensity level for each color in the chromatogram **(Fig. 3)**, as described more fully below. Reactions 2, 4, 6, 8, and 10 were negative controls used to determine the extent of non-specific amplification. In these reactions, the primers provided were not specific for the analytes provided.

[00178] Experiments 11, 12, 13, and 14 were multiplex detection experiments. Encoding was performed as indicated in Table 16. The measured fluorescence intensity for each assay, in each color, was plotted as black circles in the chromatogram of **Fig. 3.**

Table 16. Coding scheme for detection of pathogens.

	FAM	**Cy3**	**ROX**	**Cy5**
HIV poly-protease	1	0	0	0
HIV p17	1	1	0	0
Plasmodium	1	0	1	0
HSV	1	0	0	1
Mycobacterium	1	1	0	1
Dengue Virus	1	1	1	1

[00179] To construct the chromatogram, the signals of the positive controls (reactions 1, 3, 5, 7, and 9) were used to assemble every possible combination of present sequences, in each color. The cumulative signals were plotted in a diagram, called a "chromatogram" (**Fig. 3**). Cumulative signals corresponding to all possible combinations of the same rank in the same color were organized into their own "band." Doing this for all four colors produced a level and band structure, against which the experimental results of each combination (reactions 11, 12, 13, and 14) could be evaluated. **Fig. 3** shows three replicates of a chromatogram constructed using the positive control samples, with experimental results 4112, 3121, and 3102 superimposed on the chromatogram. The experimental results are indicated by black dots.

[00180] The intensity value for each color in reactions 11-14 was measure in arbitrary fluorescence units (AFU), plotted on the chromatogram, and then assigned to a band based on two criteria. First, if an experimental result was within a band, then the result was assigned to the

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 63

multiplicity of the band (e.g., the Cy3 value of 1 in the left-most chromatogram of **Fig. 3**). Second, if an experimental result was between two bands, a two-step analysis was employed. If the experimental result was between two legitimate results, the experimental result was assigned to the closest band. (e.g., the Cy5 value of 2 in the right-most chromatogram of **Fig. 3**). If the experimental result was between a legitimate result and an illegitimate result, the result was assigned to the band with the legitimate result (e.g., the FAM value of 3 in the middle chromatogram of **Fig. 3**). **Fig. 8** schematically shows a legitimate result of 4112 (top) indicating the presence of HIV polyprotease, *Plasmodium*, herpes simplex, and *Mycobacterium*; and an illegitimate result of 4110 (bottom), which cannot be decoded and indicates a malfunction in the assay.

[00181] Using these criteria, the output of Experiment 12 was designated 4112, which was the correct answer. The output of Experiment 14 could have been designated as 2121 or 3121, but 2121 is illegitimate, so the result must be 3121, which is the correct answer. The output of Experiment 13 could be designated as 3102 or 3101, both of which are legitimate, but the result is clearly closer to 2 than 1 in the Cy5 channel, so 3102, the correct answer, was designated as the result. These results provide experimental validation of the methods of the provided herein.

[00182] The results indicate a trend of slightly lower combined signal than what would be expected from a simple summation of positive-control signals. This may be because the combination of signals is not perfectly linear. This may be a result of additional effects for which the chromatogram construction does not account for at present. These effects may be related to the absolute concentration of quenchers. The higher the absolute concentration of quenchers, the larger the percentage of reaction volume they quench, so that the same percentage of released fluorophores fails to contribute as much as expected to the cumulative signal. One solution to this problem is to use low concentrations of probes, so that the percentile loss in signal is negligible, regardless of how much quencher is released.

[00183] **Fig. 3** shows that each rank containing multiple combinations has a relatively wide band. The width of the band is ultimately the difference between the highest and lowest cumulative signals for the positive control samples. If all positive controls in the particular color produced exactly the same fluorescence signal, then the width of each band would be zero. Since instead those signals are somewhat different, the resulting bands are relatively wide. However, there is a simple means to tighten the bands. Instead of loading all probes at the same concentration, as was done for the experiment above, the concentration of each probe can be adjusted so that each of the resulting positive-control signals is the same as the others in the same color. This approach would significantly tighten the bands, making it easier to determine the multiplicity of a signal for a sample containing analytes.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 64

Example 3: Simultaneous Detection of Polynucleotide Markers from Pathogens Based on an Encoding Method Using One Color Per Analyte and Different Intensities

[00184] This example describes the simultaneous detection of up to three polynucleotides selected from dengue virus, HIV p17, and HIV polyprotease. The coding scheme was as depicted in Table 11. As shown in Table 17, each analyte was encoded as a single value of fluorescence intensity in a single color. In this case, as described elsewhere in this disclosure, the coding scheme is non-degenerate by design. This is achieved by assigning an intensity value to each analyte that is equal to the cumulative intensity values for the prior analytes, plus one. Of course, as described elsewhere in the specification, the initial assigned value can be greater than one, for example to better distinguish over noise. Similarly, the values assigned to each analyte can be greater than the cumulative intensity values for the prior analytes plus one, for example, to increase the separation between the intensities and enable more simple assignment of intensities.

Table 17. Coding scheme for detection of dengue virus, HIV polyprotease, and HIV p17.

	FAM
Dengue Virus	1
HIV p17	2
HIV polyprotease	4

[00185] The same reagents (e.g., templates, primers, and probes) used in Example 2 were used here. The fluorescence signals generated by the respective probes in positive-control end-point PCR reactions were measured and used to calculate the probe concentrations that would produce 1x, 2x, and 4x signal intensities. Experiments were then set up to detect the seven non-null combinations of analyte occurrences. Tables 18-24 provide the PCR reaction components for each of these mixtures.

Table 18. Binary FAM Experiment 1 Cocktail: HIV Polyprotease, HIV p17, and dengue virus

Reagents	Concentration	Volume Added
UltraPure Water	-	50μL
Taq 5x Master Mix	-	25μL
Templates	**Concentration**	**Volume Added**
Poly Protease Template	17nM	4μL
P17 Template	15nM	4μL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 65

Dengue Virus Template	17nM	4µL
Primers	Concentration	Volume Added
Poly Protease FWD Primer	1µM	4µL
Poly Protease RVS Primer	1µM	4µL
P17 FWD Primer	1µM	4µL
P17 RVS Primer	1µM	4µL
Dengue Virus FWD Primer	1µM	4µL
Dengue Virus RVS Primer	1µM	4µL
Probes	Concentration	Volume Added
Poly Protease Probe 1	800nM	4µL
P17 Probe 1	400nM	4µL
Dengue Virus Probe 1	200nM	4µL

Table 19. Binary FAM Experiment 2 Cocktail: dengue virus and HIV p17

Reagents	Concentration	Volume Added
UltraPure Water	-	54µL
Taq 5x Master Mix	-	25µL
Templates	Concentration	Volume Added
Poly Protease Template	17nM	-
P17 Template	15nM	4µL
Dengue Virus Template	17nM	4µL
Primers	Concentration	Volume Added
Poly Protease FWD Primer	1µM	4µL
Poly Protease RVS Primer	1µM	4µL
P17 FWD Primer	1µM	4µL
P17 RVS Primer	1µM	4µL
Dengue Virus FWD Primer	1µM	4µL
Dengue Virus RVS Primer	1µM	4µL
Probes	Concentration	Volume Added
Poly Protease Probe 1	800nM	4µL
P17 Probe 1	400nM	4µL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 66

Dengue Virus Probe 1	200nM	4µL

Table 20. Binary FAM Experiment 3 Cocktail: dengue virus and HIV polyprotease

Reagents	Concentration	Volume Added
UltraPure Water	-	54µL
Taq 5x Master Mix	-	25µL
Templates	Concentration	Volume Added
Poly Protease Template	17nM	4µL
P17 Template	15nM	-
Dengue Virus Template	17nM	4µL
Primers	Concentration	Volume Added
Poly Protease FWD Primer	1µM	4µL
Poly Protease RVS Primer	1µM	4µL
P17 FWD Primer	1µM	4µL
P17 RVS Primer	1µM	4µL
Dengue Virus FWD Primer	1µM	4µL
Dengue Virus RVS Primer	1µM	4µL
Probes	Concentration	Volume Added
Poly Protease Probe 1	800nM	4µL
P17 Probe 1	400nM	4µL
Dengue Virus Probe 1	200nM	4µL

Table 21. Binary FAM Experiment 4 Cocktail: HIV polyprotease and HIV p17

Reagents	Concentration	Volume Added
UltraPure Water	-	54µL
Taq 5x Master Mix	-	25µL
Templates	Concentration	Volume Added
Poly Protease Template	17nM	4µL
P17 Template	15nM	4µL
Dengue Virus Template	17nM	-
Primers	Concentration	Volume Added

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 67

Poly Protease FWD Primer	1μM	4μL
Poly Protease RVS Primer	1μM	4μL
P17 FWD Primer	1μM	4μL
P17 RVS Primer	1μM	4μL
Dengue Virus FWD Primer	1μM	4μL
Dengue Virus RVS Primer	1μM	4μL
Probes	Concentration	Volume Added
Poly Protease Probe 1	800nM	4μL
P17 Probe 1	400nM	4μL
Dengue Virus Probe 1	200nM	4μL

Table 22. Binary FAM Experiment 5 Cocktail: dengue virus

Reagents	Concentration	Volume Added
UltraPure Water	-	58μL
Taq 5x Master Mix	-	25μL
Templates	Concentration	Volume Added
Poly Protease Template	17nM	-
P17 Template	15nM	-
Dengue Virus Template	17nM	4μL
Primers	Concentration	Volume Added
Poly Protease FWD Primer	1μM	4μL
Poly Protease RVS Primer	1μM	4μL
P17 FWD Primer	1μM	4μL
P17 RVS Primer	1μM	4μL
Dengue Virus FWD Primer	1μM	4μL
Dengue Virus RVS Primer	1μM	4μL
Probes	Concentration	Volume Added
Poly Protease Probe 1	800nM	4μL
P17 Probe 1	400nM	4μL
Dengue Virus Probe 1	200nM	4μL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 68

Table 23. Binary FAM Experiment 6 Cocktail: HIV p17

Reagents	Concentration	Volume Added
UltraPure Water	-	58µL
Taq 5x Master Mix	-	25µL
Templates	**Concentration**	**Volume Added**
Poly Protease Template	17nM	-
P17 Template	15nM	4µL
Dengue Virus Template	17nM	-
Primers	**Concentration**	**Volume Added**
Poly Protease FWD Primer	1µM	4µL
Poly Protease RVS Primer	1µM	4µL
P17 FWD Primer	1µM	4µL
P17 RVS Primer	1µM	4µL
Dengue Virus FWD Primer	1µM	4µL
Dengue Virus RVS Primer	1µM	4µL
Probes	**Concentration**	**Volume Added**
Poly Protease Probe 1	800nM	4µL
P17 Probe 1	400nM	4µL
Dengue Virus Probe 1	200nM	4µL

Table 24. Binary FAM Experiment 7 Cocktail: HIV polyprotease

Reagents	Concentration	Volume Added
UltraPure Water	-	58µL
Taq 5x Master Mix	-	25µL
Templates	**Concentration**	**Volume Added**
Poly Protease Template	17nM	4µL
P17 Template	15nM	-
Dengue Virus Template	17nM	-
Primers	**Concentration**	**Volume Added**
Poly Protease FWD Primer	1µM	4µL
Poly Protease RVS Primer	1µM	4µL

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 69

P17 FWD Primer	1μM	4μL
P17 RVS Primer	1μM	4μL
Dengue Virus FWD Primer	1μM	4μL
Dengue Virus RVS Primer	1μM	4μL
Probes	Concentration	Volume Added
Poly Protease Probe 1	800nM	4μL
P17 Probe 1	400nM	4μL
Dengue Virus Probe 1	200nM	4μL

[00186] The total fluorescence signal of each experiment was plotted on a chromatogram **(Fig. 7)** that was constructed as described above. The width of the bands around each positive control measurement is the propagated uncertainty of the 1X measurement. That uncertainty was equated to the standard deviation of the fluorescence signals of the last five PCR cycles in saturation.

[00187] **Fig. 7** shows the results of this experiment. The three right-most results show the measured intensities for dengue virus alone (DEN), HIV p17 alone (P17), and HIV polyprotease (TPP). The assigned intensities for each of these analytes were as indicated in Table 11, and **Fig. 7** confirms that the three analytes yielded the expected intensity results when alone (see DEN, P17, and TPP). At the end of the assay, essentially all of the probes have been hydrolyzed, releasing essentially all of the fluorophores, which emit a signal. A fixed amount of a fluorophore produces a fixed and predictable amount of signal. Therefore a target that is present should generally produce a fixed amount of signal, as determined by the amount of probe added to the reaction volume.

[00188] As expected, the intensity values for the combinations of analytes are equivalent to the sum of the assigned intensities. Since this coding scheme is designed to be non-degenerate, each result can be unambiguously decoded into the presence or absence of a particular analyte. For example, starting from the left side of **Fig. 7**, a result of 7 indicates the presence of all three analytes. A result of 3 indicates the presence of dengue virus and HIV p17 analytes, and the absence of HIV polyprotease. A result of 5 indicates the presence of dengue virus and HIV polyprotease, and the absence of HIV p17. Finally a result of 6 indicates the presence of HIV p17 and HIV polyprotease, and the absence of dengue virus. This coding scheme can be extended to encode additional analytes, as described elsewhere in this disclosure.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 70

CLAIMS

WHAT IS CLAIMED IS:

1. An assay that is capable of unambiguously detecting the presence or absence of each of at least 7 analytes, in any combination of presence or absence, in a single sample volume without immobilization, separation, mass spectrometry, or melting curve analysis.
2. The assay of claim 1, wherein each of said analytes is encoded as a value of one component of a signal.
3. The method of claim 2, wherein said method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = \log_2 (F + 1)$ and F is the maximum cumulative value of said one component of said signal when all of said analytes are present.
4. The assay of claim 1, wherein each of said analytes is encoded as at least one first value in a first component of said signal and at least one second value in a second component of said signal.
5. The assay of claim 4, wherein each of said analytes is encoded as a first value in a first component of said signal at each of a plurality of second values in a second component of said signal.
6. The method of claim 4, wherein said method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = C * \log_2 (F + 1)$, where C is the number of said second values used to encode said analytes and F is the maximum cumulative value of said first component of said signal, for any second value, when all of said analytes are present.
7. The method of claim 4, wherein said method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = (P * T) + 1$, where P is the number of codes per tier in a coding scheme and T is the number of tiers.
8. The method of claim 7, wherein the number of tiers $T = \log_4 (F + 1)$, and F is the maximum cumulative value of said first component of said signal, for any second value, when all of said analytes are present.
9. The method of claim 4, wherein said first value is an intensity or range of intensities.
10. The method of claim 9, wherein said coding scheme comprises at least three intensities or ranges of intensities.
11. The method of claim 4, wherein said second value is a wavelength or range of wavelengths.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 71

12. The method of claim 11, wherein said coding scheme comprises at least five wavelengths or ranges of wavelengths.

13. The method of claim 1, wherein said signal is an electromagnetic signal.

14. The method of claim 13, wherein said electromagnetic signal is a fluorescence emission signal.

15. The method of claim 14, wherein intensity of said fluorescence emission signal is measured at at least four wavelengths or ranges of wavelengths.

16. The method of claim 1, wherein said detecting is performed with reagents that are lyophilized prior to use.

17. The method of claim 1, wherein said detecting is performed with reagents comprising hybridization probes.

18. The method of claim 17, wherein the number of said hybridization probes is greater than the number of said analytes.

19. The method of claim 17, wherein said hybridization probes comprise one or more hybridization probes specific for different analytes and comprising an identical fluorophore or combination of fluorophores.

20. The method of claim 17, wherein said sample is contacted with at least 18 of said hybridization probes.

21. The method of claim 17, wherein said reagents further comprise at least one pair of primers capable of amplifying a region complementary to at least three of said hybridization probes.

22. The method of claim 21, wherein said signal is generated during a polymerase chain reaction.

23. The method of claim 22, wherein said polymerase chain reaction is selected from the group consisting of an end-point polymerase chain reaction, a real-time polymerase chain reaction, and a digital polymerase chain reaction.

24. The method of claim 1, wherein said cumulative measurement is performed on a solution.

25. The method of claim 1, wherein at least one of said analytes is encoded by at least one additional value, wherein said at least one additional value is selected from the group consisting of a value from at least one additional component of said signal, a value from at least one component of an additional signal, and combinations thereof.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 72

26. The method of claim 25, wherein said at least one additional value is selected from the group consisting of a fluorescence emission intensity, a fluorescence emission wavelength, a Förster resonance energy transfer (FRET) emission intensity, a FRET emission wavelength, an electrochemical signal, a chemiluminescence wavelength, a chemiluminescence intensity, a fluorescence bleaching rate, a chemiluminescence bleaching rate, and combinations thereof.

27. The method of claim 1, further comprising constructing a chromatogram.

28. The method of claim 27, wherein said chromatogram is constructed by plotting all possible combinations of said first values and said second values for positive control samples.

29. The method of claim 1, wherein at least one of said analytes comprises a polynucleotide.

30. The method of claim 29, wherein said polynucleotide is from a source selected from the group consisting of an animal, a plant, a bacteria, a fungus, a parasite, and a virus.

31. The method of claim 23, wherein said polynucleotide is from a source selected from the group consisting of human immunodeficiency virus, herpes simplex virus, human papilloma virus, Plasmodium, Mycobacterium, dengue virus, hepatitis virus, and influenza virus.

32. The method of claim 31, wherein said polynucleotide is selected from the group consisting of human immunodeficiency virus polyprotease, human immunodeficiency virus p17, human papilloma virus E6, and human papilloma virus E7.

33. The method of claim 1, wherein said sample is selected from the group consisting of a clinical sample, a food sample, an environmental sample, a pharmaceutical sample, and a sample from a consumer product.

34. The method of claim 1, further comprising transmitting information concerning the presence or absence of said analyte through a computer network.

35. The method of claim 1, further comprising providing information concerning the presence or absence of said analyte to a physician.

36. The method of claim 35, further comprising making a clinical decision based on said information.

37. The method of claim 1, wherein at least one step of said method is performed using instructions on a computer readable medium.

38. The method of claim 37, wherein said instructions are located on a remote server.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 73

39. The method of claim 37, wherein said instructions are located on a thermal cycler.

40. The method of claim 37, wherein said instructions are located on a computer in communication with a thermal cycler.

41. The method of claim 1, wherein at least one of said analytes is a positive control analyte.

42. The method of claim 1, wherein said coding scheme is non-degenerate.

43. The method of claim 42, wherein said coding scheme is designed to be non-degenerate.

44. The method of claim 42, wherein said coding scheme is made non-degenerate by enumerating every legitimate result that can be obtained from said coding scheme, identifying each legitimate result that is degenerate, and eliminating at least one potential analyte code from said coding scheme to eliminate degeneracy.

45. The method of claim 1, wherein said assay is an end-point assay.

46. The method of claim 1, wherein said assay is ended at a threshold number of cycles set by the limit of detection of an instrument.

47. The method of claim 1, wherein said assay is a liquid phase assay.

48. The method of claim 1, wherein said assay is quantitative.

49. A method of detecting the presence or absence of each analyte of a plurality of analytes, comprising:
    a. encoding each of said analytes as a first value of a signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said first value, wherein said encoding is performed in a manner that eliminates degeneracy;
    b. providing a sample comprising, or potentially comprising, at least one of said analytes;
    c. contacting said sample with analyte-specific reagents that generate said first value, as represented in said coding scheme, when each of said analytes is present;
    d. cumulatively measuring said first values of said signal within said sample, thereby providing a cumulative measurement; and
    e. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme.

50. The method of claim 49, wherein said method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = \log_2 (F + 1)$ and F is the maximum cumulative value of said first values when all of said analytes are present.

51. The method of claim 49, wherein said first value is an intensity or range of intensities.

52. The method of claim 49, wherein said first value has a minimum value in said coding scheme that is

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 74

selected from the group consisting of at least 1, at least 2, at least 4, at least 8, at least 16, at least 32, and at least 64.

53. The method of claim 49, wherein said first value is incremented in said coding scheme by an amount equal to the cumulative maximum of said preceding first values plus one.

54. The method of claim 49, wherein said first value is incremented in said coding scheme by an amount greater than the cumulative maximum of said preceding first values plus one.

55. A method of detecting the presence or absence of each analyte of a plurality of analytes, comprising:

    a. encoding each of said analytes as at least one first value and at least one second value, wherein said first value is a value from a first component of a signal and said second value is a value from a second component of said signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said at least one first value and said at least one second value, wherein said encoding is performed in a manner that reduces or eliminates degeneracy;

    b. providing a sample comprising, or potentially comprising, at least one of said analytes;

    c. contacting said sample with analyte-specific reagents that generate said at least one first value and at least one said second value, as represented in said coding scheme, when each of said analytes is present;

    d. cumulatively measuring said first values and said second values within said sample, thereby providing a cumulative measurement; and

    e. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme,

    wherein, when degeneracy is eliminated, said method is capable of unambiguously detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

56. The method of claim 55, wherein each of said analytes is encoded as a first value in a first component of said signal at each of a plurality of second values in a second component of said signal.

57. The method of claim 55, wherein, when degeneracy is eliminated, said method is capable of unambiguously detecting the presence or absence of each of at least seven analytes in a single volume using four second values.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 75

58. The method of claim 55, wherein, when degeneracy is eliminated, said coding scheme comprises T non-degenerate tiers, wherein $T = \log_4 (F + 1)$ and F is the maximum cumulative value of said first component of said signal, for any second value, when all of said analytes are present.

59. The method of claim 56, wherein, when degeneracy is eliminated, said method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = (P * T) + 1$ and P is the number of codes per tier.

60. The method of claim 55, wherein, when degeneracy is eliminated, said method is capable of unambiguously detecting the presence or absence of M analytes, in any combination of presence or absence, where $M = C * \log_2 (F + 1)$, C is the number of second values in said coding scheme, and F is the maximum cumulative value of said first component of said signal, for any second value, when all of said analytes are present.

61. A non-degenerate coding scheme capable of unambiguously encoding the presence or absence of each of at least 7 analytes, in any combination of presence or absence, in a single sample volume without immobilization, separation, or melting curve analysis.

62. The non-degenerate coding scheme of claim 61, wherein said coding scheme is generated by a method comprising:

    a. generating a code for each potential analyte, wherein each potential analyte is encoded by at least one value of at least one component of a signal;

    b. enumerating every legitimate cumulative result for all possible combinations of presence or absence of each analyte;

    c. identifying each legitimate result that is degenerate; and

    d. eliminating at least one code to eliminate degeneracy.

63. The method of claim 62, further comprising expanding said coding scheme by adding additional codes that are at least one unit greater than the sum of all previous codes in said at least one component of said signal.

64. The method of claim 61, wherein said coding scheme is generated by a method of mathematical iteration that guarantees non-degeneracy by construction.

65. A system for detecting the presence or absence of each analyte of a plurality of analytes, comprising:

    a. encoding each of said analytes as a first value of a signal, thereby generating a coding

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 76

scheme, wherein each of said analytes is represented in said coding scheme by said first value, wherein said encoding is performed in a manner that eliminates degeneracy;

b. providing a sample comprising, or potentially comprising, at least one of said analytes;

c. contacting said sample with analyte-specific reagents that generate said first value, as represented in said coding scheme, when each of said analytes is present;

d. cumulatively measuring said first values of said signal within said sample, thereby providing a cumulative measurement; and

e. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme.

66. A system for detecting the presence or absence of each analyte of a plurality of analytes, comprising:

a. encoding each of said analytes as at least one first value and at least one second value, wherein said first value is a value from a first component of a signal and said second value is a value from a second component of said signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said at least one first value and said at least one second value, wherein said encoding is performed in a manner that reduces or eliminates degeneracy;

b. providing a sample comprising, or potentially comprising, at least one of said analytes;

c. contacting said sample with analyte-specific reagents that generate said at least one first value and said at least one second value, as represented in said coding scheme, when each of said analytes is present;

d. cumulatively measuring said first values and said second values within said sample, thereby providing a cumulative measurement; and

e. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme,

wherein, when degeneracy is eliminated, said system is capable of unambiguously detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

67. A method of detecting the presence or absence of each analyte of a plurality of analytes, comprising:

a. obtaining the identity of each of said analytes from a party;

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 77

b. encoding each of said analytes as a first value of a signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said first value, wherein said encoding is performed in a manner that eliminates degeneracy;

c. providing said party with analyte-specific reagents that generate said first value, as represented in said coding scheme, when each of said analytes is present, said party:

   i. contacting a sample comprising, or potentially comprising, at least one of said analytes with said reagents; and

   ii. cumulatively measuring said first values within said sample, thereby providing a cumulative measurement; and

d. obtaining said cumulative measurement from said party;

e. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme; and

f. providing said party with information about the presence or absence of each of said analytes.

68. A method of detecting the presence or absence of each analyte of a plurality of analytes, comprising:

g. obtaining the identity of each of said analytes from a party;

h. encoding each of said analytes as at least one first value and at least one second value, wherein said first value is a value from a first component of a signal and said second value is a value from a second component of said signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said at least one first value and said at least one second value, wherein said encoding is performed in a manner that reduces or eliminates degeneracy;

i. providing said party with analyte-specific reagents that generate said at least one first value and said at least one second value, as represented in said coding scheme, when each of said analytes is present, said party:

   i. contacting a sample comprising, or potentially comprising, at least one of said analytes with said reagents; and

   ii. cumulatively measuring said first values and said second values within said sample, thereby providing a cumulative measurement; and

j. obtaining said cumulative measurement from said party;

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 78

k. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme; and

l. providing said party with information about the presence or absence of each of said analytes.

69. A composition for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, each reagent generating a signal comprising a first value, wherein said reagents are capable of unambiguously detecting the presence or absence of each of at least seven analytes in a single sample volume, in any combination of presence or absence, without immobilization, separation, or melting curve analysis.

70. A composition for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, each reagent generating a signal comprising at least one first value that is a value from a first component of said signal and at least one second value that is a value from a second component of said signal, wherein said reagents are capable of detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

71. A kit for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, packaging, and instructions, each reagent generating a signal comprising a first value, wherein said kit is capable of unambiguously detecting the presence or absence of each of at least seven analytes in a single sample volume, in any combination of presence or absence, without immobilization, separation, or melting curve analysis.

72. A kit for detecting the presence or absence of each analyte of a plurality of analytes, comprising analyte-specific reagents, packaging, and instructions, each reagent generating a signal comprising at least one first value that is a value from a first component of said signal and at least one second value that is a value from a second component of said signal, wherein said kit is capable of detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

73. A kit for detecting the presence or absence of each analyte of a plurality of analytes, comprising a kit body (601), clamping slots arranged in the kit body for placing bottles, a bottle comprising probes for the detection of an analyte (602), a bottle comprising primers for amplification (603), and a bottle comprising reagents for amplification (604), wherein said kit is capable of unambiguously detecting

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 79

the presence or absence of each of at least seven analytes in a single sample volume, in any combination of presence or absence, without immobilization, separation, or melting curve analysis.

74. A kit for detecting the presence or absence of each analyte of a plurality of analytes, a kit body (601), clamping slots arranged in the kit body for placing bottles, a bottle comprising probes for the detection of an analyte (602), a bottle comprising primers for amplification (603), and a bottle comprising reagents for amplification (604), each probe generating a signal comprising at least one first value that is a value from a first component of said signal and at least one second value that is a value from a second component of said signal, wherein said kit is capable of detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said probes generates only one second value.

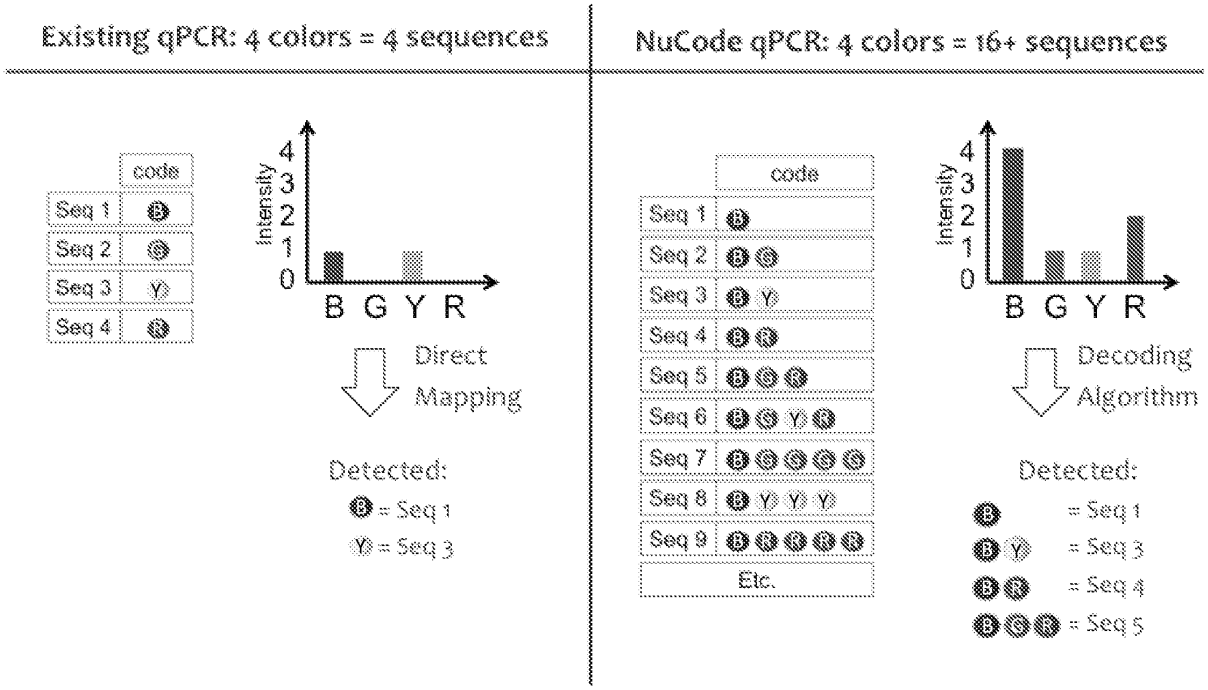
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 80

ABSTRACT

This disclosure provides methods, systems, compositions, and kits for the multiplexed detection of a plurality of analytes in a sample. In some examples, this disclosure provides methods, systems, compositions, and kits wherein multiple analytes may be detected in a single sample volume by acquiring a cumulative measurement or measurements of at least one quantifiable component of a signal. In some cases, additional components of a signal, or additional signals (or components thereof) are also quantified. Each signal or component of a signal may be used to construct a coding scheme which can then be used to determine the presence or absence of any analyte.

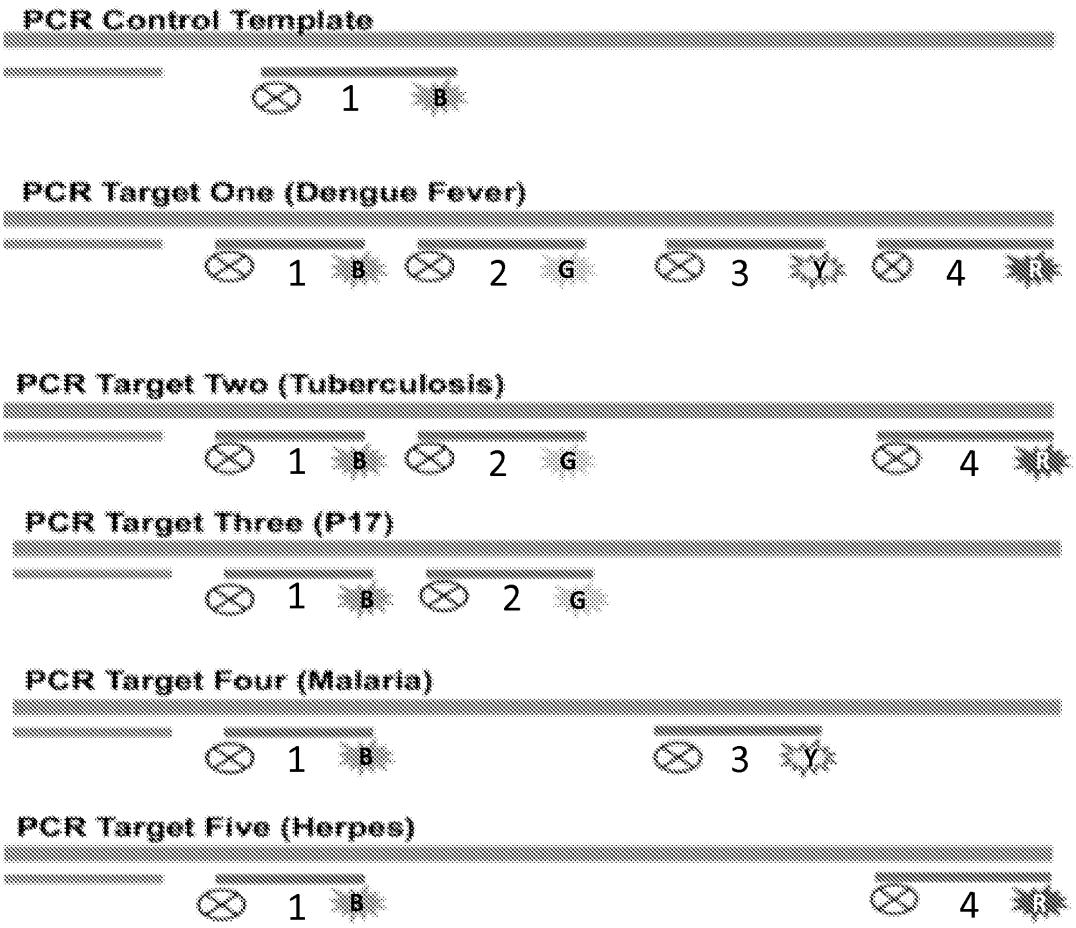
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 81

Fig. 1



B: Blue
G: Green
Y: Yellow
R: Red

Fig. 2



B: Blue
G: Green
Y: Yellow
R: Red

Fig. 3

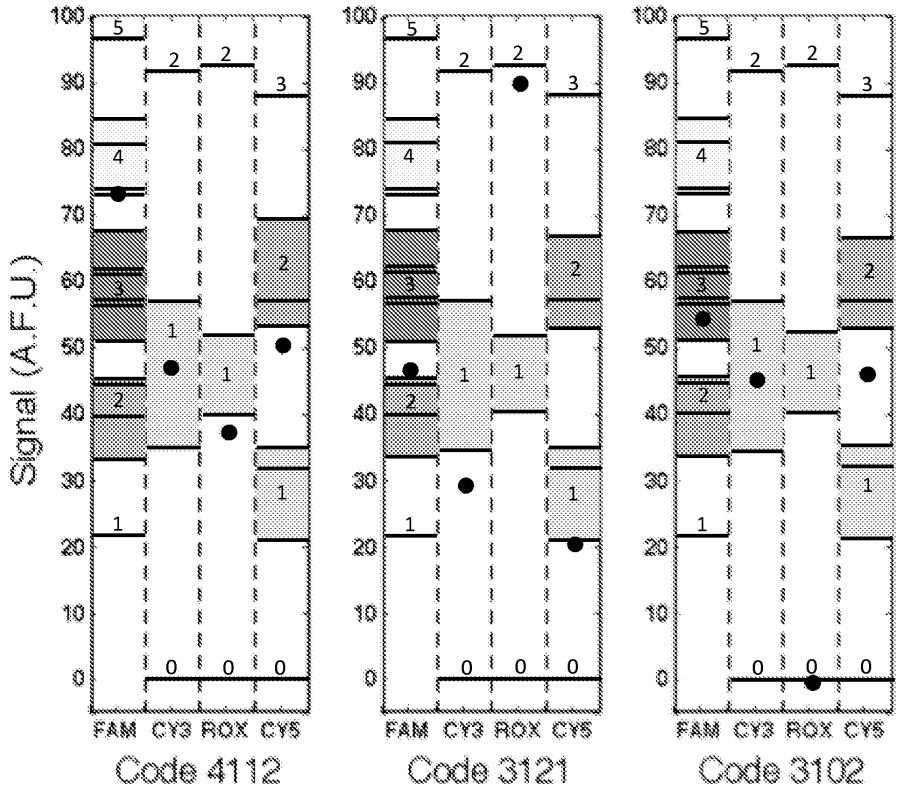


Fig. 4

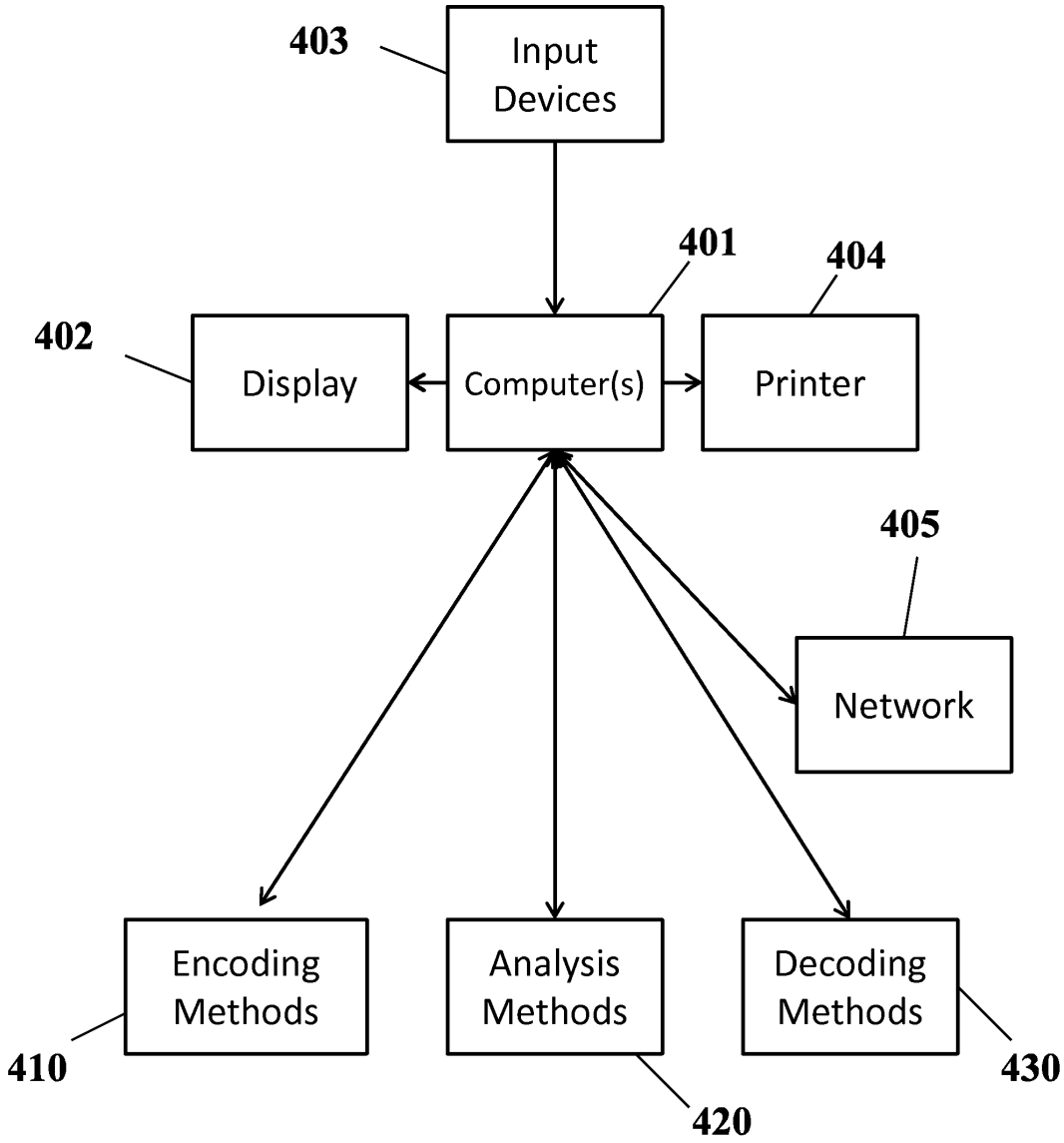

403
Input Devices
401
404
402
Display
Computer(s)
Printer
405
Network
Encoding Methods
Analysis Methods
Decoding Methods
410
420
430

Fig. 5

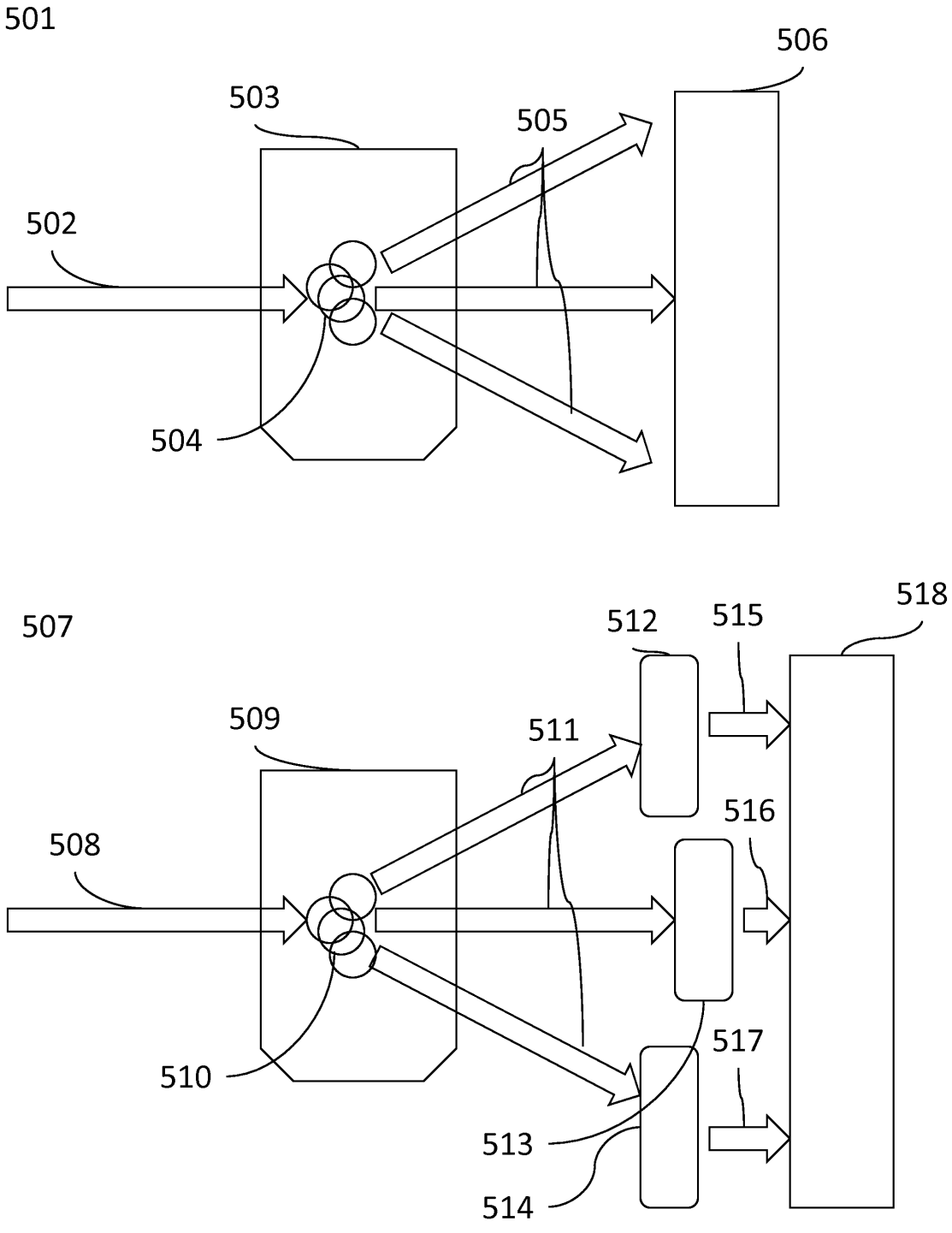

501
506
503
505
502
504
507
518
512
515
509
511
516
508
510
517
513
514

Fig. 6

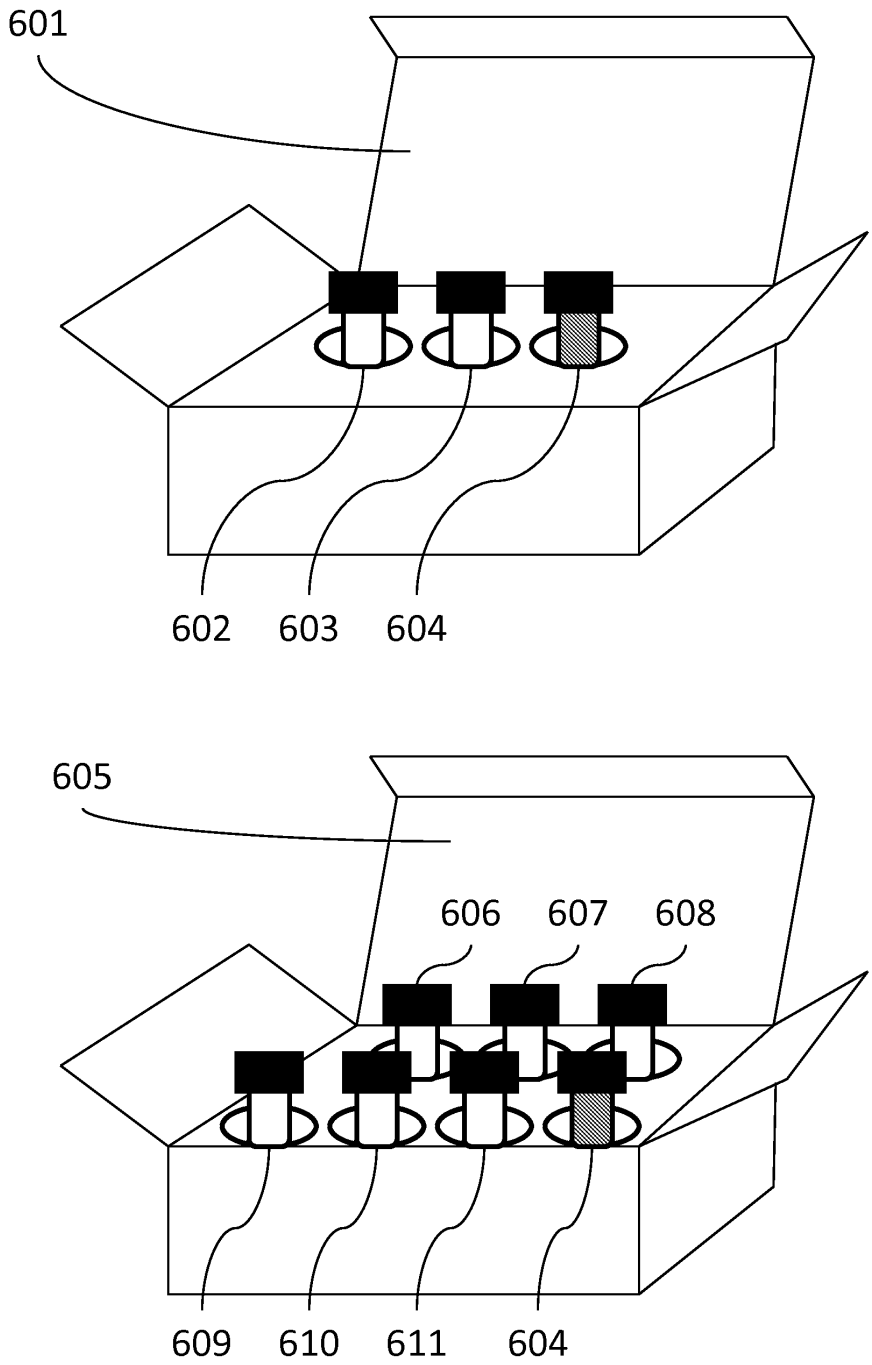


Fig. 7

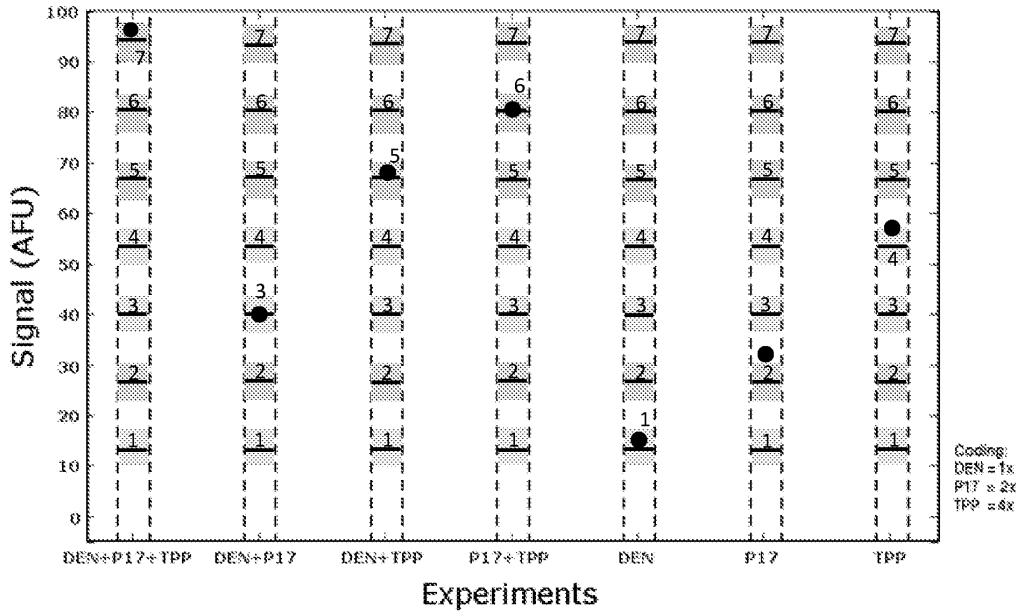
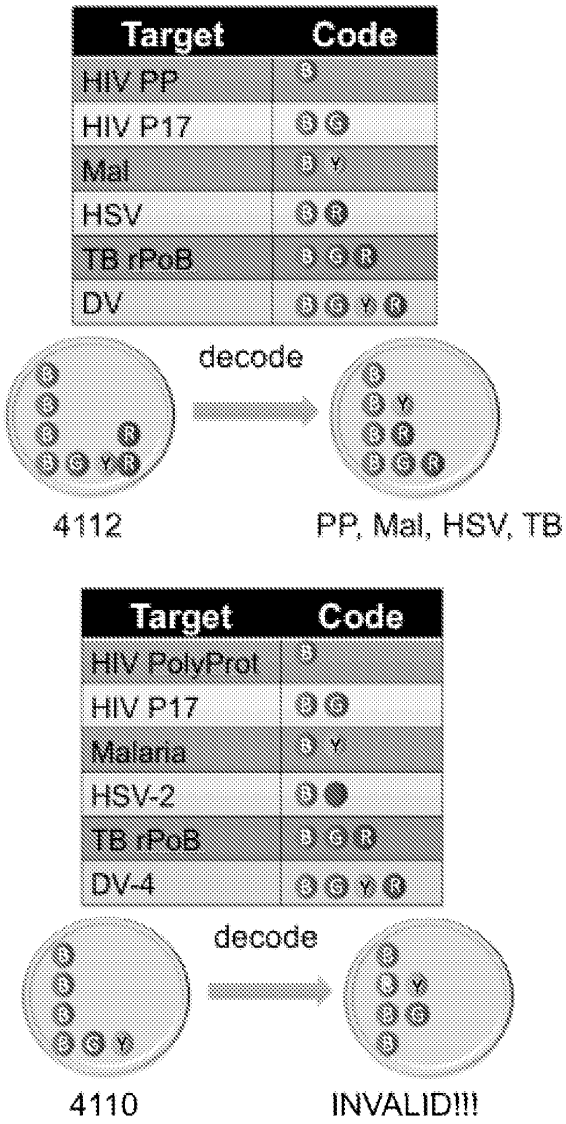

Signal (AFU)
Experiments
DEN+P17+TPP
DEN+P17
DEN+TPP
P17+TPP
DEN
P17
TPP
Coding:
DEN = 1x
P17 = 2x
TPP = 4x

Fig. 8


Target
Code
HIV PP
HIV P17
Mal
HSV
TB rPoB
DV
decode
4112
PP, Mal, HSV, TB
Target
Code
HIV PolyProt
HIV P17
Malaria
HSV-2
TB rPoB
DV-4
decode
4110
INVALID!!!

PTO/AIA/14 (12-13)
Approved for use through 01/31/2014. 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	38075-716.301
	Application Number	
Title of Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	

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
	Emil		Kartalov	

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

City	Pasadena	State/Province	CA	Country of Residence i	US

Mailing Address of Inventor:

Address 1	2133 E. Mountain Street		
Address 2			
City	Pasadena	State/Province	CA
Postal Code	91104	Country i	US

Inventor 2 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Aditya		Rajagopal	

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

City	Irvine	State/Province	CA	Country of Residence i	US

Mailing Address of Inventor:

Address 1	27 Lee		
Address 2			
City	Irvine	State/Province	CA
Postal Code	92620	Country i	US

Inventor 3 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Axel		Scherer	

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

PTO/AIA/14 (12-13)
Approved for use through 01/31/2014. 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	38075-716.301
	Application Number	
Title of Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	

City	Barnard	State/Province	VT	Country of Residence i	US

Mailing Address of Inventor:

Address 1	1661 North Road - P.O. Box 5		
Address 2			
City	Barnard	State/Province	VT
Postal Code	05031	Country i	US

Inventor 4 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix

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

City		State/Province		Country of Residence i	

Mailing Address of Inventor:

Address 1			
Address 2			
City		State/Province	
Postal Code		Country i	

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	21971		
Email Address	patentdocket@wsgr.com	Add Email	Remove Email

Application Information:

Title of the Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS		
Attorney Docket Number	38075-716.301	Small Entity Status Claimed	☒
Application Type	Nonprovisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	8	Suggested Figure for Publication (if any)	

Filing By Reference :

PTO/AIA/14 (12-13)
Approved for use through 01/31/2014. 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	38075-716.301
	Application Number	
Title of Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	

Only compete 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

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.

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	21971		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, or 365(c) or indicate National Stage entry from a PCT application. Providing this 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 blank.

Prior Application Status	Pending		Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
	Continuation of	13756760	2013-02-01
Prior Application Status	Expired		Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
13756760	Claims benefit of provisional	61703093	2012-09-19
Prior Application Status	Expired		Remove

PTO/AIA/14 (12-13)
Approved for use through 01/31/2014. 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	38075-716.301
	Application Number	
Title of Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	

Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
13756760	Claims benefit of provisional	61594480	2012-02-03

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(d). 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(h)(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)

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.

Authorization to Permit Access:

☒ Authorization to Permit Access to the Instant Application by the Participating Offices

PTO/AIA/14 (12-13)
Approved for use through 01/31/2014. 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	38075-716.301
	Application Number	
Title of Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	

If 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 World Intellectual Property Office (WIPO), and any other intellectual property offices in which a foreign application claiming priority to the instant patent application is filed access to the instant patent application. See 37 CFR 1.14(c) and (h). This box should not be checked if the applicant does not wish the EPO, JPO, KIPO, WIPO, or other intellectual property office in which a foreign application claiming priority to the instant patent application is filed to have access to the instant patent application.

In accordance with 37 CFR 1.14(h)(3), access will be provided to a copy of the instant patent application with respect to: 1) the instant patent application-as-filed; 2) any foreign application to which the instant patent application claims priority under 35 U.S.C. 119(a)-(d) if a copy of the foreign application that satisfies the certified copy requirement of 37 CFR 1.55 has been filed in the instant patent application; and 3) any U.S. application-as-filed from which benefit is sought in the instant patent application.

In accordance with 37 CFR 1.14(c), access may be provided to information concerning the date of filing this Authorization.

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	California Institute of Technology		
Mailing Address Information:			
Address 1	1200 East California Boulevard		
Address 2			
City	Pasadena	**State/Province**	CA
Country [i]	US	Postal Code	91125
Phone Number		Fax Number	

PTO/AIA/14 (12-13)
Approved for use through 01/31/2014. 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	38075-716.301
	Application Number	
Title of Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	

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 subsitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

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 form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications

Signature	/Reena Bajpai/			Date (YYYY-MM-DD)	2014-08-05
First Name	Reena	Last Name	Bajpai	Registration Number	69996

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

PTO/AIA/14 (12-13)
Approved for use through 01/31/2014. 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	38075-716.301
	Application Number	
Title of Invention	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	

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 Cooperation Treaty.

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.

DocCode - SCORE

SCORE Placeholder Sheet for IFW Content

Application Number: 14451876 Document Date: 08/05/2014

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.

- Drawings – Other than Black and White Line Drawings

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.

To access the documents in the SCORE database, refer to instructions below.

At the time of document entry (noted above):

- Examiners may access SCORE content via the eDAN interface.
- Other USPTO employees can bookmark the current SCORE URL (http://Score.uspto.gov/ScoreAccessWeb/).
- External customers may access SCORE content via the Public and Private PAIR interfaces.

Form Revision Date: September 30, 2013

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number: 14/451,876

APPLICATION AS FILED - PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	70		N/A	
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	300		N/A	
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	360		N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	74 minus 20 =	* 54	x 40 =	2160	OR		
INDEPENDENT CLAIMS (37 CFR 1.16(h))	14 minus 3 =	* 11	x 210 =	2310			
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).			0.00			
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	5200		TOTAL	

APPLICATION AS AMENDED - PART II

AMENDMENT A

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY ADDITIONAL FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE($)
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	

AMENDMENT B

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)		RATE($)	ADDITIONAL FEE($)
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
14/451,876	08/05/2014	Emil Kartalov	38075-716.301

CONFIRMATION NO. 1065

21971
WILSON, SONSINI, GOODRICH & ROSATI
650 PAGE MILL ROAD
PALO ALTO, CA 94304-1050

FORMALITIES LETTER

OC000000070191652

Date Mailed: 08/15/2014

NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

Filing Date Granted

Items Required To Avoid Abandonment:

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given **TWO MONTHS** from the date of this Notice within which to file all required items below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The statutory basic filing fee is missing.
- The application search fee must be submitted.
- The application examination fee must be submitted.
- This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 CFR 1.821(c) Applicant must provide an initial paper or compact disc copy of the "Sequence Listing", **as well as an amendment specifically directing its entry into the application** and a statement that the content of the sequence listing information recorded in computer readable form is identical to the written (on paper or compact disc) sequence listing and, where applicable, includes no new matter, as required by 37 CFR 1.821(e), 1.821(f), 1.821(g), 1.825(b), or 1.825(d). If the effective filing date is on or after September 8, 2000, see the final rulemaking notice published in the Federal Register at 65 FR 54604 (September 8, 2000) and 1238 OG 145 (September 19, 2000).
- A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 CFR 1.821(e). If the effective filing date is on or after September 8, 2000, see the final rulemaking notice published in the Federal Register at 65 FR 54604 (September 8, 2000) and 1238 OG 145 (September 19, 2000). Applicant must provide an initial computer readable form (CRF) copy of the "Sequence Listing" and a statement that the content of the sequence listing information recorded in computer readable form is identical to the written (on paper or compact disc) sequence listing and, where applicable, includes no new matter, as required by 37 CFR 1.821(e), 1.821(f), 1.821(g), 1.825(b), or 1.825(d). If applicant desires the sequence listing in the instant application to be identical with that of another application on file in the U.S. Patent and Trademark Office, such request in accordance with 37 CFR 1.821(e) may be submitted in lieu of a new CRF.

Applicant is cautioned that correction of the above items may cause the specification and drawings page count to exceed 100 pages. If the specification and drawings exceed 100 pages, applicant will need to submit the required application size fee.

For questions regarding compliance to these requirements, please contact:

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 100

- **For Rules interpretation, call (571) 270-5300**
- **For CRF submission help, call (571) 272-2510**
- **For PatentIn software program support technical assistance, call (866)217-9197.**
- **Send e-mail correspondence for PatentIn software program help to ebc@uspto.gov**
- **Patent software is available at http://www.uspto.gov/patents/resources/tools/checker/patentinrel.jsp .**

- Additional claim fees of $ **4470** as a small entity, including any required multiple dependent claim fee, are required. Applicant must submit the additional claim fees or cancel the additional claims for which fees are due.
- Surcharge as set forth in 37 CFR 1.16(f) must be submitted.
  The surcharge is due for any one of:
  - late submission of the basic filing fee, search fee, or examination fee,
  - late submission of inventor's oath or declaration,
  - filing an application that does not contain at least one claim on filing, or
  - submission of an application filed by reference to a previously filed application.

SUMMARY OF FEES DUE:

The fee(s) required within **TWO MONTHS** from the date of this Notice to avoid abandonment is/are itemized below. Small entity discount is in effect. If applicant is qualified for micro entity status, an acceptable Certification of Micro Entity Status must be submitted to establish micro entity status. (See 37 CFR 1.29 and forms PTO/SB/15A and 15B.)

- $ **70** basic filing fee.
- $ **70** surcharge.
- $ **300** search fee.
- $ **360** examination fee.
- $ **2310** for **11** independent claims over 3.
- $ **2160** for **54** total claims over 20.
- $(**0**) previous unapplied payment amount.
- $ **5270** TOTAL FEE BALANCE DUE.

Items Required To Avoid Processing Delays:

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).

- A properly executed inventor's oath or declaration has not been received for the following inventor(s):
  Emil Kartalov
  Aditya Rajagopal
  Axel Scherer

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 101

Replies must be received in the USPTO within the set time period or must include a proper Certificate of Mailing or Transmission under 37 CFR 1.8 with a mailing or transmission date within the set time period. For more information and a suggested format, see Form PTO/SB/92 and MPEP 512.

Replies should be mailed to:

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

Registered users of EFS-Web may alternatively submit their reply to this notice via EFS-Web, including a copy of this Notice and selecting the document description "Applicant response to Pre-Exam Formalities Notice". https://sportal.uspto.gov/authenticate/AuthenticateUserLocalEPF.html

For more information about EFS-Web please call the USPTO Electronic Business Center at **1-866-217-9197** or visit our website at http://www.uspto.gov/ebc.

If you are not using EFS-Web to submit your reply, you must include a copy of this notice.

/tnguyen/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

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
14/451,876	08/05/2014	1677	0.00	38075-716.301	74	14

CONFIRMATION NO. 1065

21971
WILSON, SONSINI, GOODRICH & ROSATI
650 PAGE MILL ROAD
PALO ALTO, CA 94304-1050

FILING RECEIPT

OC000000070191651

Date Mailed: 08/15/2014

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)

Emil Kartalov, Pasadena, CA;
Aditya Rajagopal, Irvine, CA;
Axel Scherer, Barnard, VT;

Applicant(s)

California Institute of Technology, Pasadena, CA

Power of Attorney: None

Domestic Priority data as claimed by applicant

This application is a CON of 13/756,760 02/01/2013
which claims benefit of 61/703,093 09/19/2012
and claims benefit of 61/594,480 02/03/2012

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 - A proper **Authorization to Permit Access to Application by Participating Offices** (PTO/SB/39 or its equivalent) has been received by the USPTO.

If Required, Foreign Filing License Granted: 08/14/2014
The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 14/451,876**
Projected Publication Date: To Be Determined - pending completion of Missing Parts

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 103

Non-Publication Request: No

Early Publication Request: No
** **SMALL ENTITY** **
Title

SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

Preliminary Class

436

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).

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 104

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.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	KARTALOV, EMIL P. *et al.*
Title:	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS
Appl No.:	14/451,876
Filing Date:	August 5, 2014
Examiner:	To be Assigned
Art Unit:	1677
Confirmation No.:	1065

ELECTRONICALLY FILED November 14, 2014

RESPONSE TO NOTICE TO FILE MISSING PARTS AND PRELIMINARY AMENDMENT

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

Sir or Madam:

This Amendment is concurrently filed with a reply to the Notice to File Missing Parts mailed August 15, 2014. The two month deadline for filing a response to the Missing Parts expired on October 15, 2014. The response is filed before or on November 15, 2014 along with a fee for one month extension of time. Accordingly, the Applicants believe that this Amendment is timely filed.

Prior to an examination on the merits of the above-identified application, please amend the application as follows:

Amendments to the Specification begin on page **2** of this paper.

Amendments to the Claims begin on page **7** of this paper.

Remarks begin on page **12** of this paper.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 106

AMENDMENTS IN THE SPECIFICATION

Please amend paragraph [0001] as follow:

[0001] This application is a continuation of U.S. Patent Application No. 13/756,760, filed February 1, 2013, which claims the benefit of U.S. provisional applications 61/594,480, filed February 3, 2012 and 61/703,093, filed September 19, 2012. The content of the three applications are ~~, each of which is~~ incorporated herein by reference in ~~its~~ entirety.

Please insert the following paragraphs on page 1, after paragraph [0001]:

SEQUENCE LISTING

[0001.1] 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 August 20, 2014, is named 38075-716-301-Seqlist.txt and is 9,260 bytes in size.

STATEMENT AS TO FEDERALLY SPONSERED RESEARCH

[0001.5] This invention was made with the support of the United States government under Grant No. 1144469 awarded by NSF. The government has certain rights in the invention.

Please delete paragraph [00174] and replace it with the following paragraph:

[00174] Oligonucleotides were synthesized by INTEGRATED DNA TECHNOLOGIES (IDT). Probes and primer pairs were chosen for each analyte using OLIGOANALYZER, from IDT. Sense probes containing a fluorophore at the 5' end and a quencher at the 3' end were synthesized for all analytes. All oligonucleotides were lyophilized and reconstituted in TE buffer before use. Tables 9 - 14 show sequence information, target sequence, primers, and probes for each of the six analytes.

Table 9. HIV-1 polyprotease sequence information.

Sequence Information	HIV-1 Poly protease 198mer synthesized from bases 2253 – 2550
Source	HIV-1 Reference Sequence, Los Alamos National Laboratory
Target sequence (3' to 5')	GGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGA GTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTT

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 107

	ATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAA AGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGG **(SEQ ID NO: 1)**
Forward Primer (3' to 5')	GGAAGCTCTATTAGATACAGGAGCAG **(SEQ ID NO: 2)**
Reverse Primer (3' to 5')	CCAATTATGTTGACAGGTGTAGGTCC **(SEQ ID NO: 3)**
Probe 1 (3' to 5')	/**56-FAM**/TGAGTTTGCCAGGAAGATGGAAACCA/3BHQ_1/ **(SEQ ID NO: 4)**

Table 10. HIV p17 sequence information.

Sequence Name	HIV-1 P17 199mer synthesized from bases 790 – 1186
Source	HIV-1 Reference Sequence, Los Alamos National Laboratory
Target sequence (3' to 5')	CAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATA TAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAG ACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAG AAAAAAGCACAGCAAGCAGCAGCTGACACAGGACACAGCAATCAGGTCA **(SEQ ID NO: 5)**
Forward Primer (3' to 5')	CAGCTACAACCATCCCTTCAGACA **(SEQ ID NO: 6)**
Reverse Primer (3' to 5')	TGACCTGATTGCTGTGTCCTGTGT **(SEQ ID NO: 7)**
Probe 1 (3' to 5')	/**56-FAM**/AGCAACCCTCTATTGTGTGCATCAAAGG/3BHQ_1 **(SEQ ID NO: 8)**
Probe 2 (3' to 5')	/**5Cy3**/AAAGCACAGCAAGCAGCAGCTGA/3BHQ_2/ **(SEQ ID NO: 9)**

Table 11. *Plasmodium* ChR7 sequence information.

Sequence Name	*Plasmodium* ChR7 199mer synthesized from bases 1139138 - 1141223
Source	UCSC Plasmodium falciparum Genome Browser Gateway
Target sequence (3' to 5')	GCCTAACATGGCTATGACGGGTAACGGGGAATTAGAGTTCGATTCCGGAG AGGGAGCCTGAGAAATAGCTACCACATCTAAGGAAGGCAGCAGGCGCGTA

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 108

	AATTACCCAATTCTAAAGAAGAGAGGTAGTGACAAGAAATAACAATGCAA GGCCAATTTAAAACCTTCCCAGAGTAACAATTGGAGGGCAAGTCTGGTG **(SEQ ID NO: 10)**
Forward Primer (3' to 5')	GCCTAACATGGCTATGACGGGTAA **(SEQ ID NO: 11)**
Reverse Primer (3' to 5')	CACCAGACTTGCCCTCCAATTGTT **(SEQ ID NO: 12)**
Probe 1 (3' to 5')	/**56-FAM**/ATTCCGGAGAGGGAGCCTGAGAAATA/3BHQ_1/ **(SEQ ID NO: 13)**
Probe 2 (3' to 5')	/**56-ROXN**/AAGGAAGGCAGCAGGCGCGTAAATTA/3BHQ_2/ **(SEQ ID NO: 14)**

Table 12. Herpes simplex virus 2 sequence information.

Sequence Name	HSV-2 193mer synthesized from HSV-2 genome
Source	EMBL Bank AJ303204
Target sequence (3' to 5')	TCAGCCCATCCTCCTTCGGCAGTATGGAGGGTGTCGCGGCGGCGAGCCGC CGTCCCCAAAGACGTGCGGGTCGTACACGTACACGTACCAGGGCGGCGGG CCTCCGACCCGGTACGCTCTCGTAAATGCTTCCCTGCTGGTGCCGATCTG GGACCGCGCCGCGGAGACATTCGAGTACCAGATCGAACTCGG **(SEQ ID NO: 15)**
Forward Primer (3' to 5')	TCAGCCCATCCTCCTTCGGCAGTAT **(SEQ ID NO: 16)**
Reverse Primer (3' to 5')	CCGAGTTCGATCTGGTACTCGAATGT **(SEQ ID NO: 17)**
Probe 1 (3' to 5')	/**56-FAM**/AAAGACGTGCGGGTCGTACACGTACA/3BHQ_1/ **(SEQ ID NO: 18)**
Probe 2 (3' to 5')	/**5Cy5**/TAAATGCTTCCCTGCTGGTGCCGAT/3IAbRQSp/ **(SEQ ID NO: 19)**

Table 13. *Mycobacterium tuberculosis* rpoB sequence information

Sequence Name	Mycobacterium Tuberculosis rpoB 200mer synthesized from Mycobacterium tuberculosis rpoB genome
Source	EMBL Bank GQ395623

Target sequence (3' to 5')	GAGTGCAAAGACAAGGACATGACGTACGCGGCCCCGCTGTTCGTCACGGC CGAGTTCATCAACAACAACACCGGCGAGATCAAGAGCCAGACGGTGTTCA TGGGTGACTTCCCGATGATGACCGAGAAGGGCACCTTCATCATCAACGGC ACCGAGCGCGTCGTGGTCAGCCAGCTGGTCCGCTCGCCCGGTGTGTACTT **(SEQ ID NO: 20)**
Forward Primer (3' to 5')	GAGTGCAAAGACAAGGACATGACG **(SEQ ID NO: 21)**
Reverse Primer (3' to 5')	AAGTACACACCGGGCGAGC **(SEQ ID NO: 22)**
Probe 1 (3' to 5')	/**56-FAM**/CGCTGTTCGTCACGGCCGAGTTCAT/3BHQ_1/ **(SEQ ID NO: 23)**
Probe 2 (3' to 5')	/**5Cy3**/AGATCAAGAGCCAGACGGTGTTCATG/3BHQ_2/ **(SEQ ID NO: 24)**
Probe 3 (3' to 5')	/**5Cy5**/AAGGGCACCTTCATCATCAACGGCA/3IAbRQSp/ **(SEQ ID NO: 25)**

Table 14. Dengue virus type 3 sequence information.

Sequence Name	Dengue Virus Type 3 200mer synthesized from Dengue Virus Type 3 genome
Source	GenBank M93130
Target sequence (3' to 5')	ATGCCAACTGTGATTGAGCACTTAGAAAGACTACAAAGGAAACATGGAGG AATGCTTGTGAGAAATCCACTCTCACGAAACTCCACGCACGAAATGTATT GGATATCCAATGGTACAGGCAACATCGTCTCATTCACAATGACACACAGG AGACCCACCATAGAGAAAGATGTGGATTTAGGAGCAGGAACCCGACATGT **(SEQ ID NO: 26)**
Forward Primer (3' to 5')	ATGCCAACTGTGATTGAGCACT **(SEQ ID NO: 27)**
Reverse Primer (3' to 5')	ACATGTCGGGTTCCTGCTCCTAAA **(SEQ ID NO: 28)**
Probe 1 (3' to 5')	/**56-FAM**/ACAAAGGAAACATGGAGGAATGCTTGTGA/3BHQ_1/ **(SEQ ID NO: 29)**
Probe 2 (3' to 5')	/**5Cy3**/ACTCTCACGAAACTCCACGCACGAAA/3BHQ_2/ **(SEQ ID NO: 30)**
Probe 3 (3' to	/**56-ROXN**/ACAATGACACACAGGAGACCCACCAT/3BHQ_2/ **(SEQ ID NO:**

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 110

5')	**31)**
Probe 4 (3' to 5')	/**5Cy5**/GGATATCCAATGGTACAGGCAACATCGT/3IAbRQSp/ **(SEQ ID NO: 32)**

3. Polymerase Chain Reactions

AMENDMENTS TO THE CLAIMS

1. **(Currently Amended)** An assay that is capable of unambiguously detecting the presence or absence of each of at least seven (7) analytes, in any combination of presence or absence, in a single sample volume without immobilization of said analytes, physical separation of said analytes, melting curve analysis, or mass spectrometry~~, or melting curve analysis~~.
2. **(Original)** The assay of claim 1, wherein each of said analytes is encoded as a value of one component of a signal.
3. **(Canceled)**
4. **(Currently Amended)** The assay of claim ~~1~~ 2, wherein each of said analytes is encoded as at least one first value in a first component of said signal and at least one second value in a second component of said signal.
5. **(Currently Amended)** The assay of claim ~~4~~ 2, wherein each of said analytes is encoded as a first value in a first component of said signal at each of a plurality of second values in a second component of said signal.

6.-8. **(Canceled)**

9. **(Currently Amended)** The ~~method~~ assay of claim 4, wherein said first value is an intensity or range of intensities and said second value is a wavelength or range of wavelengths.

10.-12. **(Canceled)**

13. **(Currently Amended)** The ~~method~~ assay of claim ~~1~~2, wherein said signal is an electromagnetic signal.

14.-16. **(Canceled)**

17. **(Currently Amended)** The ~~method~~ assay of claim 1, wherein said detecting is performed with reagents comprising hybridization probes.

18. **(Currently Amended)** The ~~method~~ assay of claim 1~~-17~~, wherein~~ the number of said hybridization probes is greater than the number of said analytes~~ said detecting is performed with a plurality of hybridization probes, comprising an identical fluorophore, wherein said plurality of hybridization probes are specific for different analytes.

19. **(Currently Amended)** The ~~method~~ assay of claim 1 ~~17~~, wherein ~~said hybridization probes comprise one or more hybridization probes specific for different analytes and comprising an~~

~~identical fluorophore or combination of fluorophores~~ said detecting is performed with a plurality of hybridization probes comprising different fluorophores, wherein said plurality of hybridization probes are specific for different analytes.

20.-21. **(Canceled)**

22. **(Currently Amended)**The ~~method~~ assay of claim 2 ~~21~~, wherein said signal is generated during a polymerase chain reaction.

23.-26. **(Canceled)**

27. **(Currently Amended)** The ~~method~~ assay of claim 4 ~~1~~, further comprising constructing a diagram ~~chromatogram~~.

28. **(Currently Amended)** The ~~method~~ assay of claim 27, wherein said diagram ~~chromatogram~~ is constructed by plotting ~~all~~ possible combinations of said first values and said second values for positive control samples.

29. **(Currently Amended)** The ~~method~~ assay of claim 1, wherein at least one of said analytes comprises a polynucleotide.

30.-33. **(Canceled)**

34. **(Currently Amended)** The assay ~~method~~ of claim 1, further comprising transmitting information concerning the presence or absence of said analyte through a computer network.

35. **(Currently Amended)**The ~~method~~ assay of claim 1, further comprising providing information concerning the presence or absence of said analyte to a ~~physician~~ health care professional, wherein the health care professional makes a clinical decision based on said information.

36. **(Canceled)**

37. **(Currently Amended)** The ~~method~~ assay of claim 1, wherein at least one step of said assay ~~method~~ is performed using ~~instructions on~~ a computer ~~readable medium~~.

38. **(Currently Amended)** The ~~method~~ assay of claim 37, wherein said computer is ~~instructions are~~ located on a remote server.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 113

39. **(Currently Amended)** The ~~method~~ assay of claim 37, wherein said computer is ~~instructions are~~ located on a thermal cycler.

40.-41. **(Canceled)**

42. **(Currently Amended)** The ~~method~~ assay of claim ~~1~~2, wherein said coding ~~scheme~~ is innately non-degenerate.

43. **(Currently Amended)** The ~~method~~ assay of claim 42, wherein said coding ~~scheme~~ is designed to be non-degenerate by enumerating legitimate results that can be obtained from said encoding, identifying legitimate results that are degenerate, and eliminating at least one potential analyte code from said encoding to eliminate degeneracy.

44. **(Canceled)**

45. **(Currently Amended)** The ~~method~~ assay of claim 1, wherein said assay is an end-point assay.

46. **(Canceled)**

47. **(Currently Amended)** The ~~method~~ assay of claim 1, wherein said assay is a liquid phase assay.

48. **(Canceled)**

49. **(Original)** A method of detecting the presence or absence of each analyte of a plurality of analytes, comprising:

    a. encoding each of said analytes as a first value of a signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said first value, wherein said encoding is performed in a manner that eliminates degeneracy;
    b. providing a sample comprising, or potentially comprising, at least one of said analytes;
    c. contacting said sample with analyte-specific reagents that generate said first value, as represented in said coding scheme, when each of said analytes is present;
    d. cumulatively measuring said first values of said signal within said sample, thereby providing a cumulative measurement; and
    e. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme.

50. **(Canceled)**

51. **(Original)** The method of claim 49, wherein said first value is an intensity or range of intensities.

52. **(Canceled)**

53. **(Currently Amended)** The method of claim 49, wherein said first value is incremented in said coding scheme ~~by~~ to an amount greater than the sum ~~equal to the cumulative maximum~~ of said preceding first values ~~plus one.~~

54. **(Canceled)**

55. **(Original)** A method of detecting the presence or absence of each analyte of a plurality of analytes, comprising:

    f. encoding each of said analytes as at least one first value and at least one second value, wherein said first value is a value from a first component of a signal and said second value is a value from a second component of said signal, thereby generating a coding scheme, wherein each of said analytes is represented in said coding scheme by said at least one first value and said at least one second value, wherein said encoding is performed in a manner that reduces or eliminates degeneracy;

    g. providing a sample comprising, or potentially comprising, at least one of said analytes;

    h. contacting said sample with analyte-specific reagents that generate said at least one first value and at least one said second value, as represented in said coding scheme, when each of said analytes is present;

    i. cumulatively measuring said first values and said second values within said sample, thereby providing a cumulative measurement; and

    j. determining whether each of said analytes is present or absent based on said cumulative measurement and said coding scheme,

    wherein, when degeneracy is eliminated, said method is capable of unambiguously detecting the presence or absence of each of at least six analytes in a single volume, in any combination of presence or absence, when each of said reagents generates only one second value.

56. **(Original)** The method of claim 55, wherein each of said analytes is encoded as a first value

in a first component of said signal at each of a plurality of second values in a second component of said signal.

57.-74. **(Canceled)**

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 116

REMARKS

Claims 3, 6-8, 10-12, 14-16, 20, 21, 23-26, 30-33, 36, 40, 41, 44, 46, 48, 50, 52, 54 and 57-74 are canceled. Claims 1, 4, 5, 9, 13, 17-19, 22, 27-29, 34, 35, 37-39, 42, 43, 45, 47 and 53 are amended. No new claims are added. Upon entry of this amendment, claims 1, 2, 4, 5, 9, 13, 17-19, 22, 27-29, 34, 35, 37-39, 42, 43, 45, 47, 49, 51, 53, 55 and 56 will be pending. Amended claims are supported by the specification and claims as originally filed. Illustrative support for amendments can be found in the claim support chart on page 13 of this letter. Accordingly, no new matter is introduced by this Preliminary Amendment. Examination and favorable consideration is respectfully requested in view of the present amendment.

Sequence Listing

The foregoing amendments are made to insert the required SEQ ID NO identifiers associated with each listed sequence, as well as to incorporate by reference the Sequence Listing filed concurrently herewith. No new matter has been added.

CONCLUSION

Applicants respectfully request that the Examiner advance the application to issuance. The Commissioner is authorized to charge any additional fees which may be required, including petition fees and extension of time fees, to Deposit Account No. 23-2415 (Docket No. 38075-716.301).

Respectfully submitted,
WILSON SONSINI GOODRICH & ROSATI

Date: November 14, 2014

By: /Reena Bajpai/
Reena Bajpai, Ph.D., Patent Agent
Registration No.: 69,996

650 Page Mill Road
Palo Alto, California 94304
Telephone No.: (650) 493-9300
Facsimile No.: (650) 493-6811
Customer No. 21971

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 117

Amended Claim	Support found
1	Original claim 1
4 and 5	Original claims 4 and 5
9	Original claims 9 and 11
13	Original claim 13
17-19	Original claims 17-19
22	Original Claim 22
27-28	Original claims 27 and 28 and paragraph [0179]
29	Original claim 29
34	Original claim 34
35	Original claims 35 and 36
37-39	Original claims 37-39 and paragraph [0048] and [0021]
42	Original claim 42
43	Original claims 43 and 44
45	Original claim 45
47	Original claim 47
53	Original claims 53 and 54

<110> KARTALOV, EMIL P.
RAJAGOPAL, ADITYA
SCHERER, AXEL

<120> SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

<130> 38075-716.301

<140> 14/451,876
<141> 2015-08-05

<150> 13/756,760
<151> 2013-02-01

<150> 61/703,093
<151> 2012-09-19

<150> 61/594,480
<151> 2012-02-03

<160> 32

<170> PatentIn version 3.5

<210> 1
<211> 198
<212> DNA
<213> Human immunodeficiency virus 1

<400> 1
ggaagctcta ttagatacag gagcagatga tacagtatta gaagaaatga gtttgccagg 60

aagatggaaa ccaaaaatga tagggggaat tggaggtttt atcaaagtaa gacagtatga 120

tcagatactc atagaaatct gtggacataa agctataggt acagtattag taggacctac 180

acctgtcaac ataattgg 198

<210> 2
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 2
ggaagctcta ttagatacag gagcag 26

<210> 3
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 119

<400> 3
ccaattatgt tgacaggtgt aggtcc 26

<210> 4
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 4
tgagtttgcc aggaagatgg aaacca 26

<210> 5
<211> 199
<212> DNA
<213> Human immunodeficiency virus 1

<400> 5
cagctacaac catcccttca gacaggatca gaagaactta gatcattata taatacagta 60

gcaaccctct attgtgtgca tcaaaggata gagataaaag acaccaagga agctttagac 120

aagatagagg aagagcaaaa caaaagtaag aaaaaagcac agcaagcagc agctgacaca 180

ggacacagca atcaggtca 199

<210> 6
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 6
cagctacaac catcccttca gaca 24

<210> 7
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 7
tgacctgatt gctgtgtcct gtgt 24

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 120

<210> 8
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 8
agcaaccctc tattgtgtgc atcaaagg 28

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy3

<220>
<223> 3'BHQ_2

<400> 9
aaagcacagc aagcagcagc tga 23

<210> 10
<211> 199
<212> DNA
<213> Plasmodium falciparum

<400> 10
gcctaacatg gctatgacgg gtaacgggga attagagttc gattccggag agggagcctg 60

agaaatagct accacatcta aggaaggcag caggcgcgta aattacccaa ttctaaagaa 120

gagaggtagt gacaagaaat aacaatgcaa ggccaattta aaaccttccc agagtaacaa 180

ttggagggca agtctggtg 199

<210> 11
<211> 24
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
primer

<400> 11
gcctaacatg gctatgacgg gtaa 24

<210> 12
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 12
caccagactt gccctccaat tgtt 24

<210> 13
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 13
attccggaga gggagcctga gaaata 26

<210> 14
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-ROXN

<220>
<223> 3'BHQ_2

<400> 14
aaggaaggca gcaggcgcgt aaatta 26

<210> 15
<211> 192
<212> DNA

<213> Herpes simplex virus 2

<400> 15
tcagcccatc ctccttcggc agtatggagg gtgtcgcggc ggcgagccgc cgtccccaaa 60

gacgtgcggg tcgtacacgt acacgtacca gggcggcggg cctccgaccc ggtacgctct 120

cgtaaatgct tccctgctgg tgccgatctg ggaccgcgcc gcggagacat tcgagtacca 180

gatcgaactc gg 192

<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 16
tcagcccatc ctccttcggc agtat 25

<210> 17
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 17
ccgagttcga tctggtactc gaatgt 26

<210> 18
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 18
aaagacgtgc gggtcgtaca cgtaca 26

<210> 19
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy5

<220>
<223> 3'3IAbRQSp

<400> 19
taaatgcttc cctgctggtg ccgat 25

<210> 20
<211> 200
<212> DNA
<213> Mycobacterium tuberculosis

<400> 20
gagtgcaaag acaaggacat gacgtacgcg gccccgctgt tcgtcacggc cgagttcatc 60

aacaacaaca ccggcgagat caagagccag acggtgttca tgggtgactt cccgatgatg 120

accgagaagg gcaccttcat catcaacggc accgagcgcg tcgtggtcag ccagctggtc 180

cgctcgcccg gtgtgtactt 200

<210> 21
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 21
gagtgcaaag acaaggacat gacg 24

<210> 22
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 22
aagtacacac cgggcgagc 19

<210> 23
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 23
cgctgttcgt cacggccgag ttcat 25

<210> 24
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy3

<220>
<223> 3'BHQ_2

<400> 24
agatcaagag ccagacggtg ttcatg 26

<210> 25
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy5

<220>
<223> 3'IAbRQSp

<400> 25
aagggcacct tcatcatcaa cggca 25

<210> 26
<211> 200
<212> DNA
<213> Dengue virus 3

<400> 26
atgccaactg tgattgagca cttagaaaga ctacaaagga aacatggagg aatgcttgtg 60

agaaatccac tctcacgaaa ctccacgcac gaaatgtatt ggatatccaa tggtacaggc 120

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 125

aacatcgtct cattcacaat gacacacagg agacccacca tagagaaaga tgtggattta 180

ggagcaggaa cccgacatgt 200

<210> 27
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 27
atgccaactg tgattgagca ct 22

<210> 28
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 28
acatgtcggg ttcctgctcc taaa 24

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

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 29
acaaaggaaa catggaggaa tgcttgtga 29

<210> 30
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'Cy3

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 126

<220>
<223> 3'BHQ_2

<400> 30
actctcacga aactccacgc acgaaa 26

<210> 31
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-ROXN

<220>
<223> 3'BHQ_2

<400> 31
acaatgacac acaggagacc caccat 26

<210> 32
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy5

<220>
<223> 3'IAbRQSp

<400> 32
ggatatccaa tggtacaggc aacatcgt 28

Electronically submitted: November 14, 2014

PATENT
Attorney Docket No.: 38075-716.301

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor: Emil Kartalov et al.	Group Art Unit: 1677
Serial Number: 14/451,876	Examiner: To be Assigned
Filing Date: August 5, 2014	**CONFIRMATION NO.: 1065**
Title: SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS	**CUSTOMER NO.: 21971**

STATEMENT REGARDING SEQUENCE SUBMISSION

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

Madame:

In connection with the sequence listing submitted concurrently herewith, the undersigned hereby states that:

1. The submission, filed herewith in accordance with 37 C.F.R. § 1.821 (g), does not include new matter; and

2. The content of the electronically filed Sequence Listing is submitted in accordance with 37 C.F.R. § 1.821(e).

The Commissioner is hereby authorized to charge any required fees due in connection with this submission, including petition and extension of time fees, and to credit any overpayment, to Deposit Account No. 23-2415 (Docket No.: 38075-716.301).

Respectfully submitted,

Date: November 14, 2014

By: /Reena Bajpai/
Reena Bajpai, Ph.D., Patent Agent
Reg. No.: 69,996

WILSON SONSINI GOODRICH & ROSATI
650 Page Mill Road
Palo Alto, CA 94304-1050
Telephone: (650) 493-9300
Customer No. 021971

6585411_1.doc

PTO/AIA/01 (06-12)
Approved for use through 01/31/2014. 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	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

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 13/756,760

filed on February 1, 2013.

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: Emil Kartalov Date (Optional): _____

Signature: Emil P. Kartalov

Note: An application data sheet (PTO/AIA/14 or equivalent), including naming the entire inventive entity, must accompany this form. Use an additional PTO/SB/AIA01 form for each additional inventor.

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. 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 1 minute 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.

PTO/AIA/01 (06-12)
Approved for use through 01/31/2014. 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	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

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 13/756,760 filed on February 1, 2013.

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: Aditya Rajagopal Date (Optional): 02/28/2013

Signature: _____

Note: An application data sheet (PTO/AIA/14 or equivalent), including naming the entire inventive entity, must accompany this form. Use an additional PTO/SB/AIA01 form for each additional inventor.

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. 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 1 minute 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.

PTO/AIA/01 (06-12)
Approved for use through 01/31/2014. 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	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

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 13/756,760 filed on February 1, 2013.

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: Axel Scherer Date (Optional): 2/28/13

Signature: Axel Scherer

Note: An application data sheet (PTO/AIA/14 or equivalent), including naming the entire inventive entity, must accompany this form. Use an additional PTO/SB/AIA01 form for each additional inventor.

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. 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 1 minute 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.

Electronic Patent Application Fee Transmittal

Application Number:	14451876
Filing Date:	05-Aug-2014
Title of Invention:	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS
First Named Inventor/Applicant Name:	Emil Kartalov
Filer:	Reena Bajpai/Mai Pham
Attorney Docket Number:	38075-716.301

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Utility filing Fee (Electronic filing)	4011	1	70	70
Utility Search Fee	2111	1	300	300
Utility Examination Fee	2311	1	360	360
Pages:				
Claims:				
Claims in excess of 20	2202	7	40	280
Miscellaneous-Filing:				
Late Filing Fee for Oath or Declaration	2051	1	70	70

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 1 month with $0 paid	2251	1	100	100
Miscellaneous:				
		Total in USD ($)		**1180**

Electronic Acknowledgement Receipt	
EFS ID:	20707594
Application Number:	14451876
International Application Number:	
Confirmation Number:	1065
Title of Invention:	SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS
First Named Inventor/Applicant Name:	Emil Kartalov
Customer Number:	21971
Filer:	Reena Bajpai/Mai Pham
Filer Authorized By:	Reena Bajpai
Attorney Docket Number:	38075-716.301
Receipt Date:	14-NOV-2014
Filing Date:	05-AUG-2014
Time Stamp:	19:41:07
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$1180
RAM confirmation Number	5785
Deposit Account	232415
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)

1	Sequence Listing (Text File)	38075-716-301-Seqlist.txt	9298 76cb4ddc9e51d19ffd6cbdbe5c0a83944341b7ef	no	0
Warnings:					
Information:					
2	Applicant Response to Pre-Exam Formalities Notice	38075716301_MP_TRANSMITTAL.pdf	119951 9bf1257a89c9ee39be2f2c7f9692466fbea81b4a	no	1
Warnings:					
Information:					
3		38075716301_PRELIMINARY_AMENDMENT.pdf	183384 618a9b8860a393c1db3f9a01c8c73144028c618d	yes	13
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Preliminary Amendment		1	1	
	Specification		2	6	
	Claims		7	11	
	Applicant Arguments/Remarks Made in an Amendment		12	13	
Warnings:					
Information:					
4	Sequence Listing	38075716301_SEQ_LISTING.pdf	24108 3701a64dafa80bef45da7f5fc00126173d7c0a7b	no	9
Warnings:					
Information:					
5	CRF Statement Paper and CRF are the same	38075716301_CFR_STATEMENT.pdf	89686 777fbb6529fe42074204010e5effb4e0723e80d7	no	1
Warnings:					
Information:					
6	Oath or Declaration filed	38075716301_DECLARATIONS.pdf	150647 4ad620a63f19ca74a3d8f7b26c3ba6e9b1010f76	no	3
Warnings:					
Information:					
7	Fee Worksheet (SB06)	fee-info.pdf	40415	no	2
Warnings:					

Information:	
Total Files Size (in bytes):	617489

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.

PATENT
Attorney Docket No. 38075-716.301

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application:

Inventor: Emil Kartalov, et al.

Application No.: 14/451,876

Filed: August 5, 2014

Title: **SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS**

Confirmation No.: **1065**

Customer No. **021971**

FILED ELECTRONICALLY ON: November 14, 2014

RESPONSE TO NOTICE TO FILE MISSING PARTS

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

In response to the Notice to File Missing Parts of Non-Provisional Application mailed Augusted 15, 2014, Applicants respond as follows:

1. ☐ Copy of Notice to File Missing Parts is provided
2. ☒ **Filing Fees are authorized herewith**
3. ☐ **Applicant claims small entity status**. See 37 CFR 1.27.
4. ☐ **Replacement Specification is provided**/*Total Pages]* Both the claims and abstract must start on a new page. *(For information on the preferred arrangement, see MPEP 608.01(a))*
5. ☐ **Replacement Drawing(s)** (35 U.S.C. 113) is provided *[Total Pages]*
6. ☒ **Oath or Declaration is provided** *[Total Pages 03]*
   a. ☐ Newly executed (original or copy)
   b. ☒ Copy from a prior application (37 CFR 1.63(d))
7. ☐ **Information Disclosure** (PTO/SB/08 or PTO-1449) is provided
   ☐ Copies of citations attached
8. ☒ **Preliminary Amendment is provided**
9. ☐ **CD-Rom or CD-R** in duplicate, large table or Computer Program *(Appendix)* is provided
   ☐ Landscape table on CD
10. **Nucleotide and/or Amino Acid Sequence Submission is provided** *(if applicable, items a.-c. are required)*
    a. ☒ Computer Readable Form (CRF)
    b. ☐ Specification Sequence Listing on:
       i. ☐ CD-ROM or CD-R (2 copies); or
       ii. ☒ Paper
    c. ☒ Statement verifying identity of above copies
11. ☐ **Assignment Papers** (cover sheet & document(s) are provided
12. ☐ **Power of Attorney is provided**
    ☐ **37 CFR 3.73(c) Statement** *(when there is an assignee)*
13. ☐ **English Translation Document** *(if applicable)* is provided
14. ☐ **Certified Copy of Priority Document(s)** *(if foreign priority is claimed)* is provided
15. ☐ Applicants request an Extension of Time under 37 C.F.R. §1.136 of:
    ☐ 1 Month ☐ 2 Months ☐ 3 Months
    ☐ 4 Months ☐ 5 Months
16. ☐ Other: **REPLACEMENT SHEET**

FEE AUTHORIZATION

The Commissioner is authorized to charge any additional fees which may be required, including petition fees and extension of time fees, to Deposit Account No. **23-2415** (Docket No. 38075-716.301).

Respectfully submitted,

Date: November 14, 2014

By: /Reena Bajpai/
Reena Bajpai, Ph.D. Patent Agent
Registration No.: 69,996

WILSON SONSINI GOODRICH & ROSATI
650 Page Mill Road
Palo Alto, CA 94304-1050
Direct Dial: (650) 493-9300

6630426_1.doc

DocCode – SEQ.TXT

SCORE Placeholder Sheet for IFW Content

Application Number: 14451876 Document Date: 11/14/2014

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/SB/06 (09-11)
Approved for use through 1/31/2014. 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.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 14/451,876	Filing Date 08/05/2014	☐ To be Mailed

ENTITY: ☐ LARGE ☒ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =	
☐ 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).			
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	11/14/2014						
	Total (37 CFR 1.16(i))	* 27	Minus	** 27	= 0	X $40 =	0
	Independent (37 CFR 1.16(h))	* 3	Minus	***3	= 0	X $210 =	0
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						TOTAL ADD'L FEE	**0**

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =	
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						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 number found in the appropriate box in column 1.

LIE
/RAMONA WILSON/

This collection of information is required by 37 CFR 1.16. 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 12 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.

```
============================================================
Sequence Listing was accepted.
If you need help call the Patent Electronic Business Center at (866)
217-9197 (toll free).
Reviewer: Anjum, Durreshwar (CGI Federal)
Timestamp: [year=2014; month=11; day=17; hr=9; min=4; sec=3; ms=306; ]
============================================================
```

Validated By CRFValidator v 1.0.4

Application No: 14451876 Version No: 1.0

Input Set:

Output Set:

Started: 2014-11-14 19:41:48.064
Finished: 2014-11-14 19:41:49.500
Elapsed: 0 hr(s) 0 min(s) 1 sec(s) 436 ms
Total Warnings: 30
Total Errors: 0
No. of SeqIDs Defined: 32
Actual SeqID Count: 32

Error code	Error Description
W 402	Undefined organism 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 402	Undefined organism 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 (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 402	Undefined organism 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)
W 213	Artificial or Unknown found in <213> in SEQ ID (21)
W 213	Artificial or Unknown found in <213> in SEQ ID (22)

Input Set:

Output Set:

Started: 2014-11-14 19:41:48.064
Finished: 2014-11-14 19:41:49.500
Elapsed: 0 hr(s) 0 min(s) 1 sec(s) 436 ms
Total Warnings: 30
Total Errors: 0
No. of SeqIDs Defined: 32
Actual SeqID Count: 32

Error code	Error Description
W 213	Artificial or Unknown found in <213> in SEQ ID (23)
W 213	Artificial or Unknown found in <213> in SEQ ID (24)
W 213	Artificial or Unknown found in <213> in SEQ ID (25) This error has occured more than 20 times, will not be displayed
W 402	Undefined organism found in <213> in SEQ ID (26)

SEQUENCE LISTING

<110> KARTALOV, EMIL P.
RAJAGOPAL, ADITYA
SCHERER, AXEL

<120> SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

<130> 38075-716.301
<140> US 14/451,876
<141> 2014-08-05

<150> 13/756,760
<151> 2013-02-01

<150> 61/703,093
<151> 2012-09-19

<150> 61/594,480
<151> 2012-02-03

<160> 32

<170> PatentIn version 3.5

<210> 1
<211> 198
<212> DNA
<213> Human immunodeficiency virus 1

<400> 1
ggaagctcta ttagatacag gagcagatga tacagtatta gaagaaatga gtttgccagg 60

aagatggaaa ccaaaaatga taggggggaat tggaggtttt atcaaagtaa gacagtatga 120

tcagatactc atagaaatct gtggacataa agctataggt acagtattag taggacctac 180

acctgtcaac ataattgg 198

<210> 2
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 2
ggaagctcta ttagatacag gagcag 26

<210> 3
<211> 26
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 3
ccaattatgt tgacaggtgt aggtcc 26

<210> 4
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 4
tgagtttgcc aggaagatgg aaacca 26

<210> 5
<211> 199
<212> DNA
<213> Human immunodeficiency virus 1

<400> 5
cagctacaac catcccttca gacaggatca gaagaactta gatcattata taatacagta 60

gcaaccctct attgtgtgca tcaaaggata gagataaaag acaccaagga agctttagac 120

aagatagagg aagagcaaaa caaaagtaag aaaaaagcac agcaagcagc agctgacaca 180

ggacacagca atcaggtca 199

<210> 6
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 6
cagctacaac catcccttca gaca 24

<210> 7
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 7
tgacctgatt gctgtgtcct gtgt 24

<210> 8
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 8
agcaaccctc tattgtgtgc atcaaagg 28

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy3

<220>
<223> 3'BHQ_2

<400> 9
aaagcacagc aagcagcagc tga 23

<210> 10
<211> 199
<212> DNA

<213> Plasmodium falciparum

<400> 10
gcctaacatg gctatgacgg gtaacgggga attagagttc gattccggag agggagcctg 60

agaaatagct accacatcta aggaaggcag caggcgcgta aattacccaa ttctaaagaa 120

gagaggtagt gacaagaaat aacaatgcaa ggccaattta aaaccttccc agagtaacaa 180

ttggagggca agtctggtg 199

<210> 11
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 11
gcctaacatg gctatgacgg gtaa 24

<210> 12
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 12
caccagactt gccctccaat tgtt 24

<210> 13
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 13
attccggaga gggagcctga gaaata 26

<210> 14
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-ROXN

<220>
<223> 3'BHQ_2

<400> 14
aaggaaggca gcaggcgcgt aaatta 26

<210> 15
<211> 192
<212> DNA
<213> Herpes simplex virus 2

<400> 15
tcagcccatc ctccttcggc agtatggagg gtgtcgcggc ggcgagccgc cgtccccaaa 60

gacgtgcggg tcgtacacgt acacgtacca gggcggcggg cctccgaccc ggtacgctct 120

cgtaaatgct tccctgctgg tgccgatctg ggaccgcgcc gcggagacat tcgagtacca 180

gatcgaactc gg 192

<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 16
tcagcccatc ctccttcggc agtat 25

<210> 17
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 17
ccgagttcga tctggtactc gaatgt 26

<210> 18
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 18
aaagacgtgc gggtcgtaca cgtaca 26

<210> 19
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy5

<220>
<223> 3'3IAbRQSp

<400> 19
taaatgcttc cctgctggtg ccgat 25

<210> 20
<211> 200
<212> DNA
<213> Mycobacterium tuberculosis

<400> 20
gagtgcaaag acaaggacat gacgtacgcg gccccgctgt tcgtcacggc cgagttcatc 60

aacaacaaca ccggcgagat caagagccag acggtgttca tgggtgactt cccgatgatg 120

accgagaagg gcaccttcat catcaacggc accgagcgcg tcgtggtcag ccagctggtc 180

cgctcgcccg gtgtgtactt 200

<210> 21
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 21
gagtgcaaag acaaggacat gacg 24

<210> 22
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 22
aagtacacac cgggcgagc 19

<210> 23
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 23
cgctgttcgt cacggccgag ttcat 25

<210> 24
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'Cy3

<220>
<223> 3'BHQ_2

<400> 24
agatcaagag ccagacggtg ttcatg 26

<210> 25
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy5

<220>
<223> 3'IAbRQSp

<400> 25
aagggcacct tcatcatcaa cggca 25

<210> 26
<211> 200
<212> DNA
<213> Dengue virus 3

<400> 26
atgccaactg tgattgagca cttagaaaga ctacaaagga aacatggagg aatgcttgtg 60

agaaatccac tctcacgaaa ctccacgcac gaaatgtatt ggatatccaa tggtacaggc 120

aacatcgtct cattcacaat gacacacagg agacccacca tagagaaaga tgtggattta 180

ggagcaggaa cccgacatgt 200

<210> 27
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 27

atgccaactg tgattgagca ct 22

<210> 28
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 28
acatgtcggg ttcctgctcc taaa 24

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

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'6-FAM

<220>
<223> 3'BHQ_1

<400> 29
acaaaggaaa catggaggaa tgcttgtga 29

<210> 30
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic probe

<220>
<223> 5'Cy3

<220>
<223> 3'BHQ_2

<400> 30
actctcacga aactccacgc acgaaa 26

<210> 31

<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'6-ROXN

<220>
<223> 3'BHQ_2

<400> 31
acaatgacac acaggagacc caccat 26

<210> 32
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
probe

<220>
<223> 5'Cy5

<220>
<223> 3'IAbRQSp

<400> 32
ggatatccaa tggtacaggc aacatcgt 28

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
14/451,876	08/05/2014	1677	1080	38075-716.301	27	3

CONFIRMATION NO. 1065

UPDATED FILING RECEIPT

21971
WILSON, SONSINI, GOODRICH & ROSATI
650 PAGE MILL ROAD
PALO ALTO, CA 94304-1050

OC000000071902843

Date Mailed: 11/21/2014

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)

Emil Kartalov, Pasadena, CA;
Aditya Rajagopal, Irvine, CA;
Axel Scherer, Barnard, VT;

Applicant(s)

California Institute of Technology, Pasadena, CA

Power of Attorney: None

Domestic Priority data as claimed by applicant

This application is a CON of 13/756,760 02/01/2013 PAT 8838394
which claims benefit of 61/703,093 09/19/2012
and claims benefit of 61/594,480 02/03/2012

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 - A proper **Authorization to Permit Access to Application by Participating Offices** (PTO/SB/39 or its equivalent) has been received by the USPTO.

If Required, Foreign Filing License Granted: 08/14/2014
The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 14/451,876**
Projected Publication Date: 02/26/2015

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 153

Non-Publication Request: No

Early Publication Request: No
** **SMALL ENTITY** **
Title

SIGNAL ENCODING AND DECODING IN MULTIPLEXED BIOCHEMICAL ASSAYS

Preliminary Class

436

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).

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 154

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: 14/451,876

APPLICATION AS FILED - PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	70		N/A	
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	300		N/A	
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	360		N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	27 minus 20 =	* 7	x 40 =	280	OR		
INDEPENDENT CLAIMS (37 CFR 1.16(h))	3 minus 3 =	*	x 210 =	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).			0.00			
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	1010		TOTAL	

APPLICATION AS AMENDED - PART II

AMENDMENT A

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY ADDITIONAL FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE($)
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	

AMENDMENT B

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)		RATE($)	ADDITIONAL FEE($)
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.

PTO/SB/08a (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons 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	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	1	of	9	Attorney Docket Number	38075-716.301

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.	US 2002/0022273 A1	02/21/2002	Empedocles et al.	
	2.	US 2003/0148280 A1	08/07/2003	Harris et al.	
	3.	US 2003/0148544 A1	08/07/2003	Nie et al.	
	4.	US 2005/0064435 A1	03/24/2005	Su et al.	
	5.	US 2005/0164264 A1	07/28/2005	Shipwash	
	6.	US 2005/0214753 A1	09/29/2005	Schultz et al.	
	7.	US 2005/0260640 A1	11/24/2005	Andersen et al.	
	8.	US 2006/0216708 A1	09/28/2006	Venema	
	9.	US 2007/0178485 A1	08/02/2007	El-deiry et al.	
	10.	US 2007/0231824 A1	10/04/2007	Chee et al.	
	11.	US 2008/0050737 A1	02/28/2008	Arieli et al.	
	12.	US 2008/0069733 A1	03/20/2008	Maltezos et al.	
	13.	US 2009/0048785 A1	02/19/2009	Katzir et al.	
	14.	US 2009/0062129 A1	03/05/2009	McKernan et al.	
	15.	US 2010/0015607 A1	01/21/2010	Geiss et al.	
	16.	US 2010/0047924 A1	02/25/2010	Webster et al.	
	17.	US 2010/0112710 A1	05/06/2010	Geiss et al.	
	18.	US 2010/0159447 A1	06/24/2010	Li et al.	
	19.	US 2010/0210472 A1	08/19/2010	Empedocles et al.	
	20.	US 2010/0248257 A1	09/30/2010	Jacobsen et al.	
	21.	US 2010/0261026 A1	10/14/2010	Ferree et al.	
	22.	US 2010/0267064 A1	10/21/2010	Kartalov et al.	
	23.	US 2010/0273173 A1	10/28/2010	Hirai et al.	
	24.	US 2011/0104684 A1	05/05/2011	Hooper	
	25.	US 2011/0151459 A1	06/23/2011	Rothmann et a.	
	26.	US 2011/0207623 A1	08/25/2011	Dimitrov et al.	
	27.	US 2011/0237459 A1	09/29/2011	Nova et al.	
	28.	US 2012/0003646 A1	01/05/2012	Joo et al.	
	29.	US 2012/0045756 A1	02/23/2012	Rothmann et a.	

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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 157

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	2	of	9	Attorney Docket Number	38075-716.301

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
	30.	US 2012/0077195 A1	03/29/2012	Li et al.	
	31.	US 2012/0077692 A1	03/29/2012	Hassibi et al.	
	32.	US 2012/0122704 A1	05/17/2012	Atchley et al.	
	33.	US 2012/0141995 A1	06/07/2012	Li et al.	
	34.	US 2012/0171677 A1	07/05/2012	Ludowise	
	35.	US 2012/0184017 A1	07/19/2012	Chatterjee	
	36.	US 2012/0190030 A1	07/26/2012	Chun et al.	
	37.	US 2012/0196283 A1	08/02/2012	Babiel et al.	
	38.	US 2012/0252017 A1	10/04/2012	Reed et al.	
	39.	US 2012/0258457 A1	10/11/2012	Jarosch et al.	
	40.	US 2013/0116780 A1	08/08/2013	Miller et al.	
	41.	US 5,079,351	01/07/1992	Sninsky et al.	
	42.	US 5,538,848	07/23/1996	Livak et al.	
	43.	US 5,618,711	04/08/1997	Gelfand et al.	
	44.	US 5,677,152	10/14/1997	Birch et al.	
	45.	US 5,723,591	03/03/1998	Livak et al.	
	46.	US 5,773,258	06/30/1998	Birch et al.	
	47.	US 5,789,224	08/04/1998	Gelfand et al.	
	48.	US 5,804,375	09/08/1998	Gelfand et al.	
	49.	US 5,876,930	03/02/1999	Livak et al.	
	50.	US 5,928,862	07/27/1999	Morrison	
	51.	US 5,948,360	09/07/1999	Roa et al.	
	52.	US 5,981,180	11/09/1999	Chandler et al.	
	53.	US 5,994,056	11/30/1999	Higuchi	
	54.	US 6,030,787	02/29/2000	Livak et al.	
	55.	US 6,127,155	10/03/2000	Gelfand et al.	
	56.	US 6,171,785	01/09/2001	Higuchi	
	57.	US 6,258,569	07/10/2001	Livak et al.	
	58.	US 6,534,266	03/18/2003	Singer	

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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 158

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	3	of	9	Attorney Docket Number	38075-716.301

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
	59.	US 6,548,259	04/15/2003	Ward et al.	
	60.	US 6,642,062	11/04/2003	Kauvar et al.	
	61.	US 7,101,663	09/05/2006	Godfrey et al.	
	62.	US 7,141,377	11/28/2006	Gelfand et al.	
	63.	US 7,348,141	03/25/2008	French et al.	
	64.	US 7,385,043	06/10/2008	Kramer	
	65.	US 7,410,764	08/12/2008	Gocke et al.	
	66.	US 7,413,708	08/19/2008	Mayrand	
	67.	US 7,473,767	01/06/2009	Dimitrov	
	68.	US 7,575,864	08/18/2009	Bedzyk et al.	
	69.	US 7,667,024	02/23/2010	Mao et al.	
	70.	US 7,767,423	08/03/2010	Kopreski	
	71.	US 7,771,949	08/10/2010	Kramer	
	72.	US 7,919,237	04/05/2011	Dimitrov et al.	
	73.	US 7,919,244	04/05/2011	Madejon et al.	
	74.	US 7,941,279	05/10/2011	Hwang et al.	
	75.	US 8,148,512	04/03/2012	Dimitrov	

UNPUBLISHED PATENT APPLICATIONS

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
	76.	N/A.			

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]
	77.	EP 1448581 B1	11/12/2008	Dimitrov et al.		
	78.	EP 1963531 B1	09/21/2011	Geiss et al.		
	79.	WO 02/056014 A2	07/18/2002	Nelsen et al.		
	80.	WO 02/056014 A3	10/09/2003	Nelsen 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 159

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	4	of	9	Attorney Docket Number	38075-716.301

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 2007/076128 A2	07/05/2007	Geiss et al.		
	82.	WO 2007/076128 A3	11/29/2007	Geiss et al.		
	83.	WO 2007/076129 A2	07/05/2007	Geiss et al.		
	84.	WO 2007/076129 A3	03/06/2008	Geiss et al.		
	85.	WO 2007/076132 A2	07/05/2007	Ferree et al.		
	86.	WO 2007/076132 A3	09/07/2009	Ferree et al.		
	87.	WO 2007/139766 A2	12/06/2007	Hwang et al.		
	88.	WO 2007/139766 A3	12/11/2008	Hwang et al.		
	89.	WO 2008/124847 A2	10/16/2008	Geiss et al.		
	90.	WO 2008/124847 A3	02/19/2009	Geiss et al.		
	91.	WO 2010/019826 A1	02/18/2010	Webster et al.		
	92.	WO 2011/047087 A2	04/21/2011	Dunaway		
	93.	WO 2011/047087 A3	08/18/2011	Dunaway		
	94.	WO 2011/100541 A2	08/18/2011	Webster et al.		
	95.	WO 2011/100541 A3	01/05/2012	Webster et al.		
	96.	WO 2011/116088 A2	09/22/2011	Hengen et al.		
	97.	WO 2011/116088 A3	02/02/2012	Hengen et al.		
	98.	WO 2012/058638 A2	05/03/2012	Lin et al.		
	99.	WO 2012/058638 A3	12/20/2012	Lin et al.		
	100.	WO 2012/106428 A2	08/09/2012	Snyder et al.		
	101.	WO 2012/135340 A2	10/04/2012	Webster et al.		
	102.	WO 2012/135340 A3	12/20/2012	Webster et al.		
	103.	WO 99/19515 A1	04/22/1999	Chandler et al.		
	104.	WO 99/52708 A1	10/21/1999	Chandler		

Continued on the next page with more references.

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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 160

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	5	of	9	Attorney Docket Number	38075-716.301

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the 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.	BEIGE, et al. Clinical evaluation of a Mycobacterium tuberculosis PCR assay. J Clin Microbiol. 1995 Jan;33(1):90-5.	
	106.	CHAMBERLAIN, et al. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. Nucleic Acids Res. 1988 Dec 9;16(23):11141-56.	
	107.	CHEN, et al. A Homogeneous, ligase-mediated DNA diagnostic test. Genome Research, 1998, volume 8, pages 549-556.	
	108.	CHONG, et al. Single-tube multiplex-PCR screen for common deletional determinants of alpha-thalassemia. Blood. 2000 Jan 1;95(1):360-2.	
	109.	Chromatogram, 2011, 2 pages. Dorland's illustrated medical dictionary. Retrieved online on 22 January 2014 from «http://www.credoreference.com».	
	110.	CHUN, et al. Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. Nucleic Acids Res. 2007;35(6):e40. Epub 2007 Feb 7.	
	111.	CRAIG, et al. Ordering of cosmid clones covering Herpes simplex virus type I (HSV-I) genome: a test case for fingerprinting by hybridisation. Nucleic Acids Research, 1990, volume 18, pages 2653-2660.	
	112.	DOS SANTOS, et al. A simple one-step real-time RT-PCR for diagnosis of dengue virus infection. J Med Virol. 2008 Aug;80(8):1426-33. doi: 10.1002/jmv.21203.	
	113.	EL-HAJJ, et al. Detection of rifampin resistance in Mycobacterium tuberculosis in a single tube with molecular beacons. J Clin Microbiol. 2001 Nov;39(11):4131-7.	
	114.	EMBL-Bank: AJ303204. http://www.ebi.ac.uk/Tools/dbfetch/emblfetch?db=embl&id=AJ303204&format=default&style=default&Retrieve=Retrieve. Accessed February 2012.	
	115.	EMBL-Bank: GQ395623. http://www.ebi.ac.uk/Tools/dbfetch/emblfetch?db=embl&id=GQ395623&format=default&style=default&Retrieve=Retrieve. Accessed February 2012.	

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]Applicant is to place a checkmark here if English language Translation is attached.

This collection of information is required by 37 CFR 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 161

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	6	of	9	Attorney Docket Number	38075-716.301

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	116.	FODOR, et al. Multiplexed biochemical assays with biological chips. Nature. 1993 Aug 5;364(6437):555-6.	
	117.	FORTINA, et al. Digital mRNA profiling. Nat Biotechnol. 2008 Mar;26(3):293-4. doi: 10.1038/nbt0308-293.	
	118.	GenBank: M93130.1. Dengue type 3 virus complete genome RNA, complete cds. http://www.ncbi.nlm.nih.gov/nuccore/M93130. Accessed February 2012.	
	119.	HAN, et al. Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. Nat Biotechnol. 2001 Jul;19(7):631-5.	
	120.	HARTMAN, et al. Development of a novel internal positive control for Taqman based assays. Mol Cell Probes. 2005 Feb;19(1):51-9. Epub 2004 Dec 10.	
	121.	HEIDARI, et al. Detection of Plasmodium falciparum Directly from Blood Samples Using the Polymerase Chain Reaction. Journal of Sciences, Islamic Republic of Iran. 2005 16(1):21-24.	
	122.	HENEGARIU, et al. Multiplex PCR: critical parameters and step-by-step protocol. Biotechniques. 1997 Sep;23(3):504-11.	
	123.	HIV databases. http://www.hiv.lanl.gov/content/index. Accessed February 2012.	
	124.	HOREJSH, et al. A molecular beacon, bead-based assay for the detection of nucleic acids by flow cytometry. Nucleic Acids Res. 2005 Jan 19;33(2):e13.	
	125.	HUANG, et al. Identification of 8 foodborne pathogens by multicolor combinational probe coding technology in a single real-time PCR. Clin Chem. 2007 Oct;53(10):1741-8. Epub 2007 Aug 10.	
	126.	HUANG, et al. Multicolor combinatorial probe coding for real-time PCR. PLoS One. 2011 Jan 14;6(1):e16033. doi: 10.1371/journal.pone.0016033.	
	127.	International search report and written opinion dated 04/12/2013 for PCT/US2013/024509.	
	128.	JOTHIKUMAR, et al. Design of FRET-TaqMan probes for multiplex real-time PCR	

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]Applicant is to place a checkmark here if English language Translation is attached.

This collection of information is required by 37 CFR 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 162

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	7	of	9	Attorney Docket Number	38075-716.301

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
		using an internal positive control. Biotechniques. 2009 Jun;46(7):519-24. doi: 10.2144/000113127.	
	129.	KUHN, et al. Hybridization of DNA and PNA molecular beacons to single-stranded and double-stranded DNA targets. J Am Chem Soc. 2002 Feb 13;124(6):1097-103.	
	130.	LAO, et al. Multiplexing RT-PCR for the detection of multiple miRNA species in small samples. Biochem Biophys Res Commun. 2006 Apr 28;343(1):85-9. Epub 2006 Feb 28.	
	131.	LEE, et al. Novel multiplex PCR using dual-priming oligonucleotides for detection and discrimination of the Mycobacterium tuberculosis complex and M. bovis BCG. J Clin Microbiol. 2010 Dec;48(12):4612-4. doi: 10.1128/JCM.00872-10. Epub 2010 Oct 13.	
	132.	LEE, et al. Seven-color, homogeneous detection of six PCR products. Biotechniques. 1999 Aug;27(2):342-9.	
	133.	LIEW, et al. Validating a custom multiplex ELISA against individual commercial immunoassays using clinical samples. Biotechniques. 2007 Mar;42(3):327-8, 330-3.	
	134.	LIVAK, et al. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl. 1995 Jun;4(6):357-62.	
	135.	MORRISON, et al. Two-color ratio-coding of chromosome targets in fluorescence in situ hybridization: quantitative analysis and reproducibility. Cytometry. 1997 Apr 1;27(4):314-26.	
	136.	NOORDHOEK, et al. Sensitivity and specificity of PCR for detection of Mycobacterium tuberculosis: a blind comparison study among seven laboratories. J Clin Microbiol. 1994 Feb;32(2):277-84.	
	137.	Office action dated 01/24/2014 for US Application No. 13/756,760.	
	138.	OLIVEIRA, et al. Multiplex PCR strategy for rapid identification of structural types and variants of the mec element in methicillin-resistant Staphylococcus aureus. Antimicrob Agents Chemother. 2002 Jul;46(7):2155-61.	
	139.	OU, et al. DNA amplification for direct detection of HIV-1 in DNA of peripheral blood	

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]Applicant is to place a checkmark here if English language Translation is attached.

This collection of information is required by 37 CFR 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 163

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	8	of	9	Attorney Docket Number	38075-716.301

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
		mononuclear cells. Science. 1988 Jan 15;239(4837):295-7.	
	140.	PATON, et al. Detection and characterization of Shiga toxigenic Escherichia coli by using multiplex PCR assays for stx1, stx2, eaeA, enterohemorrhagic E. coli hlyA, rfbO111, and rfbO157. J Clin Microbiol. 1998 Feb;36(2):598-602.	
	141.	PATTERSON, et al. Detection of HIV-1 DNA and messenger RNA in individual cells by PCR-driven in situ hybridization and flow cytometry. Science. 1993 May 14;260(5110):976-9.	
	142.	PETERSEN, et al. Short PNA molecular beacons for real-time PCR allelic discrimination of single nucleotide polymorphisms. Mol Cell Probes. 2004 Apr;18(2):117-22.	
	143.	Plasmodium falciparum (Plasmodium falciparum) Genome Browser Gateway. http://microbes.ucsc.edu/cgi-bin/hgGateway?hgsid=612764&clade=eukaryota-protista&org=0&db=0. Accessed February 2012.	
	144.	PTAK, et al. Inhibition of human immunodeficiency virus type 1 replication in human cells by Debio-025, a novel cyclophilin binding agent. Antimicrob Agents Chemother. 2008 Apr;52(4):1302-17. doi: 10.1128/AAC.01324-07. Epub 2008 Jan 22.	
	145.	ROSENSTRAUS, et al. An internal control for routine diagnostic PCR: design, properties, and effect on clinical performance. J Clin Microbiol. 1998 Jan;36(1):191-7.	
	146.	ROTH, et al. Feasibility and efficacy of routine PCR screening of blood donations for hepatitis C virus, hepatitis B virus, and HIV-1 in a blood-bank setting. Lancet. 1999 Jan 30;353(9150):359-63.	
	147.	SAMBROOK, et al. Molecular Cloning: A Laboratory Manual.	
	148.	SANG, et al. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463-7.	
	149.	SPEICHER, et al. Karyotyping human chromosomes by combinatorial multi-fluor FISH. Nat Genet. 1996 Apr;12(4):368-75.	
	150.	TIRASOPHON, et al. A novel detection of a single Plasmodium falciparum in infected	

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]Applicant is to place a checkmark here if English language Translation is attached.

This collection of information is required by 37 CFR 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 164

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				***Complete if Known***	
				Application Number	14/451,876
				Filing Date	August 5, 2014
				First Named Inventor	Emil Kartalov
				Art Unit	1631
				Examiner Name	Unassigned
Sheet	9	of	9	Attorney Docket Number	38075-716.301

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
		blood. Biochem Biophys Res Commun. 1991 Feb 28;175(1):179-84.	
	151.	TYAGI, et al. Multicolor molecular beacons for allele discrimination. Nat Biotechnol. 1998 Jan;16(1):49-53.	
	152.	TYAGI, et al. Wavelength-shifting molecular beacons. Nat Biotechnol. 2000 Nov;18(11):1191-6.	
	153.	URDEA, et al. Requirements for high impact diagnostics in the developing world. Nature. 2006 Nov 23;444 Suppl 1:73-9.	
	154.	VET, et al. Multiplex detection of four pathogenic retroviruses using molecular beacons. Proc Natl Acad Sci U S A. 1999 May 25;96(11):6394-9.	
	155.	WANG, et al. Locked nucleic acid molecular beacons. J Am Chem Soc. 2005 Nov 16;127(45):15664-5.	
	156.	WATERS, et al. Microchip device for cell lysis, multiplex PCR amplification, and electrophoretic sizing. Anal Chem. 1998 Jan 1;70(1):158-62.	
	157.	WEIDMANN, et al. Rapid detection of herpes simplex virus and varicella-zoster virus infections by real-time PCR. J Clin Microbiol. 2003 Apr;41(4):1565-8.	
	158.	WIESE, et al. Simultaneous multianalyte ELISA performed on a microarray platform. Clin Chem. 2001 Aug;47(8):1451-7.	
	159.	ZHANG, et al. Novel Multiplex PCR Assay for Characterization and Concomitant Subtyping of Staphylococcal Cassette Chromosome mec Types I to V in Methicillin-Resistant Staphylococcus aureus. J Clin Microbiol. 2005 Oct;43(10):5026-33.	

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]Applicant is to place a checkmark here if English language Translation is attached.

This collection of information is required by 37 CFR 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 165

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 448 581 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
12.11.2008 Bulletin 2008/46

(21) Application number: **02763230.6**

(22) Date of filing: **03.07.2002**

(51) Int Cl.:
C07H 21/02 *(2006.01)* ***C12Q 1/68*** *(2006.01)*

(86) International application number:
PCT/US2002/021278

(87) International publication number:
WO 2003/003810 (16.01.2003 Gazette 2003/03)

(54) METHODS FOR DETECTION AND QUANTIFICATION OF ANALYTES IN COMPLEX MIXTURES

VERFAHREN ZUM NACHWEIS UND ZUR QUANTIFIZIERUNG VON ANALYTEN IN KOMPLEXEN MISCHUNGEN

PROCEDES DE DETECTION ET DE QUANTIFICATION D'ANALYTES DANS DES MELANGES COMPLEXES

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR

(30) Priority: **03.07.2001 US 898743**

(43) Date of publication of application:
25.08.2004 Bulletin 2004/35

(73) Proprietor: **The Institute for Systems Biology**
Seattle, WA 98103-8904 (US)

(72) Inventors:
- **DIMITROV, Krassen**
  Seattle, WA 98133 (US)
- **DUNAWAY, Dwayne**
  Seattle, WA 98115 (US)

(74) Representative: **UEXKÜLL & STOLBERG**
Patentanwälte
Beselerstrasse 4
22607 Hamburg (DE)

(56) References cited:
WO-A-99/47643
WO-A-99/62926
WO-A1-99/19515
WO-A2-02/056014
US-A- 5 401 847
US-A- 5 487 973

- **HIRSCHHORN JOEL N ET AL: "SBE-TAGS: An array-based method for efficient single-nucleotide polymorphism genotyping" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, DC, US, vol. 97, no. 22, 24 October 2000 (2000-10-24), pages 12164-12169, XP002158215 ISSN: 0027-8424**
- **BRENNER SYDNEY ET AL: "Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays" NATURE BIOTECHNOLOGY, NATURE PUB. CO, NEW YORK, NY, US, vol. 18, no. 6, June 2000 (2000-06), pages 630-634, XP002199492 ISSN: 1087-0156**
- **COLLINS M L ET AL: "A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 25, no. 15, 1997, pages 2979-2984, XP002208495 ISSN: 0305-1048**
- **CHEN J ET AL: "A MICROSPHERE-BASED ASSAY FOR MULTIPLEXED SINGLE NUCLEOTIDE POLYMORPHISM ANALYSIS USING SINGLE BASE CHAIN EXTENSION" GENOME RESEARCH, COLD SPRING HARBOR LABORATORY PRESS, WOODBURY, NY, US, vol. 10, no. 4, April 2000 (2000-04), pages 549-557, XP000927257 ISSN: 1088-9051**
- **TONG ANTHONY K ET AL: "Combinatorial fluorescence energy transfer tags for multiplex biological assays" NATURE BIOTECHNOLOGY, NATURE PUB. CO, NEW YORK, NY, US, vol. 19, no. 8, August 2001 (2001-08), pages 756-759, XP002193783 ISSN: 1087-0156**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

EP 1 448 581 B1

• BONETTA L: "Gene expression: One size does not fit all" NATURE METHODS 2006 UNITED KINGDOM, vol. 3, no. 5, 2006, pages 401-409, XP008070014 ISSN: 1548-7091 1549-1676

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 167

Description

BACKGROUND OF THE INVENTION

[0001] This invention relates generally to the field of genomics and, more specifically to detection, identification, and quantification of target analytes in mixtures.

[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. These variations in gene expression are at the core of altered physiologic and pathologic processes. Therefore, identifying and quantifying the expression of genes in normal cells compared to diseased cells can aid the discovery of new drug and diagnostic targets.

[0003] Nucleic acids can be detected and quantified based on their specific polynucleotide sequences. The basic principle underlying existing methods of detection and quantification is the hybridization of a labeled complementary probe sequence to a target sequence of interest in a sample. The formation of a duplex indicates the presence of the target sequence in the sample and the degree of duplex formation, as measured by the amount of label incorporated in it, is proportional to the amount of the target sequence.

[0004] This technique, called molecular hybridization, has been a useful tool for identifying and analyzing specific nucleic acid sequences in complex mixtures. This technique has been used in diagnostics, for example, to detect nucleic acid sequences of various microbes in biological samples. In addition, hybridization techniques have been used to map genetic differences or polymorphisms between individuals. Furthermore, these techniques have been used to monitor changes in gene expression in different populations of cells or in cells treated with different agents.

[0005] In the past, only a few genes could be detected in a complex sample at one time. However, DNA microarrays, devices that consist of thousands of immobilized DNA sequences present on a miniaturized surface, have made this process more efficient. Using a microarray, it is possible in a single experiment to detect the presence or absence of thousands of genes in a biological sample. This allows researchers to simultaneously perform several diagnostic tests on one sample, or to observe expression level changes in thousands of genes in one experiment. Generally, microarrays are prepared by binding DNA sequences to a surface such as a nylon membrane or glass slide at precisely defined locations on a grid. Then nucleic acids in a biological sample are labeled and hybridized to the array. The labeled sample DNA marks the exact position on the array where hybridization occurs, allowing automatic detection.

[0006] Unfortunately, despite the miniaturization of array formats, this method still requires significant amounts of the biological sample. However, in several cases, such as biopsies of diseased tissues or samples of a discrete cell type, the biological sample is in limited supply. In addition, the kinetics of hybridization on the surface of a microarray is less efficient than hybridization in small amounts of aqueous solution. Furthermore, microarrays require a large dynamic range of detection to account for large difference in abundance of the different molecular species. This results in decreased sensitivity since there is a trade-off between sensitivity and dynamic range. A further problem with microarray methods is that the output is quantitative analog data that has undergone several intermediary transformations. In microarrays, the amount of nucleic acid hybridized to each spot is determined by measuring its label and so any nonlinear correlation between the amount of DNA hybridized and the amount of the label detected will skew the data output. Such non-linearity has been widely documented.

[0007] Collins et al. (Nucleic Acids Research 25(15): 2979- 2984 (1997)) disclose target-specific oligonucleotides called label extenders which include a non-target extension which bind to a preamplifier which in turn binds to many amplifiers so as to generate signal amplification; the same label is bound to all label extenders, wherein detection is based on immobilization (localization by immobilization). The method disclosed by Collins et al. does not provide for simultaneous (multiplexing) signal amplification for direct detection of nucleic acids, in particular when present in complex mixtures.

[0008] Thus, there exists a need for accurate and sensitive detection, identification and quantification of analytes in complex mixtures. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

[0009] The invention provides a diverse population of uniquely labeled probes, containing about thirty or more target specific nucleic acid probes each attached to a unique label bound to a nucleic acid. Also provided is a method of detecting a nucleic acid analyte. The method consists of (a) contacting a mixture of nucleic acid analytes with the diverse population of uniquely labelled probes of the invention under conditions sufficient for hybridization of the probes to target nucleic acid analytes; (b) measuring the resulting signal from one or more of the target specific probes hybridized to an analyte , where the signal uniquely identifies the analyte.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 168

BRIEF DESCRIPTION OF THE DRAWINGS

[0010]

Figure 1 shows components of a specifier. Figure 1A shows an association between a target and a labeled specifier. Figure 1B shows a structure of a genedigit. Figure 1C shows an association between a genedigit and labeled anti-genedigit.

Figure 2 shows a DNA probe molecule (specifier) that contains several different types of spatially separated labels.

Figure 3 shows DNA probe molecules containing a label on one end that have been stretched on a coverslip using a flow-stretch technique.

Figure 4 shows labeled DNA molecules that have been aligned using an electro-stretch technique. Figure 4A shows labeled DNA probe molecules in a flow cell with the voltage off. Figure 4B shows the labeled DNA probe molecule in the flow cell with the voltage on.

Figure 5 shows a schematic of probes being aligned by constriction in the flow of a liquid containing the probes in the presence of an oscillating electric field.

DETAILED DESCRIPTION OF THE INVENTION

[0011] This invention is directed to a diverse population of uniquely labeled probes, comprising thirty or more target-specific nucleic acid probes, each attached to a unique label bound to a nucleic acid, each nucleic acid probe being a nucleic acid molecule comprising a target-specific region and a region comprising genedigits, said genedigits being nucleic acid regions of predetermined nucleotide sequences that are specifically bound by complementary anti-genedigit nucleotide sequences, each said anti-genedigit nucleotide sequence being attached to a label monomer or a combination of label monomers, wherein the population is in solution, and wherein the label has a detectable signal that is fluorescence, luminescence, colorimetric, voltage, current, or a magnetic field can be used for the detection, identification, and direct quantification of a wide variety of target analytes. The diverse population is advantageous in that it provides a large number of unique labels of about the same unit signal starting from just a small number of different labels. Enough labels are provided so that each analyte in a complex mixture can be uniquely bound by a label and thus identified. The labels are designed so that they can be used in a small volume of solution which increases the efficiency of the binding reaction and is useful when only small amounts of sample are available for analysis. After the individual molecules in a sample are identified, they can be directly counted resulting in a digital read-out of each molecular species in a mixture. Since the diverse labels are based on a small number of starting labels, this allows the detection method to operate in a narrow dynamic range resulting in improved sensitivity of the system since the trade-off between sensitivity and dynamic range is avoided. The methods of the invention therefore provide an accurate and sensitive system for the detection, identification and quantitation of analytes in a mixture.

[0012] In one embodiment, the invention is directed to detecting nucleic acid analytes in a complex mixture by first contacting the mixture with the diverse population of probes of the invention under conditions sufficient for hybridization of the probes to target nucleic acid analytes. These target specific nucleic acid probes, called specifiers, contain a target specific region and a region containing one or more unique "genedigit" sequences. The genedigits are nucleic acid regions of predetermined nucleotide sequences that are specifically bound by complementary anti-genedigit nucleotide sequences, each said anti-genedigit nucleotide sequence being attached to a label monomer or a combination of label monomers, wherein the label has a detectable signal that is fluorescence, luminescence, colorimetric, voltage, current, or a magnetic field.

[0013] As used herein, the term "bound" when referring to a unique label or nucleic acid is intended to mean that a label monomer is attached to a nucleotide in a 1:1 correspondence. A label monomer as used herein is intended to mean an individual measurable moiety, such as a radioisotope, fluorochrome, dye, enzyme, nanoparticle, chemiluminescent marker, biotin, or other moiety known in the art that is measurable by analytical methods. A label monomer can be directly attached to a nucleotide using methods well known in the art. Nucleotides can also be chemically modified or derivitized in order to attach a label monomer. For example, a fluorescent monomer such as a fluorescein molecule can be attached to dUTP (deoxyuridine-triphosphate) using a four-atom aminoalkynyl group. Each label monomer is attached to a nucleotide making a label monomer:nucleotide complex. This label monomer:nucleotide can be incorporated into nucleic acids in a variety of ways. For example, a label monomer:nucleotide can be incorporated at only one location within a nucleic acid or at two or more locations within a nucleic acid. A nucleotide can be attached to a label monomer first and then be incorporated into a nucleic acid, or an existing nucleic acid can be labeled by attaching a label monomer

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 169

to a nucleotide within the nucleic acid. In addition, for example, a label monomer:nucleotide can be incorporated into a nucleic acid and a different type of label monomer:nucleotide can be incorporated into the same nucleic acid.

[0014] As used herein, "analyte" or target when referring to an analyte is intended to mean any nucleic acid molecule whose presence is measured. A nucleic acid analyte molecule can be essentially any nucleic acid molecule for which a detectable probe or assay exists or can be produced by one skilled in the art. Measurement can be quantitative or qualitative. An analyte can be part of a sample that contains other components or can be the sole or major component of the sample. Therefore, an analyte 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 analyte can be attached in solution or solid-phase, including, for example, to a solid surface such as a chip, microarray or bead. Also the analyte can have either a known or unknown structure or sequence.

[0015] As used herein, the term "target specific" is intended to mean an agent that binds to a target analyte selectively. This agent will bind with preferential affinity toward the target while showing little to no detectable cross-reactivity toward other molecules. For example, when the target is a nucleic acid, a target specific sequence is one that is complementary to the sequence of the target and able to hybridize to the target sequence with little to no detectable cross-reactivity with other nucleic acid molecules. A nucleic acid target could also be bound in a target specific manner by a protein, for example by the DNA binding domain of a transcription factor.

[0016] The term "complementary" refers to two nucleotides that can form multiple thermodynamically favorable interactions with one another. For example, adenine is complementary to thymine as they can form two hydrogen bonds. Similarly, guanine and cytosine are complementary since they can form three hydrogen bonds. A nucleotide sequence is the complement of another nucleotide sequence if the nucleotides of the first sequence are complementary to the nucleotides of the second sequence. The percent of complementarity (i.e. how many nucleotides from one strand form multiple thermodynamically favorable interactions with the other strand compared with the total number of nucleotides present in the sequence) indicates the extent of complementarity of two sequences.

[0017] As used herein, the term "repeat sequence" is intended to mean two or more copies of a core element. A repeat sequence can have direct repetition of the core sequence without any intervening sequence, or the repeat sequence can have non-consecutive repetition of the core element with intervening sequences. A core element can be made of nucleic acids such as an oligonucleotide or an aptamer. If, for example, the core element is a nucleic acid sequence of 8 base pairs then three direct repeats of this sequence would be a 24 base sequence. A "complimentary repeat sequence" is a sequence that binds specifically to the repeat sequence. For the example above where the repeat sequence is repetition of a nucleic acid core element, the complimentary repeat sequence can contain one or more copies of the complementary strand of the core element that will hybridize specifically to the repeat sequence.

[0018] As used herein, the term "genedigit" is intended to mean a region of pre-determined nucleotide sequence that serves as an attachment point for a label. The genedigit can have any structure including, for example, a single unique sequence or a sequence containing repeated core elements. Each genedigit has a unique sequence which differentiates it from other genedigits. An "anti-genedigit" is a nucleotide sequence or structure that binds specifically to the gene digit. For example, if the genedigit is a nucleic acid, the anti-genedigit can be a nucleic acid sequence that is complementary to the genedigit sequence. If the genedigit is a nucleic acid that contains repeated core elements then the anti-genedigit can be a series of repeat sequences that are complementary to the repeat sequences in the genedigit. An anti-genedigit can contain the same number, or a lesser number, of repeat sequences compared to the genedigit as long as the anti-genedigit is able to specifically bind to the genedigit.

[0019] As used herein, the term "specifier" is intended to mean the linkage of one or more genedigits to a target specific sequence. The genedigits can be directly linked or can be attached using an intervening or adapting sequence. A specifier can contain a target specific sequence which will allow it to bind to a target analyate. An "anti-specifier" has a complementary sequence to all or part of the specifier such that it specifically binds to the specifier.

[0020] As used herein, the term "mixture" is intended to mean a composition that contains more than one molecule. A mixture can be homogenous, containing a single species, or heterogeneous, containing different species. Examples of homogeneous samples include, for example, isolated populations of nucleic acids. Heterogeneous mixtures include extracts from tissues, cells, lysates and fractionated portions thereof.

[0021] As used herein, the term "label" is intended to mean a molecule or molecules that render an analyte detectable by an analytical method. An appropriate label depends on the particular assay format and are well known by those skilled in the art. For example, a label specific for a nucleic acid molecule can be a complementary nucleic acid molecule attached to a label monomer or measurable moiety, such as a radioisotope, fluorochrome, dye, enzyme, nanoparticle, chemiluminescent marker, biotin, or other moiety known in the art that is measurable by analytical methods. In addition, a label can include any combination of label monomers.

[0022] As used herein, "unique" when used in reference to label is intended to mean a label that has a detectable signal that distinguishes it from other labels in the same mixture. Therefore, a unique label is a relative term since it is dependent upon the other labels that are present in the mixture and the sensitivity of the detection equipment that is used. In the case of a fluorescent label, a unique label is a label that has spectral properties that significantly differentiate

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 170

it from other fluorescent labels in the same mixture. For example, a fluorescein label can be a unique label if it is included in a mixture that contains a rhodamine label since these fluorescent labels emit light at distinct, essentially non-overlapping wavelengths. However, if another fluorescent label was added to the mixture that emitted light at the same or very similar wavelength to fluorescein, for example the Oregon Green fluorophore, then the fluorescein would no longer be a unique label since Oregon Green and fluorescein could not be distinguished from each other. A unique label is also relative to the sensitivity of the detection equipment used. For example, a FACS machine can be used to detect the emission peaks from different fluorophore-containing labels. If a particular set of labels have emission peaks that are separated by, for example, 2 nm these labels would not be unique if detected on a FACS machine that can distinguish peaks that are separated by 10nm or greater, but these labels would be unique if detected on a FACS machine that can distinguish peaks separated by 1nm or greater.

[0023] As used herein, the term "signal" is intended to mean a detectable, physical quantity or impulse by which information on the presence of an analyte can be determined. Therefore, a signal is the read-out or measurable component of detection. A signal includes, for example, fluorescence, luminescence, colorimetric, density, image, sound, voltage, current, magnetic field and mass. Therefore, the term "unit signal," as used herein is intended to mean a specified quantity of a signal in terms of which the magnitudes of other quantities of signals of the same kind can be stated. Detection equipment can count signals of the same type and display the amount of signal in terms of a common unit. For example, a nucleic acid can be radioactively labeled at one nucleotide position and another nucleic acid can be radioactively labeled at three nucleotide positions. The radioactive particles emitted by each nucleic acid can be detected and quantified, for example in a scintillation counter, and displayed as the number of counts per minute (cpm). The nucleic acid labeled at three positions will emit about three times the number of radioactive particles as the nucleic acid labeled at one position and hence about three times the number of cpms will be recorded.

[0024] As used herein, the term "dendrimer" is intended to mean a branched nucleic acid. These structures are composed of layers of nucleic acid, each layer being composed of partially single-stranded heteroduplexes which are called dendrimer monomers. The outermost layer of a given dendrimer can have multiple single-stranded arms capable of hybridization with a complementary nucleic acid sequence. Dendrimer monomers have the property that sequential addition of monomers yields a three-dimensional structure composed of nucleic acid. Various configurations of nucleic acid molecules can give rise to a large number of differently shaped dendritic structures. For example, using commercially available dendrimeric synthons, a dendrimer with 1 stem and 81 branches can be synthesized. Also fork-like, comb-like and bubbled structures are possible.

[0025] As used herein, the term "nucleic acid" is intended to mean a single- or double-stranded DNA or RNA molecule including, for example, genomic DNA, cDNA and mRNA. The term is intended to include nucleic acid molecules of both synthetic and natural origin. A nucleic acid molecule of the invention can be of linear, circular or branched configuration, and can represent either the sense or antisense strand, or both, of a native nucleic acid molecule. A nucleic acid molecule of the invention can further incorporate a detectable moiety such as a radiolabel, a fluorochrome, a ferromagnetic substance, a luminescent tag or a detectable moiety such as biotin.

[0026] As used herein, the term "hybridizing" is intended to mean joining different components together. Any number of components can be joined together, for example two components can be joined together to make a duplex, three components can be joined together to make a triplex, and so on. Nucleic acids can form a hybrid or duplex, for example, by hydrogen bonding between complementary nucleotides. The formation of nucleic acid hybrids is dependent on several conditions known in the art, including temperature, salt concentration, and pH.

[0027] As used herein, "complexity" refers to the degree of repeated elements between two nucleic acids that are being hybridized together in a solution (see Anderson, M.L.M., Nucleic Acid Hybridization, Springer-Verlag, New York (1999)). When the nucleic acid molecules that are to be hybridized contain repeated core elements or homopolymeric regions, there are many possible pairing opportunities and so the hybridization proceeds quickly. When the nucleic acid molecules that are to be hybridized do not contain any repeated core elements then there is only one way that the two sequences can be hybridized and so the hybridization proceeds more slowly. Sequences that hybridize quickly are said to have a low complexity, while sequences that take longer to hybridize have a higher complexity. For example, a 40 base pair genedigit sequence made up of five direct repeats of an 8 base pair core element, can be hybridized to a 24 base pair anti-genedigit containing three repeats of the 8 base pair core repeat in three different registers (see, for example, Figure 1C). Thus the anti-genedigit can hybridize to the 40 base pair genedigit through a 24 base pair sequence that only has a complexity of an 8 base pair sequence.

[0028] The invention provides a diverse population of uniquely labeled probes providing a large number of unique labels of about the same unit signal starting from just a small number of different labels. Enough labels are provided so that each analyte in a complex mixture can be uniquely bound by a label and thus identified. These labels can be used in very small volumes which improves kinetics in the binding reaction. In addition, the design of these labels allows for improved accuracy in detection.

[0029] A large diversity of unique labels can be desirable in order to provide a unique label to each species in a complex mixture. The invention provides methods for combining different labels in pre-determined ratios to generate a large

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 171

diversity of unique labels. The labels are designed in a modular fashion which allows for flexibility in the number of unique labels that are generated. For example, if a large number of modules are used, a large number of ratios of the different labels is possible which leads to a large number of unique labels. The number of labels generated can be adjusted to cover mixtures with different numbers of species.

[0030] The invention provides a diverse population of labels that contains thirty or more unique labels where each unique label is bound to a nucleic acid. A diverse population of labels is a mixture of distinct label species. This population can have as few as about thirty distinct label species or as high as 10^{17} distinct label species. The actual number of molecules of each label species can vary as long as at least one molecule of the label species is present. In addition, the invention provides a diverse population of labels that contains 40, 60, 80, 100, 120, 140, or about 150 unique labels. A portion of this population can be made up of different individual label monomers. The invention also provides unique labels made from combinations of different labels which can increase the number of unique labels substantially.

[0031] The labels of the invention are bound to nucleic acids. In particular, the labels are bound to nucleic acids through the attachment of a label monomer to a nucleotide within a nucleic acid in a 1:1 correspondence. A nucleic acid can contain several label monomers, however each label monomer is attached directly to a nucleotide.

[0032] A label monomer can be attached to any nucleotide including both natural and non-natural nucleotides. A nucleotide contains three parts, a phosphate group, a pentose five-carbon sugar molecule, and an organic base. In RNA, the pentose is ribose and in DNA it is deoxyribose and so nucleotides for incorporation into RNA are called ribonucletides and nucleotides for incorporation into DNA are called deoxyribonucleotides. Three bases adenine, guanine, and cytosine are found in both DNA and RNA while thymine is normally found only in DNA and uracil is normally found only in RNA. Nucleotides can have one, two or three attached phosphate groups and are sometimes referred to as nucleoside phosphates. Nucleotides can contain modified nucleosides having modified bases (e.g., 5-methyl cytosine) and modified sugar groups (e.g., 2'O-methyl ribosyl, 2'O-methoxyethyl ribosyl, 2'fluoro ribosyl, 2'amino ribosyl, and the like). An example of non-natural bases that are used in the art are isocytidine and isoguanine.

[0033] A label monomer as used herein is intended to mean an individual measurable moiety, such as a radioisotope, fluorochrome, dye, enzyme, nanoparticle, chemiluminescent marker, biotin, or other moiety known in the art that is measurable by analytical methods. A label monomer can be attached to a nucleotide using methods well known in the art and exemplified herein.

[0034] Radioisotopes are an example of label monomers that can be utilized by the invention. Several radioisotopes can be used as label monomers for labeling nucleotides including, for example, ^{32}P, ^{33}P, ^{35}S, ^{3}H, and ^{125}I. These radioisotopes have different half-lives, types of decay, and levels of energy which can be tailored to match the needs of a particular experiment. For example, ^{3}H is a low energy emitter which results in low background levels, however this low energy also results in long time periods for autoradiography. Radioactively labeled ribonucleotides and deoxyribonucleotides are commercially available. Nucleotides are available that are radioactively labeled at the first, or α, phosphate group, or the third, or γ, phosphate group. For example, both [α-^{32}P]dATP and [γ-^{32}P]dATP are commercially available. In addition, different specific activities for radioactively labeled nucleotides are also available commercially and can be tailored for different experiments.

[0035] Another example of label monomers that can be utilized by the invention are fluorophores. Several fluorophores can be used as label monomers for labeling nucleotides including, for example, fluorescein, tetramethylrhodamine, and Texas Red. Several different fluorophores are known, and more continue to be produced, that span the entire spectrum. Also different formulations of the same fluorophore have been produced for different applications. For example, fluorescein, can be used in its isothiocynanate form (FITC), as mixed isomer or single isomer forms of carboxyfluorescein succinimidyl ester (FAM), or as isomeric dichlorotriazine forms of fluorescein (DTAF). These labels are chemically distinct, but all emit light with a peak between 515-520 nm. In addition to the chemical modifications of fluorescein, completely different fluorophores have been synthesized that have the same or very similar emission peaks as fluorescein. For example, the Oregon Green dye has virtually superimposable excitation and emission spectra compared to fluorescein. Other fluorophores such as Rhodol Green and Rhodamine Green are only slightly shifted in their emission peaks and so also serve functionally as substitutes for fluorescein. In addition, different formulations or related dyes have been developed around other fluorophores that emit light in other parts of the spectrum.

[0036] Amine-reactive and thiol-reactive fluorophores are available and used for labeling nucleotides and biomolecules. Generally, nucleotides are fluorescently labeled during chemical synthesis, for example, incorporation of amines or thiols during nucleotide synthesis permit addition of fluorophores. Fluorescently labeled nucleotides are commercially available. For example, uridine and deoxyuridine triphosphates are available that are conjugated to ten different fluorophores that cover the spectrum.

[0037] Fluorescent dyes that can be bound directly to nucleotides can also be utilized as label monomers. For example, FAM, JOE, TAMRA, and ROX are amine reactive fluorescent dyes that have been attached to nucleotides and are used in automated DNA sequencing. These fluorecently labeled nucleotides, for example, ROX-ddATP, ROX-ddCTP, ROX-ddGTP and ROX-ddUTP, are commercially available.

[0038] Non-radioactive and non-fluorescent label monomers are also available. For example, biotin can be attached

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 172

directly to nucleotides and detected by specific and high affinity binding to avidin or streptavidin which has been chemically coupled to an enzyme catalyzing a colorimetric reaction (such as phosphatase, luciferase, or peroxidase). Digoxigenin labeled nucleotides can also similarly be used for non-isotopic detection of nucleic acids. Biotinylated and digoxigenin-labeled nucleotides are commercially available.

[0039] Very small particles, termed nanoparticles, also can be used as label monomers to label nucleic acids. These particles range from 1 - 1000nm in size and include diverse chemical structures such as gold and silver particles and quantum dots.

[0040] When irradiated with angled incident white light, silver or gold nanoparticles ranging from 40-120nm will scatter monochromatic light with high intensity. The wavelength of the scattered light is dependent on the size of the particle. Four to five different particles in close proximity will each scatter monochromatic light which when superimposed will give a specific, unique color. The particles are being manufactured by companies such as Genicon Sciences. Derivatized silver or gold particles can be attached to a broad array of molecular probe molecules including, proteins, antibodies, small molecules, receptor ligands, and nucleic acids. For example, the surface of the particle can be chemically derivitized to allow attachment to a nucleotide.

[0041] Another type of nanoparticle that can be used as a label monomer are quantum dots. Quantum dots are fluorescing crystals 1-5nm in diameter that are excitable by a large range of wavelengths of light. These crystals emit light, such as monochromatic light, with a wavelength dependent on their chemical composition and size. Quantum dots such as CdSe, ZnSe, InP, or InAs possess unique optical properties. These particles have been used in the semiconductor industry for several years, but are just now being applied to molecular biology applications.

[0042] Many dozens of classes of particles can be created according to the number of size classes of the quantum dot crystals. The size classes of the crystals are created either 1) by tight control of crystal formation parameters to create each desired size class of particle, or 2) by creation of batches of crystals under loosely controlled crystal formation parameters, followed by sorting according to desired size and/or emission wavelengths. Use of quantum dots for labeling particles, in the context of the present invention, is new, but is old in the art of semiconductors. Two examples of earlier references in which quantum dots are embedded within intrinsic silicon epitaxial layers of semiconductor light emitting/detecting devices are United States Patent Nos. 5,293,050 and 5,354,707 to Chapple-Sokol, et al.

[0043] Due to their very small size the quantum dots can be coupled into oligonucleotides directly without affecting the solubility or use of the oligonucleotide. The invention requires that only one oligonucleotide molecule is coupled to each nanoparticle. To synthesize an oligonucleotide-nanoparticle complex in a 1:1 ratio by conventional batch chemistry, both the oligonucleotide and the nanoparticle require a single reactive group of different kinds that can be reacted with each other. For example, if an oligonucleotide has an amino group and a nanoparticle has an aldehyde group, these groups can react to form a Schiff base. An oligonucleotide can be derivitized to attach a single amino or other functional group using chemistry well known in the art. However, when a nanoparticle is derivatized, it is covered with a chemical reagent which results in coating the entire surface of the nanoparticle with several functional groups.

[0044] In this context, a method of coupling one oligonucleotide to one nanoparticle by chemically coupling the oligonucleotide on a solid surface such as the glass support used for the oligonucleotide synthesis is provided. For example, commercially available resins for oligonucleotide synthesis such as long chain alkylamino controlled pore glass (lcaa CPG) can be used. Alternatively, a flat surface such as a derivitized microscope slide can be used. The surface density of the nascent oligonucleotide chains should be lower than the diameter of the nanoparticle. This can be achieved by either choosing a glass support with low surface density of the reactive groups, or by using diluted reagent for the first step of the oligonucleotide synthesis so that the surface is not saturated. Another point of consideration when using the standard glass matrices for oligonucleotide synthesis is to use a pore diameter higher than the nanoparticle diameter to ensure the flow of the reagents. For example, an oligonucleotide can be synthesized on a diluted basis relative to the solid support, for example one tenth of a normal synthesis, to ensure good spacing of the oligonucleotides on the glass support. After the oligonucleotide is synthesized with a reactive functional group, for example, an amino group, derivitized nanoparticles are passed over the glass support to react with the oligonucleotides. A sufficiently large pore size of the glass support can be chosen to prevent clogging with nanoparticles. For example, a pore size of about 200 nm can be used. After the reaction is complete, un-reacted groups on the nanoparticle can be blocked and the complexes can be uncoupled from the glass support.

[0045] The labels of the invention are bound to nucleic acids through nucleotides in the nucleic acid. A nucleotide can be attached to a label monomer first and then the label monomer:nucleotide can be incorporated into a nucleic acid, or an existing nucleic acid can be labeled by attaching a label monomer to a nucleotide within the nucleic acid.

[0046] A label monomer can be attached to a nucleotide using a variety of methods well known in the art and described herein. For example, the label monomer can be directly attached to the nucleotide in a 1:1 correspondence by incorporation of a radioactive phosphate into the phosphate backbone of the nucleotide. Also, for example, a general method for labeling phosphates with a fluorescent label that employs an imidazole derivative prepared from a BODIPY FL hydrazide has been reported (Wang and Giese, Anal. Chem. 65: 3518 (1993).

[0047] Depending on the labeling moiety used, it can be desirable to derivitize or chemically modify a nucleotide in

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 173

order to bind the label monomer. These methods and chemistries are known in the art. In addition, a linker can be used to attach a label monomer to a nucleotide in a 1:1 correspondence. For example, a fluorescently labeled nucleotide such as fluorescein-12-dUTP can have a fluorophore monomer attached via a four-atom aminoalkynyl group to the dUTP molecule.

[0048] These nucleotides attached to label monomers can be incorporated into a nucleic acid using several methods for labeling nucleic acids well known in the art. For example, enzymes such as DNA or RNA polymerases, Taq polymerases, terminal deoxynucleotidyl transferases, or reverse transcriptases can be used to incorporate labeled nucleotides into nucleic acids.

[0049] Labeled nucleotides can be incorporated into nucleic acids, for example, by nick translation. In this procedure DNAse I is used to create single-strand nicks in double stranded DNA and then the 5' to 3' exonuclease and 5' to 3' polymerase actions of E. coli DNA polymerase I are used to remove stretches of single stranded DNA starting at the nicks and replace them with new strands made by incorporation of labeled nucleotides. Nick translation can utilize any labeled nucleotide including radioactively labeled nucleotides and biotinylated or digoxigenin labeled nucleotides. In a similar way T4 DNA polymerase can be used to incorporate labeled nucleotides. In addition, labeled nucleotides can be incorporated into nucleic acids using the polymerase chain reaction (PCR) and Taq polymerases. The degree of labeling can be controlled by including one, or up to all four labeled nucleotides. In addition, the degree of labeling can be controlled by increasing or decreasing the concentration of the labeled nucleotide(s).

[0050] Other methods for labeling nucleic acids include generating single-stranded cDNA from RNA by using a reverse transcriptase in the presence of labeled nucleotides. In addition, DNA can be cloned into a vector with SP6 or T7 polymerase sites. Transcription in the presence of SP6 or T7 RNA polymerase and labeled nucleotides results in a labeled RNA transcript. The transcript can be labeled to different degrees by including one or more labeled nucleotides. In addition, several nucleotides within a nucleic acid can be labeled, for example, by cloning DNA into a bacteriophage M13 based vector. Then the Klenow fragment of DNA polymerase I and the M13 universal probe primer can be used to synthesize the complementary stand with incorporation of labeled nucleotides.

[0051] Several methods are described above for incorporation of labeled nucleotides into newly synthesized nucleic acids. Existing nucleic acids can also be labeled using several methods known in the art. For example, RNA or DNA can be end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. This kinase can be used to transfer the radioactive phosphate of ATP to a free 5' OH group in either DNA or RNA. The enzyme also has a phosphatase activity and so two reactions are possible. In the forward reaction, the enzyme catalyzes phosphorylation following removal of 5' terminal phosphates with alkaline phosphatase (or other phosphatase). In the exchange reaction, the kinase catalyzes the exchange of an existing 5' phosphate with the third or γ phosphate of ATP. The latter reaction is carried out in the presence of excess ATP and ADP for efficient phosphorylation. Using this method the radioactive phosphate of ATP is transfered to the end of the nucleic acid molecule.

[0052] Nucleic acids can also be labeled with terminal deoxynucleotidyl transferease which adds labeled nucleotides onto the 3' end of DNA fragments. Both single and double-stranded DNAs are substrates for this enzyme. The large (Klenow) fragment of E. coli DNA polymerase I can also be used to label the ends of nucleic acids. Since this enzyme has a 5' to 3' polymerase activity it can be use to "fill in" the 3' ends of DNA fragments opposite of 5' extensions or overhangs with labeled nucleotides. End-labeling of nucleic acids using polynucleotide kinase or terminal deoxynucleotidyl transferease results in the incorporation of one label per nucleic acid. The "fill in" reaction can be used to label the nucleic acid at one nucleotide per nucleic acid or at more than one nucleotide per nucleic acid.

[0053] In addition, nucleic acids can be labeled by modification of nucleotides within the nucleic acid. For example, cytidine residues in DNA and RNA can be modified by reaction with sodium bisulfite to form sulfonate intermediates that can then be directly coupled to hydrazides or aliphatic amines. Virtually any of the fluorescent, biotin or other hydrazides or aliphatic amines can be used in this reaction. The bisulfite-activated cytidylic acid can also be coupled to aliphatic diamines such as ethylenediamine. The amine-modified DNA or RNA can then be modified with any of the amine-reactive dyes. In addition, phosphate groups can be targeted in nucleic acids for labeling. Although phosphate groups of nucleotides are not very reactive in aqueous solution, their terminal phosphate groups can react with carbodiimides and similar reagents in combination with nucleophiles to yield labeled phsophodiesters, phosphoramidates and phosphorothioates. For example, DNA can be reacted quantitatively with carbonyl diimidazole and a diamine such as ethylenediamine to yield a phosphoramidate that has a free primary amine and that this amine can then be modified with amino-reactive reagents. Fluorescent or biotinylated amines have been coupled to the 5' phosphate of tRNA using dithiodipyridine and triphenylphosphine.

[0054] The bond between labels and nucleic acids can be covalent bonds or non-covalent bonds that are stable to hybridization and washing conditions. The labels can be bound to a nucleic acid in a sequence specific manner, for example by the incorporation of a labeled nucleotide into DNA that has been digested by a restriction enzyme. Alternatively the labels can be bound to a nucleic acid in a non-sequence specific manner, for example by the incorporation of a label onto the terminal phosphate of a nucleic acid using [γ-^{32}P]ATP and T4 polynucleotide kinase.

[0055] Several types of nucleic acids can be utilized with this invention including a single- or double-stranded DNA or

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 174

RNA molecule which can include, for example, genomic DNA, cDNA and mRNA. Nucleic acid molecules of both synthetic and natural origin can be used. In connection with the present invention, a nucleic acid molecule can be of linear, circular or branched configuration, and can represent either the sense or antisense strand, or both, of a native nucleic acid molecule. Nucleotide bases not found in nature such as isocytidine and isoguanine can be incorporated into the nucleic acid.

[0056] Labels can bind to nucleic acids in a number of diverse ways. For example, a particular label monomer can be bound to a nucleic acid at only one position in the nucleic acid or at many positions in the nucleic acid. In addition, a particular label monomer can be bound to a nucleic acid and one or more other label monomers can also bound to the same nucleic acid. In this case the label can contain a mixture of two or more different labels. Furthermore, nucleic acids labeled with any or all of these combinations can be bound to another nucleic acid through hybridization.

[0057] Additional diversity is introduced when the nucleic acid is branched. One example of a branched nucleic acid is a dendrimer. Dendrimers are composed of layers of nucleic acid, each layer being composed of partially single-stranded heteroduplexes yielding a three-dimensional structure composed of nucleic acid. Various configurations of nucleic acid molecules can give rise to a large number of differently shaped dendritic structures including, for example, a dendrimer with 1 stem and 81 branches, or dendrimers with fork-like, comb-like or bubbled structures. The outermost layer of a given dendrimer can have multiple single-stranded arms capable of hybridization with a complementary nucleic acid sequence. Due to the relatively large size of nucleic acid molecules, nucleic acid dendrimers can contain numerous labels with limited steric hindrance. Use of dendrimers can multiply the signal generated by a labeled nucleic acid by a pre-determined factor equal to the number of branches.

[0058] Several unique combinations of labels can be formed using branched nucleic acids. For example, by using different chemical protective groups, one label monomer can bind to one branch while one or more other label monomers bind to other branches. Furthermore, labeled nucleic acids can be attached in various combinations to the branches of a dendrimer through hybridization.

[0059] The invention provides a diverse population of 30, or about 40, 60, 80, 100, 120, 140, or 150 unique labels bound to a nucleic acid. Part of this population can be made up of different individual label monomers. The invention also provides unique labels made from a combination of two or more different labels. This can increase the number of unique labels substantially.

[0060] A unique label is a label that generates a signal that is distinguishable from other labels in the same mixture. Therefore, designation as a unique label is dependent upon the sensitivity of the detection equipment that is used. For example, where fluorescent or nanoparticle labels are used a CCD camera can be used to detect the labels. The sensitivity of this equipment depends on the manufacturer, model, and design of the equipment. In addition, several parameters can be set by the user in order to achieve maximum sensitivity. For example, the use of different filter sets can increase the sensitivity of detection for certain experiments.

[0061] The ability to distinguish different labels also depends on the particular properties of the label. For example, some fluorophores emit light within a broad peak or range of wavelengths while other fluorophores emit light within a narrow peak. Fluorophores that emit light within a broad peak can obscure neighboring peaks. In addition, the shapes of the emission peaks from different fluorophores varies. For example, some fluorophores have a peak that has a sharp ascent, but a broad trailing end that can obscure neighboring peaks. If two peaks cannot be distinguished then the two labels associated with those peaks cannot be considered unique. Where fluorescent labels are used, the labels would emit fluorescent light at distinct, essentially non-overlapping wavelengths distant from each other by at least 1nm, 5nm, 10nm, 15nm, 20nm, 25nm, and preferably 30nm, 35nm, 40nm, 45nm, and more preferably by at least 50nm. For example, the emission peak of dye #1 could be 585nm, and the peak emission of dye #2 could be 630nm.

[0062] A unique label is a label that generates a signal that is distinguishable from other labels in the same mixture. A unique label, therefore, is dependent on the other labels that are included in the mixture. For example, fluorescein which has an emission spectra peak at 518nm and rhodamine red with an emission spectra peak of 590nm have clearly distinct emission peaks and hence both are considered to be unique labels when included together in a mixture. However, if another fluorophore is added that has an emission peak that overlaps with fluorescein or rhodamine red, these labels would not be unique. For example, if Oregon green (emission spectra peak 522nm) is added to the above mixture the emission peak of fluorescein and Oregon Green can overlap so much that, depending on the detection equipment used, they cannot be clearly distinguished from each other and hence neither fluorescein nor Oregon Green can act as unique labels in that particular mixture.

[0063] Often several formulations of the same label or related labels are synthesized for use in different applications, and while these labels have different chemical properties, they are not distinct in terms of detection. For example, fluorescein, can be used in its isothiocynanate form (FITC), as mixed isomer or single isomer forms of carboxyfluorescein succinimidyl ester (FAM), or isomeric dichlorotriazine forms of fluorescein (DTAF). These labels are chemically distinct, but all emit light with a peak between 515-520 nm and hence would overlap so as to appear identical on most currently available detection equipment. In addition to the chemical modifications of fluorescein, completely different fluorophores have been synthesized that have the same or very similar emission peaks as fluorescein. For example, the Oregon

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 175

Green dye has virtually superimposable excitation and emission spectra compared to fluorescein. Other fluorophores such as Rhodol Green and Rhodamine Green are only slightly shifted in their emission peaks and so also serve functionally as substitutes for fluorescein.

[0064] A limited number of different label monomers are known that can be used together in a mixture and still provide unique signals. For example, five analytes can be distinctly labeled using the BODIPY fluorophore set from Molecular Probes (Eugene, OR). These fluorophores have the following distinct emission peaks: BODIPY FL(513), BODIPY R6G (550), BODIPY TMR (574), BODIPY 581/591 (592) and BODIPY TR (617). In order to obtain clear results using currently available detection equipment, the number of different fluorophores that can be used is less than thirty.

[0065] Unique labels also can be generated by combining two or more different label monomers to make a new label. The signal from the resulting label must be distinguishable from the signals of other labels used in the same experiment in order to be a unique label. For example, a nucleic acid labeled with both fluorescein and rhodamine will emit light at a different wavelength than a nucleic acid labeled with either fluorescein or rhodamine alone.

[0066] In this invention, various ratios of different label monomers bound to nucleic acids can be combined to generate a diverse population of unique labels that can include up to 10^{17} or more unique labels. For example, a nucleic acid labeled with two fluorescein labeled nucleotides and three rhodamine labeled nucleotides will emit light at a different wavelength compared to a nucleic acid labeled with three fluorescein nucleotides and two rhodamine nucleotides. In another example, a nucleic acid could be labeled with different ratios of three or more label monomer:nucleotides which greatly increases the variety of unique labels that can be generated.

[0067] The signal generated by each of the label monomers bound to the nucleotides can be normalized to have about the same unit signal. For example, if fluorescent monomer A is known to emit a different quanta of light as fluorescent monomer B, the signal from unique labels that contain a pre-determined number of these monomers can be normalized based on the known signal properties of the label monomers and the number of each monomer present in the unique label. Different numbers of labels can be used in the invention and so different multiples of the same unit signal can be utilized by the invention. For example, a nucleic acid can be labeled with two fluorophores and another nucleic acid can be labeled with six fluorophores. The second nucleic acid will have three times the signal of the first nucleic acid. Since the number of label monomers attached to each nucleic acid is known, the signals from the labeled nucleic acids can be normalized based on the number of label monomers present. For example, the signal from the nucleic acid with six fluorphores can be divided by three which normalized the signal relative to the signal from the nucleic acid with two fluorophores.

[0068] This method of producing labels results in important advantages over existing techniques such as microarray formats. Since each analyte is uniquely identified by a label with about the same unit signal, it allows the labels to be directly counted resulting in a digital read-out of each molecular species in a mixture. In contrast, microarray data must undergo several intermediate transformations to quantitate the number of molecules which results in a less precise analog output. In addition, using a diversity of labels with the same unit signal allows the detection method to operate in a narrow dynamic range resulting in improved sensitivity of the system since the trade-off between sensitivity and dynamic range is avoided.

[0069] The invention provides a diverse population of uniquely labeled probes. This population of probes contains about 30 or more target specific nucleic acid probes each attached to a unique label bound to a nucleic acid. In addition, the invention provides a diverse population of uniquely labeled probes containing a diversity of 50, 100, 200, 500, 1,000, 2,000, 5,000, $1x10^4$, $3x10^4$ and about $1x10^5$ or more different labels. As described above, these unique labels can contain a mixture of two or more different labels and comprise about the same unit signal or multiple thereof.

[0070] The invention provides a diverse population of uniquely labeled probes in which a target specific nucleic acid contains a nucleic acid bound to a unique label. In addition, the invention provides a diverse population of uniquely labeled probes containing two attached populations of nucleic acids, one population of nucleic acids containing thirty or more target specific nucleic acid probes, and a second population of nucleic acids containing a nucleic acid bound by a unique label.

[0071] A target specific probe is intended to mean an agent that binds to the target analyte selectively. This agent will bind with preferential affinity toward the target while showing little to no detectable cross-reactivity toward other molecules.

[0072] The target analyte is a nucleic acid. For example, a target can be a nucleic acid that is recognized and bound specifically by a complementary nucleic acid including for example, an oligonucleotide or a PCR product, or a non-natural nucleic acid such as a locked nucleic acid (LNA) or a peptide nucleic acid (PNA).

[0073] Aptamers are nucleic acid sequences that have three dimensional structures capable of binding small molecular targets including metal ions, organic dyes, drugs, amino acids, co-factors, aminoglycosides, antibiotics, nucleotide base analogs, nucleotides and peptides (Jayasena, S.D., Clinical Chemistry 45:9, 1628-1650, (1999)) incorporated herein by reference. Nucleic acids can be directly attached to the aptamer or attached through the use of hybridization.

[0074] The population of labels of the invention is produced by a method consisting of synthesizing a population of nucleic acids bound to a predetermined ratio of at least two different labels. The method involves incorporating labeled nucleotides into a repeated nucleic acid structure using a DNA polymerase. The repeated nucleic acid structures can

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 176

be designed to allow incorporation of a pre-determined ratio of labels. Using this method several unique labels can be generated from a small number of starting labels.

[0075] A specific example of this method where ten unique labels are made from two different labels is provided in Example 1. Briefly, ten unique templates of a 220 base pair single-stranded DNA are synthesized. The templates consist of a pre-determined ratio of the following 20 base pair repeats where n+m=11: 5' (ACTCTCTCTCTCTCTCTCTC)n (GCTCTCTCTCTCTCTCTCTC)m 3' (SEQ ID NO:1). The second strand is synthesized using the primer GAGAGAGAGA (SEQ ID NO:2), Klenow polymerase, DNA ligase, dGTP, dATP, and dCTP and dUTP each labeled with a different fluorophore. The labeled nucleotides will be incorporated into the DNA in a unique ratio determined by the ratio of the two repeats. In this example, the end result is ten uniquely labeled nucleic acids where the set ratio of the two fluorophores is 1:10, 2:9, 3:8, 4:7, 5:6, 6:5, 7:4, 8:3, 9:2, and 10:1.

[0076] In Example 1, two different labels resulted in ten unique labels. Using the same protocol, three different labels would result in 30 unique labels, four different labels would result in 60 unique labels, five different labels would result in 100 unique labels, and so on. Several variations of the method will be apparent to one skilled in the art. For example, the number of repeats can be changed to be less than or greater than ten. An increased number of repeats will increase the number of unique ratios possible. This will result in an increase in the number of unique labels that can be generated from the same number of starting different labels. Also in Example 1, the ratio between the two fluorophores can include 0:11 and 11:0 which results in two additional labels that contain one fluorophore or the other.

[0077] One skilled in the art will recognize that the sequence of the templates can differ from that shown above. For example, the repeat sequence in the template can be $(GA)_n$ instead of $(CT)_n$. In addition, the repeat sequence could be a single nucleotide homopolymer such as $(A)_n$. With a homopolymer template, three labeled nucleotides can be incorporated in different ratios thus increasing the number of unique labels that can be generated. It is possible using the claimed method to generate a large number of unique labels including 40, 60, 80, 100, 120, 140, 150, 200, 500, 2,000, 5,000, $1x10^4$, $3x10^4$, $1x10^5$ or more labels. Again as described above, these unique labels can comprise about the same unit signal or multiple thereof.

[0078] Another possible modifications of the method is to change the length of the repeat, for example, to less than or greater than 20 base pairs. The repeat serves to separate the labeled nucleotides and therefore to decrease possible interference between the labels. In the case of a fluorescent label, this can decrease quenching between fluorophores. In addition, the protocol for incorporating the labeled nucleotides into the DNA can be modified as would be clear to one skilled in the art and as described herein.

[0079] In one embodiment, the labeled DNA described above can be attached to a dendrimer. Oligonucleotide tags can be synthesized at the branches of the dendrimer to allow binding of the labeled DNA. For example, a linker can be attached to the labeled DNA described above that corresponds to an oligonucleotide tag on the branches of the dendrimer. Several types of linkers are known to one skilled in the art. For example, a restriction enzyme linker can be attached to the labeled DNA. These linkers are double-stranded oligonucleotides that contain the recognition sequence of a particular restriction enzyme. These linkers can be ligated onto double-stranded DNA using a DNA ligase and digested using the appropriate restriction enzyme. The result is an overhanging single stranded sequence that is available to hybridize to another nucleic acid.

[0080] The labeled DNA described above can be directly attached to a target specific probe. In addition, the labeled DNA can be indirectly attached to a target specific probe, for example, through use of a bridging nucleic acid. One or more of these labels can be attached to each target specific probe. Binding of a uniquely labeled target specific probe to a target analyte results in the unique tagging of that analyte. This tagging allows identification of the target analyte from a mixture of analytes.

[0081] The number of unique labels can be further increased by combining the unique labels described above in different combinations. In connection with the invention, a label to a nucleic acid probe may be attached by a method, the method comprising hybridizing a nucleic acid probe having a genedigit to an anti-genedigit having a label. The anti-genedigits that contain the labels described above are hybridized to genedigits. Genedigits can be linked together in unique combinations creating an even larger number of unique labels. The modular design of the genedigits allows for flexibility in the number of unique labels that are generated. For example, if a large number of genedigit modules are used, a large number of unique templates will be available for attachment of unique labels. The number of templates generated can be adjusted to cover the number of species in the mixture.

[0082] A genedigit can be a region of pre-determined nucleotide or amino acid sequence that serves as an attachment point for a label. The genedigit can have any sequence including, for example, a single unique sequence or a sequence containing repeated core elements. However, each genedigit has a unique sequence which differentiates it from other genedigits. When added to a complex mixture of targets, a nucleic acid genedigit can contain non-natural bases such as isocytidine and isoguanine which can reduce hybridization to naturally occurring target sequences. The sequence, length of a core element, and number of repeated core elements can be varied according to the particular requirements of an experiment and will be clear to one skilled in the art. For example, a nucleic acid core element can be between about 5 and 12 base pairs in length and the core element can be repeated once to about ten times.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 177

[0083] Different genedigits can be synthesized that have a core unit that differs from the core units of other genedigits, for example, by at least two bases. By combining these unique genedigits in various combinations, a highly diverse number of structures can be synthesized. For example, 50 genedigits with unique sequences can be synthesized and split into five groups containing ten genedigits in each group. The genedigits of each group can be synthesized to have a short tag on each end. One genedigit from each group is then linked together using an adapter oligonucleotide that is complimentary to the tags that are common for each group. In this example, $1x10^5$ (10x10x10x10x10) unique combinations are possible.

[0084] Genedigits serve as attachment points for the unique labels described above. Since genedigits can be linked together in unique combinations this greatly increases the number of unique labels. By linking together genedigits, a large number of unique labels can be generated including 200, 500, 2,000, 5,000, $1x10^4$, $1x10^4$, $1x10^5$ or more unique labels.

[0085] In connection with the invention, a label may be attached to a nucleic acid probe by a method, the method comprising hybridizing a nucleic acid probe having a genedigit to an anti-genedigit having a label where the genedigit comprises a set of three or more repeat sequences, and the anti-genedigit comprises a cognate set of at least two complimentary repeat sequences where the anti-genedigit specifically hybridizes to the genedigit through a sequence having a complexity less than the number of hybridized base pairs.

[0086] An example of a nucleic acid genedigit is given in Example 2. Briefly, in this example an eight base pair core element containing the non-natural bases isocytidine and isoguanine is repeated five times. This results in a 40 base pair genedigit. In this example the anti-genedigit consists of a sequence that is complimentary to three of the five core elements in the genedigit. An anti-genedigit can contain the same number, or a lesser number, of repeat sequences compared to the genedigit as long as the anti-genedigit is able to specifically bind to the genedigit. In this example, the anti-genedigit is a 24 base pair sequence that can bind to the 40 base pair genedigit in three different registers (see Figure 1C).

[0087] The anti-genedigit in this specific example is a 24 base pair sequence that only has the complexity of an eight base pair sequence. Complexity refers to the degree of repeated elements between two nucleic acids that are being hybridized together in a solution. When the nucleic acid molecules that are to be hybridized contain repeated core elements or homopolymeric regions, there are many possible pairing opportunities and so the hybridization proceeds quickly. When the nucleic acid molecules that are to be hybridized do not contain any repeated core elements then there is only one way that the two sequences can be hybridized and so the hybridization proceeds more slowly. Sequences that hybridize quickly are said to have a low complexity, while sequences that take longer to hybridize have a higher complexity. In this example, a 40 base pair genedigit sequence made up of five direct repeats of an 8 base pair core element, can be hybridized to a 24 base pair anti-genedigit containing three repeats of the 8 base pair core repeat in three different registers. Thus the anti-genedigit can hybridize to the 40 base pair genedigit through a 24 base pair sequence that only has a complexity of an 8 base pair sequence. The advantage of this method is that hybridization will proceed more quickly and efficiently.

[0088] In the specific example above, an 8 base pair core element was describe; however, a core element can be more or less than 8 base pairs. For example, a core element can be between 5 and 12 base pairs. A change in the core element of a repeat will correspondingly change complexity. For example, if a core element is between 5 and 12 base pairs, complexity will be between 5 and 12. In addition, an anti-genedigit can bind to a genedigit using all or part of its sequence. In the example above, a 24 base pair anti-genedigit can bind to the genedigit with less than 24 bases, for example, 15 to 23 base pairs.

[0089] The anti-genedigit serves as a connector between the genedigit and the label. The genedigit is bound by the anti-genedigit and the anti-genedigit is bound to a label directly or indirectly by being bound to a dendrimer that has a label(s) attached. The anti-genedigit also contains a linker sequence that allows attachment to a dendrimer. For example, an oligonucleotide tag can be synthesized at the stem of a dendrimer that is complementary to the linker sequence on the anti-genedigit. As described above, several types of linkers can be used. In this way the repeat sequences of the anti-genedigit are free to hybridize with their corresponding genedigit resulting in the specific attachment of a label to a genedigit.

[0090] Genedigits and anti-genedigits can be comprised of nucleic acids, including aptamers.

[0091] In connection with the invention, a target specific probe can be attached to one or more genedigits to form a "specifier." The genedigits can be directly linked or can be attached to the target specific region using an intervening or adapting sequence. As described above, the target specific area can be a nucleic acid, including an aptamer. The target specific area is designed to specifically bind an analyte in a mixture. In this way an analyte can be labeled with a unique label.

[0092] A specifier can contain one to several genedigits. For example, a specifier can contain four or five or more gendigits. The number of genedigits in a specifier will determine the number of unique labels that are available to bind to analytes in a mixture. Therefore, in order to label each analyte in a complex mixture, a large population of specifiers can be synthesized that contain several combinations of genedigits. Alternatively, to label one or a small number of

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 178

analytes in mixture, a specifier or specifiers can be synthesized that contain one or a few genedigits. In addition, a specifier can contain a common tag such as a biotin tag. These tags can facilitate synthesis and purification of specifiers.

[0093] As outlined herein, the invention relates to a population of uniquely labeled nucleic acid probes. The method for providing said population consists of synthesizing a population of target specific nucleic acid probes each having a different specifier; then synthesizing a corresponding population of anti-genedigits each having a unique label, and finally hybridizing the populations of target nucleic acid probes to the anti-genedigits, to produce a population where each of the target specific probes is uniquely labeled.

[0094] The invention further provides a method for detecting a nucleic acid analyte, by contacting a mixture of nucleic acid analytes with a population of target specific probes each attached to a unique label under conditions sufficient for hybridization of the probes to the target and measuring the resulting signal from one or more of the target specific probes hybridized to an analyte where the signal uniquely identifies the analyte.

[0095] The nucleic acid analyte can contain any type of nucleic acid, including for example, an RNA population or a population of cDNA copies. The invention provides for at least one target specific probe for each analyte in a mixture. The invention also provides for a target specific probe that contains a nucleic acid bound to a unique label. Furthermore, the invention provides two attached populations of nucleic acids, one population of nucleic acids containing a plurality of target specific nucleic acid probes, and a second population of nucleic acids containing a nucleic acid bound by a unique label. When the target specific probes are attached to unique labels, this allows for the unique identification of the target analytes.

[0096] The methods of the invention are advantageous because hybridizations can be performed in solution in a small volume (0.01 -2.0 μl) which ensures high concentration of the nucleic acids which will drive the hybridization rate. Two different types of hybridizations are utilized by the methods of invention. The first hybridization is between a complex mixture of analytes and the specifiers and the second type of hybridization is between the specifiers and the labels.

[0097] In the first type of hybridization between a complex mixture of analytes and a population of specifiers, the specifers are in excess. For example, the specifiers can be in 10 to about 100 fold excess over the analytes in the complex mixture. The kinetics of this reaction can be described by the following equation:

$$t_{1/2} = \frac{N \ln 2}{3.5 \times 10^5 \times L^{0.5} \times C_0}$$

where N is the complexity of the probe (the specifiers), L is the length, C_0 is the concentration of the probe, and $t_{1/2}$ is the time required for 50% completion of the reaction. Using this equation it is clear that higher concentrations of probe will result in shorter time required for 50% completion of the reaction. The sample in this hybridization can be either RNA or DNA. If the sample is not poly A RNA then it must first be tagged by some method, for example, platinum-digoxygenin. After the hybridization is complete, the duplexes containing the analyte and specifier are separated using the poly A, or the digoxygenin tag as a handle and washed from the excess specifier. The washed material is then used for the second hybridization.

[0098] In the second hybridization the isolated specifiers are mixed with labels. The labels are in excess over the specifiers and the kinetics follow the same kinetics as shown above for the first hybridization, and it is performed at an even faster rate because the complexity, N, is very low. The hybridized complexes containing specifiers and labels can be isolated using a biotin tag on the specifiers as a handle and processed further for detection.

[0099] The formation of macromolecule hybrids is dependent on several conditions known in the art, including temperature, salt concentration, and pH. Different conditions for the formation of nucleic acid hybrids are well known in the art and can be found in, for example, Hames and Higgins, Nucleic Acid Hybridisation: A Practical Approach, IRL Press, Oxford (1991). In addition, conditions for nucleic acid-protein and protein-protein interactions are well known in the art and can be found in, for example, Current Protocols in Molecular Biology (ed. Ausubel et al., Greene Publ., New York 1989) which is incorporated herein by reference.

[0100] The "specifier-label" complexes can be separated from one another, for example, by spreading them on a two-dimensional surface such as glass, or by splitting them in liquid droplets in a flow cytometer. In this example, visualization can be achieved either by scanning the 2-D surface or by flow cytometry. Co-localization of specific labels will determine the identity of the particular specifier, which will determine the identity of the particular analyte that the specifier bound initially.

[0101] The "specifier-label" complexes can be detected by various devices including but not limited to visual inspection, digital cameras, video cameras, photographic film, or the use of current instruments such as laser scanning devices, fluorometers, luminometers, photodiodes, quantum counters, plate readers, epifluorescence microscopes, scanning microscopes, confocal microscopes, or by other means for amplifying the signal such as a photomultiplier tube or other light detector capable of detecting the presence, location, intensity, excitation and emission spectra, fluorescence po-

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 179

larization, fluorescence lifetime and other physical properties of the fluorescent signal. Non-fluorescent signals can be detected using a Geiger counter, scintillation counters, chemiluminescence, enzyme assays, and autoradiography.

[0102] An example of a detection method that can be utilized by the invention is a CCD (charge-coupled device) camera. Use of this device requires that the specifier-label complexes are first spread on a glass slide. Since all complexes will have approximately the same total intensity this simplifies detection, because the camera can be adjusted for maximum sensitivity (maximum gain) and minimum dynamic range.

[0103] For example, if fluorescent dendrimers are used for labeling the sample can be excited at the corresponding absorption maxima and scanned at the corresponding emission maxima for each of the, for example, 8-10 fluorofors used. A benefit of CCD cameras is that they have very wide range of detection which allows for a choice of fluorofors with distant emission peaks. Thinned layer CCD cameras can detect from the soft X-ray to the near infrared spectrum.

[0104] Alternatively if nanoparticles labels are used the sample can be illuminated by angled white light and detection occurs at a few wavelengths. The number of wavelengths will depend on the quality of the camera used and its sensitivity and linearity. Cameras of good quality can reproducible distinguish millions of colors with only the three standard filters (red, green, and blue).

[0105] In the case of unique labels that contain more than one label monomer, the overall signal generated by the unique label can be determined. For example, an overall signal for a unique label containing several fluorophores can be determined using a spectrometer. In addition to detecting an overall signal generated from the unique label, the invention provides for the determination of the spatial location of starting labels (which can be label monomers or combinations of label monomers) within a unique label. For example, the labels within a unique label can be separated from one another using a flow-stretch technique (Henegariu et al., Biotechniques 31:246-250 (2001)) a receding meniscus technique (Yokota et al., Nuc. Acids Res. 25:1064-1070 (1997)) or an electrostretching technique (Matsuura et al., Nuc. Acids Res. 29:E79 (2001)).

[0106] The use of flow-stretching, receding meniscus, or electro-stretching techniques allows for the separation of the labels within a unique label so that one can determine spatially where a particular label is positioned in the unique label (Example V, VI, VII and VIII). Therefore, unique labels that have the same combination of starting labels and the same overall signal can be differentiated from one another based on the location of those labels within the unique label. This ability to locate the position of a label within a unique label allows for the position of the starting label to be used as a distinguishing characteristic when generating a set of unique labels. Hence, a complex set of unique labels can be generated using the same combination of starting labels by varying the position of the labels within a unique label.

[0107] An observed signal can be modified using methods known in the art. For example, an observed signal can include subtraction of non-specific noise. An observed signal can also include, for example, treatment of the measured quantity by routine data analysis and statistical procedures which allow meaningful comparison and analysis of the observed values. Such procedures include, for example, normalization for direct comparison of values having different scales, and filtering for removal of aberrant or artifactual values.

[0108] In the first type of hybridization described above between a complex mixture of analytes and a population of specifiers, the specifers were in excess. In contrast, hybridization on standard microarrays occurs under conditions where the analyte is in excess compared to the labeled probe. In addition, on microarrays only a small fraction of this probe will actually hybridize to the cognate target during the course of the hybridization, and so target coverage at the end of the experiment is usually less than 5%. The target coverage using the methods of the invention are theoretically 100% because the label is in excess compared to the target analyte.

[0109] One of the advantages of 100% target coverage is that this allows the target analytes to be directly counted since each analyte is bound by a label. Using the methods of the invention, molecular species can be directly counted one by one. Direct counting, or a digital output, is preferable to the indirect quantitation methods used for microarrays because the data does not need to go through several intermediary transformations. If the detection equipment directly counts the number of particles emitted it is said to have a digital output, however if the direct counts undergo several intermediary transformations then the data has an analog output. Quantitation data from microarrays is known to be subject to several distortions due to data extrapolation.

[0110] Another consequence of low target coverage in microarrays is that highly sensitive equipment is needed to detect the low amount of signal. However, since microarrays require a high dynamic range for signal detection, sensitivity is decreased because of the trade-off between sensitivity and dynamic range. The methods of the invention use a limited number of labels to create a large number of unique label combinations. This allow the detection method to operate in a narrow dynamic range. In contrast, microarray methods require a large dynamic range (4 orders of magnitude or more) to account for the large differences in abundance of the different molecular species. The low requirements in terms of dynamic range required by the methods of the invention will improve the sensitivity of the system since the trade-off between sensitivity and dynamic range is avoided.

[0111] The invention also provides a method of detecting a nucleic acid analyte as above where the anti-genedigit(s) each have a unique label.

[0112] The methods of the invention provide for detection of analytes in mixtures. The mixture can contain several

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 180

types of analytes or the mixture can contain just one type of analyte. In addition, the mixture could contain just a single copy of an analyte. If the target analyte has an unknown sequence or structure, a large population of target specific specifiers can be added to the mixture. This population can include specifiers with target specific regions of pre-determined sequence or structure or specifiers can be used with target specific regions of random sequence or structure. Alternatively, if the target analyte has a known sequence or structure, a particular specifier containing a region that will specifically bind to that sequence or structure can be used either alone or in combination with other specifiers.

[0113] The methods of the invention are suited to nucleic acid analytes. A population of specifiers can be generated for any analyte where a target specific region can be found that specifically interacts with that analyte. In addition, the ability of nucleic acid aptamers to bind a wide variety of analytes allows these structures to be used in the target specific regions of specifiers. Several combinations are possible as long as the specifiers specifically bind to the analytes.

[0114] The labels and methods described herein can be used for diagnostic and therapeutic purposes. Analytes, or combinations of analytes, that are diagnostic of a disease can be detected and quantified from a sample derived from a subject. Many different analytes can be analyzed at one time from a single sample using the methods of the invention. This allows, for example, for several diagnostic tests to be performed on one sample. In addition, the methods of the invention can provide information that determines a course of treatment for a patient. For example, the amount of a particular marker for a tumor can be accurately quantified from even a small sample from a patient. For some disease like breast cancer overexpression of certain genes, such as Her2-neu, indicate a more aggressive course of treatment will be needed.

EXAMPLE I

Generation of Unique Labels Using Two Different Labels

[0115] In this example, ten unique labels are made from two different fluorescent labels. First, ten unique templates of a 220-base pair single-stranded DNA are synthesized. The templates consist of a pre-determined ratio of the following 20-base pair repeats:

5' (ACTCTCTCTCTCTCTCTCTC)n(GCTCTCTCTCTCTCTCTCTC)m 3' (SEQ ID NO:1) where n=1,2,3,4,5,6,7,8,9,10, m=1,2,3,4,5,6,7,8,9,10, and n+m=11.

[0116] The second strand is synthesized using the primer GAGAGAGAGA (SEQ ID NO:2), Klenow polymerase, DNA ligase, dGTP, dATP, dUTP-fluorescein and dCTP-rhodamine. After the reaction is complete the product is treated with S1 nuclease to digest the DNA with gaps, and the remaining full length DNA is then purified. The labeled nucleotides will be incorporated into the DNA in a unique ratio determined by the ratio of the two repeats. The end result is ten uniquely labeled nucleic acids where the set ratio of fluorescein to rhodamine is 1:10, 2:9, 3:8, 4:7, 5:6, 6:5, 7:4, 8:3, 9:2, and 10:1.

[0117] A linker oligonucleotide is ligated to the labeled DNA and then this linker is used to attach the labeled DNA to the branches of a dendrimer. The dendrimer has an oligonucleotide tag of 5 bases at the stem to facilitate binding of an anti-genedigit (see Example II) and tags of 10-base pairs at the branches to facilitate binding of labeled DNA.

EXAMPLE II

Generation of a Labeled Specifier

[0118] The specifiers are synthesized by ligating together one target specific sequence (synthetic oligonucleotide, peptide-nucleic acid (PNA), PCR product, or linked-nucleic acid (LNA)), and several "genedigits" (see Figure 1A). In this example, each specifier contains a unique combination of 4 different genedigits. This results in the generation of 10,000 possible unique specifiers.

[0119] The genedigits are synthetic oligonucleotides that contain only two of the natural bases, plus two bases that not found in nature: isocytidine and isoguanine. Such base composition ensures that the genedigits will not non-specifically hybridize with analytes in a complex mixture. The sequence of each genedigit is composed of 5 repeats of an 8-base pair core sequence (see Figure 1B). Each core sequence unit differs from the others by at least two bases.

[0120] In order to make 10,000 unique specifiers, forty different genedigits are synthesized and split into 4 groups containing 10 genedigits each. The genedigits of each group have a 5-base pair tag on each end. One genedigit from each group is present in each specifier. The genedigits are ligated with the help of adapter 10mer oligonucleotides that are complimentary to the 5-base pair tags that are common for each group. In this way, for a specifier with 4 genedigits, there will be 10x10x10x10 = 10,000 possible combinations. All specifiers also contain a biotin tag.

[0121] The genedigit serves as an attachment point for a label and so the number of labels synthesized corresponds

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 181

to the number of genedigits. The genedigits are labeled through the use of anti-genedigits. An anti-genedigit sequence consisting of three 8-base pair repeats complimentary to the 8-base pair core repeat of the corresponding genedigit is ligated to the stem of a labeled dendrimer (from Example I).

[0122] The 24-base pair labeled anti-genedigit hybridizes to the 40-base pair genedigit sequence in the specifier in one of three different registers (see Figure 1C). Thus when the label hybridized to the specifier it will do so through a 24-base pair sequence that only has a complexity of an 8-base pair sequence.

EXAMPLE III

Gene Expression Analysis Using Specifiers

[0123] In order to determine differences in gene expression between astrocytes and LPS-activated astrocytes, RNA is isolated from both populations of astrocytes using cell lysis in guanidine isothiocynine or phenol/chloroform. A population of specifiers is added to each RNA sample under conditions suitable for hybridization. The mRNA-specifier complexes are isolated with oligo dT beads and washed extensively to remove excess specifiers. The specifiers are eluted from the mRNA by digesting the mRNA with RNAse A. The specifiers are then are processed for labeling as described in Examples I and II and these labels are detected using a CCD camera. The number of specifiers corresponding to specific mRNAs from untreated astrocytes is then compared to the specifier pattern from LPS-treated astrocytes. Since the sequence of the target specific region of the specifier is known, this identifies the genes that are differentially expressed between the two samples.

EXAMPLE IV

Microbial Detection Using Specifiers

[0124] The invention can be used to detect strains of microorganisms with known sequence in biological samples. Total DNA is extracted from a blood sample from a patient with a suspected microbial infection. The total DNA is then tagged with digoxygenin. The DNA is denatured and hybridized with a population of specifiers containing target regions specific for a particular microorganism, or a panel of microorganisms, in a small volume (0.01 - 2.0 μl). The sample DNA-specifier complexes are isolated using anti-digoxygenin antibodies, and washed extensively to remove excess specifiers. The specifiers are then processed for labeling and imaging as described above.

EXAMPLE V

Label Configurations

[0125] This example shows a DNA probe molecule (specifier) that contains labels that are separated by a distance sufficient such that the labels can be individually resolved (Figure 2). For example, in the case of fluorescent labels, this large spatial separation of labels allows for spectral identification of each label individually.

[0126] As shown in Figure 2, a DNA probe (specifier) can be generated that contains four distinct fluorescent labels. These labels are separated by a distance of about one micron. In this example, the labels are separated by about 2,000 to 3,000 bases of DNA, however other polymers could be used to separate the labels as well.

[0127] A diverse population of unique labels can be generated by varying the fluorescent label at each of the four locations. For example, a set of 10 distinct fluorphores can be used to generate 10,000 unique labels. Specifically, for a probe with four different labels, there will be 10x10x10x10 = 10,000 possible combinations. Similarly, for a probe that has five different labels, there will be 100,000 possible combinations.

EXAMPLE VI

Flow-stretch Separation of Labels

[0128] This example shows a DNA probe molecule (specifier) with one label monomer which has been stretched on a coverslip using a flow stretch technique. The flow stretch technique can also be used for a probe molecule with multiple label monomers along the probe molecule.

[0129] A double stranded DNA probe was constructed with a single stranded section at a desired location for label attachment. In this example, the single stranded section is at the end of the probe molecule. The label molecule contains about 300 Cy3 fluorophores and a section of DNA that is complementary to the single stranded area of the probe molecule. After hybridization, the probe molecule is stained with YOYO1 dye to verify that the label is attached to the

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 182

probe. The staining procedure is not necessary when multiple labels are attached. The probes with hybridized label were then stretched on a coverslip using a flow stretch technique (Henegariu et al., *supra).* An image, figure 3, was obtained using an inverted fluorescent microscope. Probes can also be stretched using a receding meniscus technique (Yokota et al., *supra*).

EXAMPLE VII

Electro-stretch Separation of Labels

[0130] This example shows than an electric field can be used to align probe (specifier) molecules that are attached by one end to a surface. In this manner, labels that are along the probe molecule are spatially separated.

[0131] A flow cell was constructed that allowed imaging of the probe molecules through a coverslip. The flow cell also allowed for the exchange of solutions and the application of an electric field. In order to prevent bubbling, a gel was used to separate the electrodes from the solution. With gel separation of the electrodes and solution, higher voltages are possible which improves the alignment of the probe molecules. The coverslip was coated with bovine serum albumin (BSA) to minimize non-specific binding. DNA was added to the coverslip and attached to the coverslip primarily by the ends of the DNA. Un-bound DNA was washed away and the remaining bound DNA was stained with YOYO1. Imaging was performed using an epi-fluorescent microscope. When the voltage was off, tethered DNA can be seen to move randomly (Figure 4A). However, when the voltage is on, the DNA becomes aligned (Figure 4B).

EXAMPLE VIII

Cytometer Detection of Labels

[0132] Another way to align DNA probes (specifiers) spatially is by the use of a constriction in the flow of a liquid containing the DNA in conjunction with an oscillating electric field (Asbury, C.L., and van den Engh, G., Biophys. J. 74: 1024-1030 (1998)). In this way, when the probe molecules pass through the constriction the labels are spatially separated (Figure 5). Confocal optics can allow for detection with enough spatial resolution to determine the sequence of the labels.

SEQUENCE LISTING

[0133]

<110> The Institute for Systems Biology
Dimitrov, Krassen
Dunaway, Dwayne

<120> Methods For Detection and Quantification
of Analytes In Complex Mixtures

<130> FP-IS 5320

<140> PCT/US02/21278
<141> 2002-07-03

<150> US 09/898,743
<151> 2001-07-03

<160> 11

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 183

<223> synthetic primer

<400> 1

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 120
actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 180
actctctctc tctctctctc gctctctctc tctctctctc 220

<210> 2
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 2

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 120
actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 3
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 3

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 120
actctctctc tctctctctc actctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 4
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 4

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 184

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 120
actctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 5
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 5

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 120
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 6
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 6

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
actctctctc tctctctctc actctctctc tctctctctc gctctctctc tctctctctc 120
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 7
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 7

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
actctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 120
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 8
<211> 220
<212> DNA

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 185

<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 8

actctctctc tctctctctc actctctctc tctctctctc actctctctc tctctctctc 60
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 120
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 9
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 9

actctctctc tctctctctc actctctctc tctctctctc gctctctctc tctctctctc 60
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 120
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 10
<211> 220
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 10

actctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 60
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 120
gctctctctc tctctctctc gctctctctc tctctctctc gctctctctc tctctctctc 180
gctctctctc tctctctctc gctctctctc tctctctctc 220

<210> 11
<211> 10
<212> DNA
<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 11
gagagagaga 10

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 186

Claims

1. A diverse population of uniquely labeled probes, comprising thirty or more target-specific nucleic acid probes, each attached to a unique label bound to a nucleic acid, each nucleic acid probe being a nucleic acid molecule comprising a target-specific region and a region comprising genedigits, said genedigits being nucleic acid regions of predetermined nucleotide sequences that are specifically bound by complementary anti-genedigit nucleotide sequences, each said anti-genedigit nucleotide sequence being attached to a label monomer or a combination of label monomers, wherein the population is in solution, and wherein the label has a detectable signal that is fluorescence, luminescence, colorimetric, voltage, current, or a magnetic field.

2. The diverse population of claim 1, comprising thirty or more target-specific nucleic acid probes, wherein each nucleic acid probe comprises four or more genedigits, said genedigits being nucleic acid regions of predetermined nucleotide sequence that are specifically bound by complementary anti-genedigit nucleotide sequences, each said anti-genedigit nucleotide sequence being attached to a label monomer or a combination of label monomers, wherein the population is in solution, and wherein the label has a detectable signal that is fluorescence, luminescence, colorimetric, voltage, current, or a magnetic field.

3. The diverse population of any one of claims 1 to 2, wherein each probe is synthetic.

4. The diverse population of any one of claims 1 to 3, wherein each probe is DNA attached to a bridging nucleic acid.

5. The diverse population of any one of claims 1 to 4, wherein said label monomers are attached to dendrimers.

6. The diverse population of any one of claims 1 to 5 wherein the genedigits are DNA.

7. The diverse population of any one of claims 1 to 6, wherein the probes comprise tags.

8. The diverse population of claim 7, wherein the tags are biotin tags.

9. The diverse population of any one of claims 1 to 8, wherein a quantum dot is used as a label monomer.

10. The diverse population of any one of claims 1 to 8, wherein a fluorophore is used as a label monomer.

11. The diverse population of any one of claims 1 to 10, comprising 50, 100, 200, 500, 1,000, 2,000, 5,000, 1×10^4, 3×10^4, 1×10^5 or more different nucleic acid probes.

12. The diverse population of any one of claims 1 to 11, wherein the nucleic acid probes are separated from one another.

13. The diverse population of claim 12, wherein the nucleic acid probes are spread on a two-dimensional surface.

14. The diverse population of any one of claims 1 to 13, wherein one or more of the target-specific probes is hybridized to an analyte and wherein the signal from the hybridized target-specific probe uniquely identifies the analyte.

15. A method of detecting a nucleic acid analyte, comprising:

    (a) contacting a mixture of nucleic acid analytes with the diverse population of uniquely labeled probes of any one of claims 1 to 11 under conditions sufficient for hybridization of the probes to target nucleic acid analytes, and
    (b) measuring the resulting signal from one or more of the target specific probes hybridized to an analyte, where the signal uniquely identifies the analyte.

16. A method of quantifying a nucleic acid analyte, comprising:

    (a) contacting a mixture of nucleic acid analytes with the diverse population of uniquely labeled probes of any one of claims 1 to 11 under conditions sufficient for hybridization of the probes to target nucleic acid analytes;
    (b) measuring the resulting signal from one or more of the target specific probes hybridized to an analyte, where the signal uniquely identifies the analyte, thereby identifying the target nucleic acid analyte; and
    (c) counting the individual molecules of the nucleic acid analyte in the sample,
    thereby quantifying the nucleic acid analyte.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 187

17. The method according to any one of claims 15 and 16, wherein the position of the label monomer or combination of label monomers is used as a distinguishing characteristic.

18. The method according to any one of claims 15 to 17, wherein the nucleic acid probes are subjected to a flow-stretching technique, to a receding meniscus technique, to an electrostretching technique, or to constriction in the flow of a liquid containing said nucleic acid probes in conjunction with an oscillating electric field.

Patentansprüche

1. Eine verschiedenartige Population von einzigartig markierten Sonden, umfassend 30 oder mehr zielspezifische Nukleinsäuresonden, jeweils mit einer einzigartigen Markierung verbunden, die an eine Nukleinsäure gebunden ist, wobei jede Nukleinsäuresonde ein Nukleinsäuremolekül ist, welches eine zielspezifische Region und eine Genedigits umfassende Region umfasst, wobei die Genedigits Nukleinsäureregionen mit vorherbestimmten Nukleotidsequenzen sind, die spezifisch durch komplementäre Anti-Genedigit Nukleotidsequenzen gebunden werden, wobei die Anti-Genedigit Nukleotidquenz mit einem Markierungs-Monomer oder einer Kombination von Markierungs-Monomeren verbunden ist, wobei die Population sich in Lösung befindet, und wobei die Markierung ein detektierbares Signal aufweist, das Fluoreszenz, Lumineszenz, kolorimetrisch, Spannung, Strom oder ein Magnetfeld ist.

2. Die verschiedenartige Population nach Anspruch 1, umfassend 30 oder mehr zielspezifische Nukleinsäuresonden, wobei jede Nukleinsäuresonde vier oder mehr Genedigits umfasst, wobei die Genedigits Nukleinsäureregionen mit vorherbestimmten Nukleotidsequenzen sind, die spezifisch durch komplementäre Anti-Genedigit Nukleotidsequenzen gebunden werden, wobei die Anti-Genedigit Nukleotidsequenz mit einem Markierungs-Monomer oder einer Kombination von Markierungs-Monomeren verbunden ist, wobei die Population sich in Lösung befindet, und wobei die Markierung ein detektierbares Signal aufweist, das Fluoreszenz, Lumineszenz, kolorimetrisch, Spannung, Strom oder ein Magnetfeld ist.

3. Die verschiedenartige Population nach einem der Ansprüche 1 oder 2, **dadurch gekennzeichnet, dass** jede Sonde synthetisch ist.

4. Die verschiedenartige Population nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet, dass** jede Sonde DNA ist, die mit einer überbrückenden Nukleinsäure verbunden ist.

5. Die verschiedenartige Population nach einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet, dass** die Markierungs-Monomere mit Dendrimeren verbunden sind.

6. Die verschiedenartige Population nach einem der Ansprüche 1 bis 5, **dadurch gekennzeichnet, dass** die Genedigits DNA sind.

7. Die verschiedenartige Population nach einem der Ansprüche 1 bis 6, **dadurch gekennzeichnet, dass** die Sonden Kennzeichnungen umfassen.

8. Die verschiedenartige Population nach Anspruch 7,
dadurch gekennzeichnet, dass die Kennzeichnungen Biotin-Kennzeichnungen sind.

9. Die verschiedenartige Population nach einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet, dass** ein Quantenpunkt als Markierungs-Monomer verwendet wird.

10. Die verschiedenartige Population nach einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet, dass** ein Fluorophor als Markierungs-Monomer verwendet wird.

11. Die verschiedenartige Population nach einem der Ansprüche 1 bis 10, umfassend 50, 100, 200, 500, 1000, 2000, 5000, 1×10^4, 3×10^4, 1×10^5 oder mehr unterschiedliche Nukleinsäuresonden.

12. Die verschiedenartige Population nach einem der Ansprüche 1 bis 11, **dadurch gekennzeichnet, dass** die Nukleinsäuresonden voneinander getrennt sind.

13. Die verschiedenartige Population nach Anspruch 12, **dadurch gekennzeichnet, dass** die Nukleinsäuresonden auf

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 188

einer zweidimensionalen Oberfläche ausgebreitet sind.

14. Die verschiedenartige Population nach einem der Ansprüche 1 bis 13, **dadurch gekennzeichnet, dass** eine oder mehrere der zielspezifischen Sonden mit einem Analyten hybridisiert ist, und wobei das Signal von der hybridisierten zielspezifischen Sonde den Analyten einzigartig identifiziert.

15. Ein Verfahren zur Detektion eines Nukleinsäureanalyten, Schritte umfassend, bei denen man:

(a) eine Mischung von Nukleinsäureanalyten mit der verschiedenartigen Population von einzigartig markierten Sonden nach einem der Ansprüche 1 bis 11 unter Bedingungen in Kontakt bringt, die zur Hybridisierung der Sonden an Zielnukleinsäureanalyten ausreichend sind, und
(b) das sich ergebende Signal von einer oder mehreren der zielspezifischen Sonden, die mit einem Analyten hybridisiert ist, misst, wobei das Signal den Analyten einzigartig identifiziert.

16. Ein Verfahren zur Quantifizierung eines Nukleinsäureanalyten, Schritte umfassend, bei denen man:

(a) eine Mischung von Nukleinsäureanalyten mit der verschiedenartigen Population von einzigartig markierten Sonden nach einem der Ansprüche 1 bis 11 unter Bedingungen in Kontakt bringt, die zur Hybridisierung der Sonden an Zielnukleinsäureanalyten ausreichend sind;
(b) das sich ergebende Signal von einer oder mehreren der zielspezifischen Sonden, die mit einem Analyten hybridisiert ist, misst, wobei das Signal den Analyten einzigartig identifiziert, wodurch der Zielnukleinsäureanalyt identifiziert wird, und
(c) die individuellen Moleküle des Nukleinsäureanalyten in der Probe zählt, wodurch der Nukleinsäureanalyt quantifiziert wird.

17. Das Verfahren nach einem der Ansprüche 15 und 16, **dadurch gekennzeichnet, dass** die Position des Markierungs-Monomers oder der Kombination von Markierungs-Monomeren als unterscheidendes Merkmal verwendet wird.

18. Das Verfahren nach einem der Ansprüche 15 bis 17, **dadurch gekennzeichnet, dass** die Nukleinsäuresonden eine Fluss-Dehnungs-Technik (flow-stretching technique) einer Technik des zurückgehenden Meniskus (receding meniscus technique), einer Elektro-Dehnungs-Technik (electrostretching technique) oder einer Einschränkung in dem Fluss einer die Nukleinsäuresonden enthaltenden Flüssigkeit in Verbindung mit einem schwingenden elektrischen Feld unterworfen werden.

Revendications

1. Population variée de sondes marquées de manière unique comprenant trente ou plus de trente sondes d'acides nucléiques spécifiques d'une cible, chacune étant fixée à un marqueur unique lié à un acide nucléique, chaque sonde d'acide nucléique étant une molécule d'acide nucléique comprenant une région spécifique d'une cible et une région comprenant des empreintes génétiques, lesdites empreintes génétiques étant des régions d'acides nucléiques de séquences nucléotidiques prédéterminées qui se lient spécifiquement à des séquences nucléotidiques complémentaires d'anti-empreinte génétique, chaque séquence nucléotidique d'anti-empreinte génétique étant fixée à un monomère marqueur ou à une combinaison de monomères marqueurs, dans laquelle la population est en solution et dans laquelle le marqueur possède un signal détectable qui est la fluorescence, la luminescence, la colorimétrie, une tension, un courant ou un champ magnétique.

2. Population variée selon la revendication 1, comprenant trente ou plus de trente sondes d'acides nucléiques spécifiques d'une cible, dans laquelle chaque sonde d'acide nucléique comprend quatre ou plus de quatre empreintes génétiques, lesdites empreintes génétiques étant des régions d'acides nucléiques possédant une séquence nucléotidique prédéterminée qui se lie spécifiquement à des séquences nucléotidiques complémentaires d'anti-empreinte génétique, chaque séquence nucléotidique d'anti-empreinte génétique étant fixée à un monomère marqueur ou à une combinaison de monomères marqueurs, dans laquelle la population est en solution et dans laquelle le marqueur possède un signal détectable qui est la fluorescence, la luminescence, la colorimétrie, une tension, un courant ou un champ magnétique.

3. Population variée selon l'une quelconque des revendications 1 à 2, dans laquelle chaque sonde est synthétique.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 189

4. Population variée selon l'une quelconque des revendications 1 à 3, dans laquelle chaque sonde est un ADN fixé à un acide nucléique de pontage.

5. Population variée selon l'une quelconque des revendications 1 à 4, dans laquelle lesdits monomères marqueurs sont fixés à des dendrimères.

6. Population variée selon l'une quelconque des revendications 1 à 5, dans laquelle les empreintes génétiques sont des ADN.

7. Population variée selon l'une quelconque des revendications 1 à 6, dans laquelle les sondes comprennent des étiquettes.

8. Population variée selon la revendication 7, dans laquelle les étiquettes sont des étiquettes de biotine.

9. Population variée selon l'une quelconque des revendications 1 à 8, dans laquelle un point quantique est utilisé comme monomère marqueur.

10. Population variée selon l'une quelconque des revendications 1 à 8, dans laquelle un fluorophore est utilisé comme monomère marqueur.

11. Population variée selon l'une quelconque des revendications 1 à 10, comprenant 50, 100, 200, 500, 1000, 2000, 5000, 1×10^4, 3×10^4, 1×10^5 ou plus sondes d'acides nucléiques différentes.

12. Population variée selon l'une quelconque des revendications 1 à 11, dans laquelle les sondes d'acides nucléiques sont séparées les unes des autres.

13. Population variée selon la revendication 12, dans laquelle les sondes d'acides nucléiques sont étalées sur une surface bidimensionnelle.

14. Population variée selon l'une quelconque des revendications 1 à 13, dans laquelle une ou plusieurs des sondes spécifiques d'une cible s'hybrident à un analyte et dans laquelle le signal provenant de la sonde hybridée spécifique d'une cible identifie exclusivement l'analyte.

15. Procédé de détection d'un analyte d'acide nucléique, comprenant :

(a) le contact d'un mélange d'analytes d'acides nucléiques avec la population variée de sondes marquées de manière unique selon l'une quelconque des revendications 1 à 11 dans des conditions suffisantes pour l'hybridation des sondes aux analytes d'acides nucléiques cibles, et
(b) la mesure du signal résultant provenant d'une ou de plusieurs des sondes spécifiques d'une cible hybridées à un analyte, le signal identifiant exclusivement l'analyte.

16. Procédé de quantification d'un analyte d'acides nucléiques, comprenant :

(a) le contact d'un mélange d'analytes d'acides nucléiques avec la population variée de sondes marquées de manière unique selon l'une quelconque des revendications 1 à 11 dans des conditions suffisantes pour l'hybridation des sondes aux analytes d'acides nucléiques cibles ;
(b) la mesure du signal résultant provenant d'une ou de plusieurs des sondes spécifiques d'une cible hybridées à un analyte, le signal identifiant exclusivement l'analyte, ce qui permet ainsi d'identifier l'analyte d'acide nucléique cible ; et
(c) le comptage des molécules individuelles de l'analyte d'acide nucléique dans l'échantillon,
pour ainsi quantifier l'analyte d'acide nucléique.

17. Procédé selon l'une quelconque des revendications 15 et 16, dans lequel la position du monomère marqueur ou de la combinaison de monomères marqueurs est utilisée comme caractéristique distinctive.

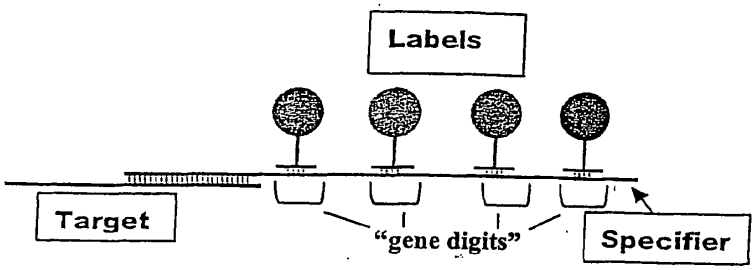
18. Procédé selon l'une quelconque des revendications 15 à 17, dans lequel les sondes d'acides nucléiques sont soumises à une technique d'étirage en écoulement, à une technique du ménisque descendant, à une technique d'électro-étirage, ou à un resserrement dans l'écoulement d'un liquide contenant lesdites sondes d'acides nucléiques

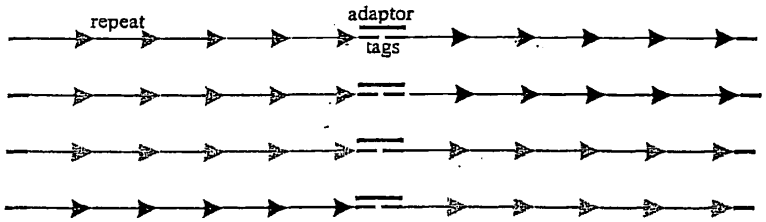
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 190

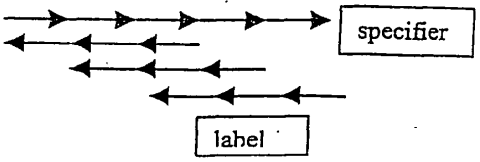
conjointement avec un champ électrique oscillant.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 191

Figure 1

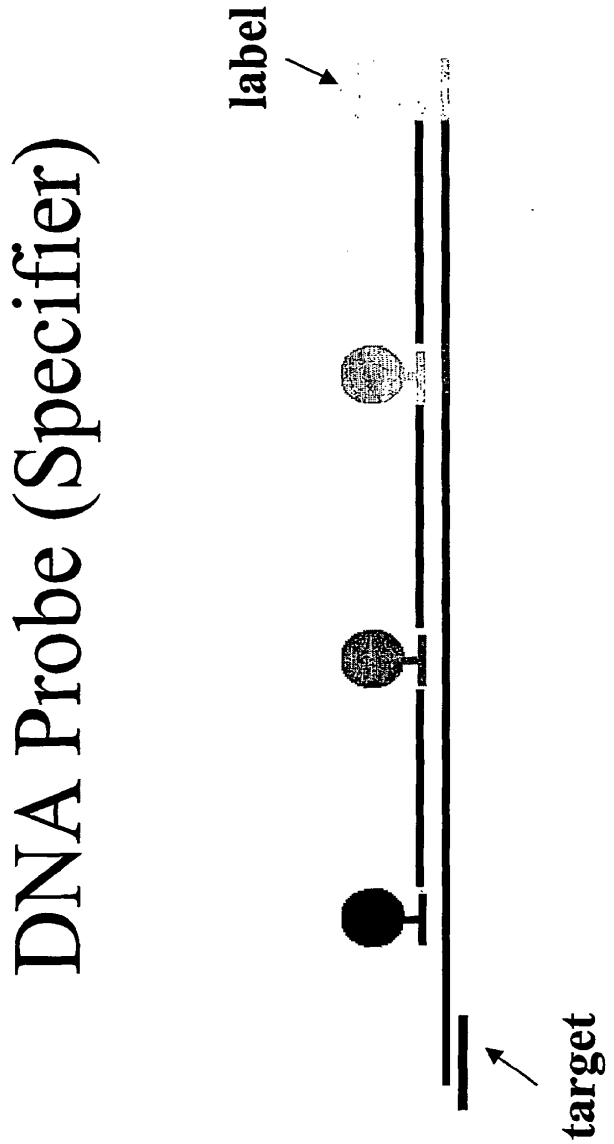






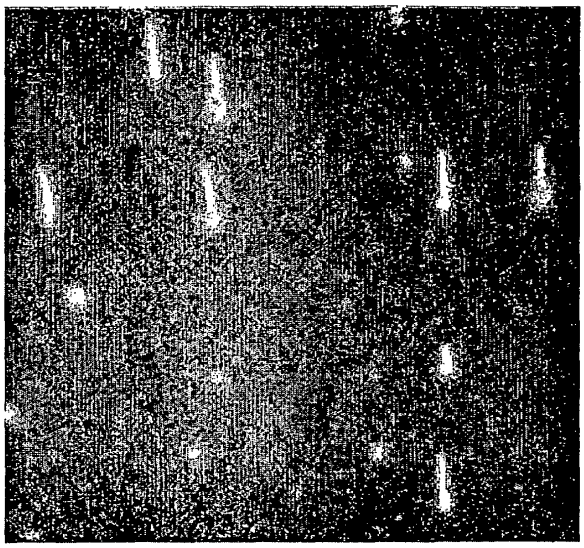
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 192

Figure 2



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 193

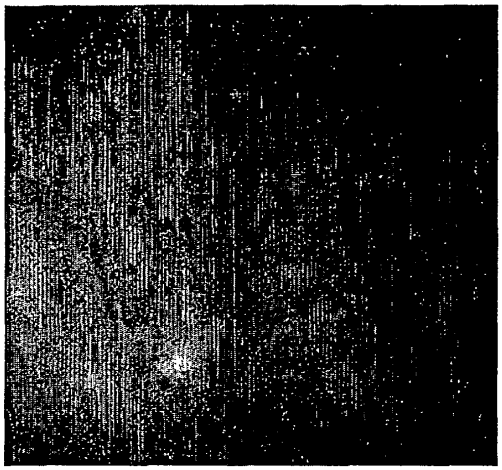
Figure 3



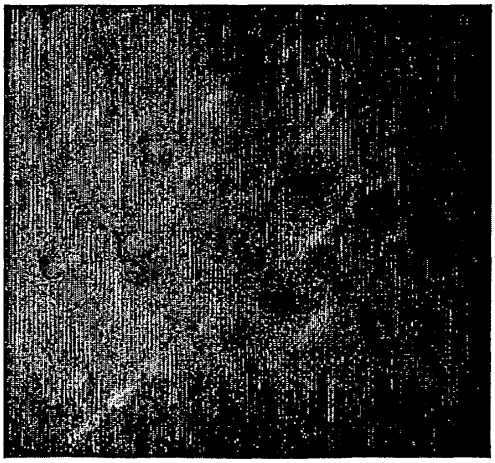
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 194

Figure 4

A

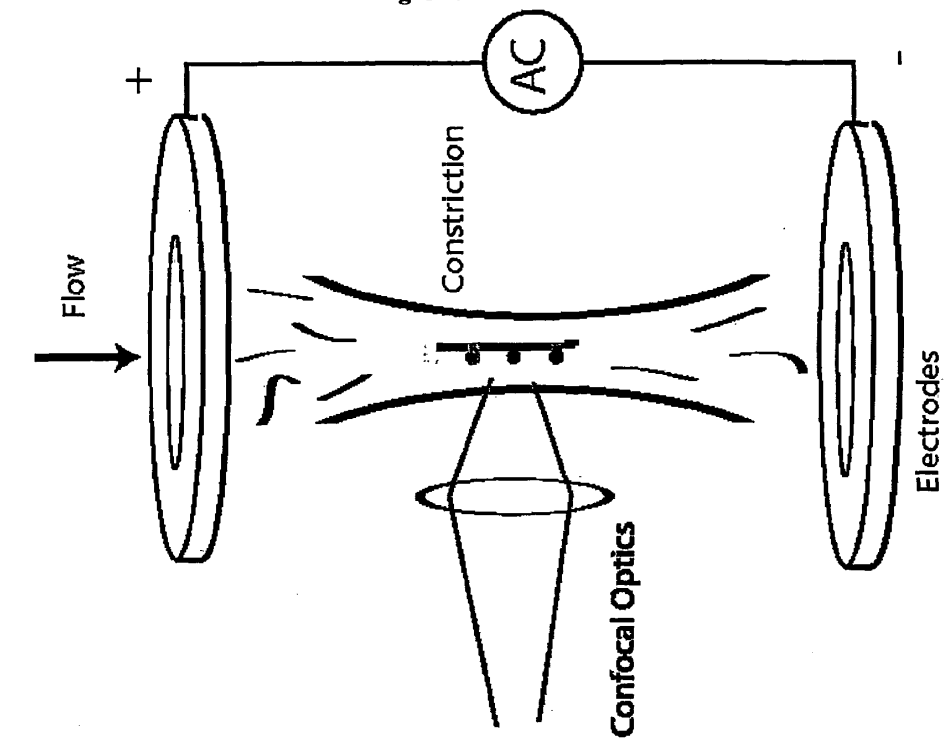


B



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 195

Figure 5



DNA alignment in AC field

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 196

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 5293050 A **[0042]**
- US 5354707 A, Chapple-Sokol **[0042]**
- US 0221278 W **[0133]**
- US 09898743 B **[0133]**

Non-patent literature cited in the description

- **COLLINS et al.** *Nucleic Acids Research,* 1997, vol. 25 (15), 2979-2984 **[0007]**
- **ANDERSON, M.L.M.** Nucleic Acid Hybridization. Springer-Verlag, 1999 **[0027]**
- **WANG ; GIESE.** *Anal. Chem.,* 1993, vol. 65, 3518 **[0046]**
- **JAYASENA, S.D.** *Clinical Chemistry,* 1999, vol. 45 (9), 1628-1650 **[0073]**
- **HAMES ; HIGGINS.** Nucleic Acid Hybridisation: A Practical Approach. IRL Press, 1991 **[0099]**
- Current Protocols in Molecular Biology. Greene Publ, 1989 **[0099]**
- **HENEGARIU et al.** *Biotechniques,* 2001, vol. 31, 246-250 **[0105]**
- **YOKOTA et al.** *Nuc. Acids Res.,* 1997, vol. 25, 1064-1070 **[0105]**
- **MATSUURA et al.** *Nuc. Acids Res.,* 2001, vol. 29, E79 **[0105]**
- **ASBURY, C.L. ; VAN DEN ENGH, G.** *Biophys. J.,* 1998, vol. 74, 1024-1030 **[0132]**

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 197

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 963 531 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
21.09.2011 Bulletin 2011/38

(21) Application number: **06848157.1**

(22) Date of filing: **22.12.2006**

(51) Int Cl.:
C12Q 1/68 *(2006.01)* ***C12P 19/34*** *(2006.01)*

(86) International application number:
PCT/US2006/049274

(87) International publication number:
WO 2007/076128 (05.07.2007 Gazette 2007/27)

(54) **NANOREPORTERS AND METHODS OF MANUFACTURING AND USE THEREOF**

NANOREPORTER UND VERFAHREN ZU DEREN HERSTELLUNG UND VERWENDUNG

NANORAPPORTEURS ET PROCEDES DE PRODUCTION ET D'UTILISATION DE CEUX-CI

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR

(30) Priority: **23.12.2005 US 753758 P**
08.09.2006 US 843528 P

(43) Date of publication of application:
03.09.2008 Bulletin 2008/36

(73) Proprietors:
- **Nanostring Technologies, Inc.**
  Seattle, WA 98119 (US)
- **Institute for Systems Biology**
  Seattle WA 98103-8904 (US)

(72) Inventors:
- **GEISS, Gary, K.**
  Seattle, WA 98117 (US)
- **FERREE, Sean, M.**
  Seattle, WA 98117 (US)
- **WEBSTER, Philippa, J.**
  Seattle, WA 98103 (US)
- **DIMITROV, Krassen, M.**
  Seattle, WA 98115 (US)

(74) Representative: **Crump, Julian Richard John et al**
Mintz Levin Cohn Ferris Glovsky & Popeo Intellectual Property LLP
Alder Castle
10 Noble Street
London EC2V 7JX (GB)

(56) References cited:
WO-A-03/003810 **WO-A-2005/071401**
US-A1- 2002 039 732 **US-A1- 2002 137 036**
US-A1- 2004 091 864 **US-B1- 6 268 147**

- **GEISS GARY K ET AL: "Direct multiplexed measurement of gene expression with color-coded probe pairs." NATURE BIOTECHNOLOGY MAR 2008, vol. 26, no. 3, March 2008 (2008-03), pages 317-325, XP002505107 ISSN: 1546-1696**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

EP 1 963 531 B1

Description

1. FIELD OF THE INVENTION

[0001] The present invention relates to compositions and methods for detection and quantification of individual target molecules in biomolecular samples. In particular, the invention relates to coded, labeled reporter molecules, referred to herein as labeled "nanoreporters," that are capable of binding individual target molecules. Through the nanoreporters' label codes, the binding of the nanoreporters to target molecules results in the identification of the target molecules. Methods of making and using such nanoreporters are also provided. The nanoreporters can be used in diagnostic, prognostic, quality control and screening applications.

2. BACKGROUND OF THE INVENTION

[0002] This invention relates generally to the field of detection, identification, and quantification of target molecules in mixtures.

[0003] 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. These variations in gene expression are at the core of altered physiologic and pathologic processes. Therefore, identifying and quantifying the expression of genes in normal cells compared to diseased cells can aid the discovery of new drug and diagnostic targets.

[0004] Nucleic acids can be detected and quantified based on their specific polynucleotide sequences. The basic principle underlying existing methods of detection and quantification is the hybridization of a labeled complementary probe sequence to a target sequence of interest in a sample. The formation of a duplex indicates the presence of the target sequence in the sample and the degree of duplex formation, as measured by the amount of label incorporated in it, is proportional to the amount of the target sequence.

[0005] This technique, called Molecular hybridization, has been a useful tool for identifying and analyzing specific nucleic acid sequences in complex mixtures. This technique has been used 111 diagnostics, for example, to detect nucleic acid sequences of various microbes in biological samples. In addition, hybridization techniques have been used to map genetic differences or polymorphisms between individuals. Furthermore, these techniques have been used to monitor changes in gene expression in different populations of cells or in cells treated with different agents.

[0006] WO 2005/071401 discloses a method for the simultaneous analysis, in a homogeneous format, for detection of multiple nucleic acid target sequences. The invention utilizes capture probes conjugate to substrates that are distinguishable based on unique spectral signature in combination with reporter probes having a separately detectable label. In particular, the substrates are subsets of microspheres, created by impregnating a combination of two fluorescent dyes in different proportions to produce a unique spectral signature for each subset, In the presence of a target sequence, detectable complexes are formed that are distinguishable for each combination of target, capture probe (carrying bead subsets containing a single spectral signature) and reporter probe (containing a single signal produced by a single fluorescent label).

[0007] US 6,268,147 relates to a method for analyzing genomic DNA and expressed sequences using auxiliary oligonucleotides, preannealed to the single-stranded target nucleic acid to form a partially duplex target molecule. Unlabeled nucleic acid analyte (the "target sequence") is denatured and annealed or hybridized with a molar excess of two or more oligonucleotide probes, at least one of which is labeled, and which bind to target sequences in one or more regions of known sequence, to form a partially duplex structure in which at least two oligonucleotide probes bind to the target sequence in tandem, forming a duplex region in which binding of at least one probe is stabilized by uninterrupted contiguous base stacking with the tandemly hybridizing probe. The probes are simply labeled with detectable labels or tags.

[0008] WO 03/003810 relates to single probes, present in populations. In particular, this document discloses a diverse population of uniquely labeled probes, containing about thirty or more target specific nucleic acid probes each attached to a unique label bound to a nucleic acid, and methods of using these probes to detect a nucleic acid analyte.

[0009] In the past, only a few genes could be detected in a complex sample at one time. Within the past decade, several technologies have made it possible to monitor the expression level of a large number of transcripts within a cell at any one time (see, e.g., Schena et al., 1995, Science 270: 467-470; Lockhart et al., 1996, Nature Biotechnology 14: 1675-1680; Blanchard et al., 1996, Nature Biotechnology 14:1649). In organisms for which most or all of the genome is known, it is possible to analyze the transcripts of large numbers of the genes within the cell. Most of these technologies employ, DNA microarrays, devices that consist of thousands of immobilized DNA sequences present on a miniaturized surface that have made this process more efficient. Using a microarray, it is possible in a single experiment to detect the presence or absence of thousands of genes in a biological sample. This allows researchers to simultaneously perform several diagnostic tests on one sample, or to observe expression level changes in thousands of genes in one experiment.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 199

Generally, microarrays are prepared by binding DNA sequences to a surface such as a nylon membrane or glass slide at precisely defined locations on a grid. Then nucleic acids in a biological sample are labelled and hybridized to the array. The labeled sample DNA marks the exact position on the array where hybridization occurs, allowing automatic detection.

[0010] Unfortunately, despite the miniaturization of array formats, this method still requires significant amounts of the biological sample. However, in several cases, such as biopsies of diseased tissues or samples of a discrete cell type, the biological sample is in limited supply, In addition, the kinetics of hybridization on the surface of a microarray is less efficient than hybridization in small amounts of aqueous solution. Moreover, while methods exist to estimate the amount of nucleic acid present in a sample based on microarray hybridization result, microarray technology thus far does not allow for detection of target molecules on an individual level, nor are there microarray-based methods for directly quantifying the amount of target molecule in a given sample.

[0011] Thus, there exists a need for accurate and sensitive detection, identification and quantification of target molecules in complex mixtures. Discussion or citation of a reference herein shall not be construed as an admission that such reference is prior art to the present invention.

3. SUMMARY OF THE INVENTION

[0012] The present invention relates to methods for the generation of a diverse populations of uniquely-labeled molecules, preferably synthetic molecules, referred to herein as nanoreporters, that can be used for the detection, identification, and direct quantification of a wide variety of target molecules. The methods are advantageous in that they generate large numbers of distinctly labeled reporter molecules, each capable of detecting a single target molecule, starting from just a small number of different types of label monomers.

In certain embodiments, the present invention provides a dual nanoreporter, or a "probe pair," comprising two components: a first probe and a second probe. The first probe is a complex comprising: (a) a first molecule, or a scaffold, comprising: (i) a first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a first signal and a second label attachment region, which is non-overlapping with the first label attachment region, to which are attached (directly, or indirectly) one or more label monomers that emit light constituting a second signal; (ii) a first target-specific sequence attached to the first molecule; and (iii) an affinity tag attached to said first molecule. The second probe is a second molecule comprising (i) a second target-specific sequence; (ii) optionally, a third label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a third signal; and (iii) an affinity tag attached to said second molecule; wherein the first target-specific sequence and the second target-specific sequence bind to different regions of the same target molecule wherein the target molecule is a naturally occurring molecule or a DNA thereof or a complement of said cDNA; wherein when said probe pair is bound to its target molecules, the identity of the first and second signals and their locations relative to each other constitutes at least part of a code that identifies the target molecule, wherein the first, second and third label attachment regions are predetermined nucleotide sequences,

wherein, the first, second, and third signals are spatially distinguishable, and

wherein a first DNA or RNA molecule is hybridized to the first label attachment region, to which said first DNA or RNA molecule is bound said one or more label monomers that emit light constituting said first signal; and wherein a second DNA or RNA molecule is hybridized to the second label attachment region, to which said second DNA or RNA molecule is bound said one or more label monomers that emit light constituting said second signal.

[0013] In alternative embodiments, the present invention provides a dual nanoreporters, or a "probe pair," comprising two components: a first probe and a second probe. The first probe is a complex comprising: (i) a first molecule comprising: a first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a first signal and a second label attachment region which is non-overlapping with the first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a second signal; (ii) a first target-specific sequence attached to the first and (iii) an affinity tag attached to said first molecule. The second probe is a second molecule comprising (i) a second target-specific sequence; (ii) optionally, a third label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a third signal ; and (iii) an affinity tag attached to said second molecule; wherein the first signals, the second signal and the third signal are spatially or spectrally distinguishable; wherein the first target-specific sequence and the second target-specific sequence bind to different regions of the same target molecule wherein the target molecule is a naturally occurring molecule or a cDNA thereof or a complement of said DNA, wherein when said probe pair is bound to its target molecule, the identity of the first and second signals and their locations relative to each other constitutes at least part of a code that identifies the target molecule, wherein the first, second, and third label attachment regions are predetermined nucleotide sequences, and

wherein a first DNA or RNA molecule is hybridized to the first label attachment region, to which said first DNA or RNA molecule is bound said one or more label monomers that emit light constituting said first signal; and wherein a second

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 200

DNA or RNA molecule is hybridized to the second label attachment region, to which said second DNA or DNA molecule is bound to said one or more label monomers that emit light constituting said second signal.

[0014] The code can consist of the first and second signals and their locations relative to each other, or can consist of the first and second signals and additional signals similarly generated from additional label attachment regions (on the first or second molecule) and the locations of the signals relative to each other.

[0015] The target molecule is a naturally occurring molecule or a cDNA of a naturally occurring molecule or the complement of said cDNA.

[0016] In a first aspect of the dual nanoreporters of the invention in which the target attachment regions are nucleotide sequences, a first DNA molecule is hybridized to the first label attachment region, to which first DNA molecule are bound (directly or indirectly) said one or more label monomers that emit light constituting said first signal; and wherein a second DNA molecule is hybridized to the second label attachment region, to which second DNA molecule are bound (directly or indirectly) said one or more label monomers that emit light constituting said second signal.

[0017] In a second aspect or the dual nanoreporters of the invention in which the target attachment regions are nucleotides sequences, a first RNA molecule is hybridized to the first label attachment region, to which first RNA molecule are bound (directly or indirectly) said one or more label monomers that emit light constituting said first signal; and wherein a second DNA molecule is hybridized to the second label attachment region, to which second DNA molecule are bound (directly or indirectly) said one or more label monomers that emit light constituting said second signal.

[0018] In a third aspect of the dual nanoreporters of the invention in which the target attachment regions are nucleotide sequences, a plurality of first DNA molecules are hybridized to the first label attachment region, to which DNA molecules are bound (directly or indirectly) said one or more label monomers that emit light constituting said first signal; and wherein a plurality of second DNA molecules are hybridized to the second label attachment region, to which second DNA molecules are bound (directly or indirectly) said one or more label monomers that emit light constituting said second signal.

[0019] In a fourth aspect of the dual nanoreporters of the invention in which the target attachment regions are nucleotide sequences, a plurality of first RNA molecules are hybridized to the first label attachment region, to which RNA molecules are bound (directly or indirectly) said one or more label monomers that emit light constituting said first signal; and wherein a plurality of second RNA molecules are hybridized to the second label attachment region, to which second RNA molecules are bound (directly or indirectly) said one or more label monomers that emit light constituting said second signal.

[0020] In certain embodiments of the dual nanoreporters of the invention, the second probe is a nucleic acid complex comprising: (a) the second nucleic acid (scaffold) molecule, wherein the second nucleic acid molecule comprises a third label attachment region to which is hybridized a third RNA molecule, to which third RNA molecule are bound (directly or indirectly) one or more label monomers that emit light constituting a third signal; and (b) the second target-specific sequence covalently attached to the second nucleic acid molecule, wherein the code comprises the identity of the first, second and third signals and their locations relative to each other. In certain embodiments, the label monomers attached to the first label attachment regions emit light at the same wavelength, which light constitutes said first signal, and the label monomers attached to the second label attachment region emit light at the same wavelength, which light constitutes the second signal, and the label monomers attached to the third label attachment region emit light at the same wavelength, which light constitutes the third signal. In other embodiments, at least one of the first signal, the second signal and the third signal comprises light at a plurality of different wavelengths. The first, second and third signals may be spectrally distinguishable. Alternatively, the first and third signals emit at the same wavelength or wavelengths.

[0021] The first probe and the second probe comprise an affinity tag. Also, in specific embodiments, the second probe comprises a third label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a third signal. Preferably, the first and second molecules are nucleic acid molecules; more preferably, the label attachment regions and target-specific sequences are predetermined nucleotide sequences. In certain embodiments of the dual nanoreporters of the invention, the one or more label monomers attached to one or both of the label attachment regions are covalently bound to nucleic acids hybridized to their respective attachment regions, or are covalently bound to nucleic acids that are bound to the respective label attachment regions indirectly through one or more bridging nucleic acids.

[0022] In certain embodiments of the dual nanoreporters of the invention, the second target-specific sequences are unlabeled with any one or more of said label monomers; however, at least one target-specific sequence in a nanoreporter is attached to a scaffold that is directly or indirectly attached to one or more label monomers.

[0023] In other embodiments of the dual nanoreporters of the invention, the second molecule further comprises a fourth label attachment region to which is attached one or more label monomers that emit light constituting a fourth signal. Preferably, when such a dual nanoreporter is bound to its target molecule, the code comprises the identity of the first signal, second signal, third signal and fourth signals and their locations relative to each other.

[0024] However, the second nucleic acid need not comprise a third label attachment region to which are attached (directly or indirectly) one or more labels monomers that emit light constituting a third signal. Such probes are referred to herein as "ghost probes." Ghost probes contain a target-specific sequence, which improves the hybridization kinetics of a nanoreporter to its target molecule, and, optionally, an affinity tag that can be used to immobilize and stretch the

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 201

dual nanoreporter. In nanoreporter embodiments employing a ghost probe, the first molecule (to which the label monomers are attached and is sometimes referred to herein as a "reporter probe") is preferably a nucleic acid molecule of about 2,000 to about 10,000 bases in length, more preferably about 5,000 to about 8,000 bases in length, and the second molecule (the "ghost probe") is preferably a nucleic acid molecule of about 40 to about 250 bases in length, more preferably about 50 to 100 bases in length. In a specific embodiment, both the first molecule and the second molecule are DNA molecules.

[0025] As used herein, the use of the terms "about" and "approximately" before a number or range means that number or range plus or minus 5%.

[0026] An affinity tag can be indirectly attached to the first probe or second probe. For example an affinity tag can be covalently attached to an oligonucleotide hybridized to the first or second probe, or it can be covalently attached to an oligonucleotide hybridized to a "patch" that is hybridized to the scaffold of the first and/or second probe. Alternatively, the affinity tag is covalently attached to the first nucleic acid molecule and/or to the second, nucleic acid molecule of the first and second probes.

[0027] The signals emitted by the label monomers attached to a given label attachment region can be the same, or different. Thus, in one embodiment, the label monomers attached to a specific (e.g., first or second) label attachment region emit light at the same wavelength, which light constitutes the signals corresponding to that Label attachment region. Alternatively, in another embodiment, at least one of the first signal and the second signal comprises light at a plurality of different wavelengths.

[0028] Preferably, in certain embodiments, signals emitted by label monomers attached to adjacent labels attachment regions are spectrally distinguishable. Thus, in specific embodiments, the first and second signals are spectrally distinguishable.

[0029] We also disclose a probe, which can be a singular nanoreporter or one or both of the two components of a dual nanoreporter, said probe being a nucleic acid complex comprising (a) a nucleic acid molecule comprising (i) a first label attachment region to which is hybridized a first RNA molecule, to which first RNA molecule are bound (directly or indirectly) one or more label monomers that emit light constituting a first signal; (ii) a second label attachment region to which is hybridized a second RNA molecule, to which second RNA molecule are bound (directly or indirectly) one or more label monomers that emit light constituting a second signal; and (b) a target-specific sequence covalently attached to the nucleic acid molecule, wherein the first and second RNA molecules preferably are each at least 20 nucleotides, wherein the target-specific sequence binds to a target molecule, which target molecules is a naturally occurring molecule or a cDNA of a naturally occurring molecule or the complement of said cDNA, and wherein when said probe is bound to its target molecules, the identify of the first and second signals and their locations relative to each other constitutes at least part of a code that identifies the target molecule. The first and second RNA molecules preferably are each at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90 or at least 100 nucleotides,

[0030] Additionally, we disclose a probe, said probe being a nucleic acid complex comprising (a) a synthetic nucleic acid molecule comprising (i) a first labels attachment region to which is hybridized a first RNA molecule, to which first RNA molecule are attached one or more label monomers that emit light constituting a first signal; (ii) a second label attachment region to which is hybridized a second RNA molecule, to which second RNA molecule are attached one or more label monomers that emit light constituting a second signal ; and (b) a target-specific sequence covalently attached to the synthetic nucleic acid molecule, wherein the first and second RNA molecules are each at least 50 nucleotides, wherein the target-spepific sequence binds to a target molecule, which target molecule is a naturally occurring molecule or a DNA of a naturally occurring molecule or the complement of said cDNA, and wherein when said probe is bound to its target molecule, the identity of the first and second signals and their locations relative to each other constitutes at least part of a code that identifies the target molecule.

[0031] Also disclosed herein is a probe, said probe being a nucleic acid complex comprising: (a) a nucleic acid molecule comprising (i) a first label attachment region to which is hybridized a first RNA molecule, to which first RNA molecule are covalently attached one or more label monomers that emit light constituting a first signal; (ii) a second label attachment region to which is hybridised a second RNA molecules, to which second RNA molecule are covalently attached one or more label monomers that emit light constituting a second signal; and (b) a target-specinc sequence covalently attached to the nucleic acid molecule, wherein the first and second RNA molecules are each at least 50 nucleotides, wherein the target-specific sequence binds to a target molecules, which target molecule is a naturally occurring molecule or a DNA of a naturally occurring molecule or the complement of said cDNA, and wherein when said probe is bound to its target molecule, the identity of the first and second signals and their locations relative to each other constitutes at least part of code that identifies the target molecule.

[0032] In certain aspects of the present invention, the label monomers are attached to the label attachment regions of the nanorsporters of the invention by way of one or more RNA patches, or RNA molecules. Such RNA patches are preferably about (+/- 5%) 100 to about 3,000 nucleotides each, more preferably about 500 to about 1,500 nucleotides each.

[0033] The affinity tag can be attached to a patch and, in a dual nanoreporter, the affinity tag can be attached to one or more patches on one or both components of the probe pair.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 202

[0034] The affinity tag can be attached by being covalently attached to an oligonucleotide hybridized to one or more patches on a nanoreporter. Alternatively, the affinity tag can be attached to the nucleic acid molecule that is the scaffold of a nanoreporter.

[0035] Also disclosed herein are nanoreporter-target molecule complexes. Optionally, one or more components of the nanoreporter complexes are covalently attached to an affinity tag for purification and/or for immobilization (either before or after stretching, as described in Section 5,12 below). Alternatively, the target molecule itself is attached to an affinity tag. For example, where the target molecule is a nucleic acid, the affinity tag can be bio tin that is incorporated into the nucleic acid as biotin-modificd nucleotides. The target molecule can be immobilized through the affinity tag prior to or after formation of the nanoreporter-target molecule complex. Where the affinity tag is a biotin moiety, the target molecule can be immobilized on a solid surface that is coated with avidin or streptavidin.

[0036] One or all of the components (e.g., scaffold, target-specific sequence) of a nanoreporter can be (or have the sequence of) a naturally-occurring molecule. However, the fully assembled and labeled nanoreporter is generally a synthetic molecule, for example a chimeric molecule made by joining naturally-occurring sequences (e.g., a viral or plasmid- based scaffold and a mammalian target-specific sequence) and/or manmade sequences.

[0037] The nanoreporters of the invention, when complexed to their target molecules, are preferably imaged when immobilized and stretched. The nanoreporter and complexes comprising the nanoreporters and target molecules can be immobilized by any method known in the art. Preferably, the nanoreporter is attached to an affinity tag (e.g., biotin or digoxigenin) that can be used to tether the nanoreporter to a surface containing a ligand for the affinity tag (e.g., streptavidin or an anti-digoxigenin antibody, respectively). The immobilized nanoreporter can be stretched by any method known in the art, including but not limited to flow-stretching, a receding meniscus technique, electrostretching, constriction in the flow of a liquid containing the nanoreporter in conjunction with an oscillating electric field.

[0038] We also disclose herein populations often or more nanoreporters, or probes, each probe comprising a synthetic molecule, said synthetic molecule comprising, sequentially in the following order: (a) a first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a first signal; (b) a second label attachment region to which is attached one or more label monomers that emit light constituting a second signal; (c) a third label attachment region to which is attached one or more label monomers that emit light constituting a third signal; and (d) a target-specific sequence that binds to a target molecule; wherein the first and second signal are spectrally distinguishable; wherein the second and third signals are spectrally distinguishable; wherein the first and second signals are not spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein the second and third signals are not spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein the first and third signals are spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein each probe comprises a target-specific sequence that binds to a different target molecule; and wherein the identities of the first, second and third signals and the location of the first and third signal relative to each other constitute at least part of a code that identifies each target molecule.

[0039] Preferably, the target molecule is a naturally occurring molecule or a cDNA of a naturally occurring molecule or the complement of said DNA.

[0040] Also preferably, the label attachment regions and target-specific sequences are predetermined nucleotide sequences.

[0041] In certain aspects of the invention, the probe pairs of the invention are purified.

[0042] In specific embodiments, the label monomers attached to the first label attachment region emit light at the same wavelength, which light constitutes said first signal; wherein the label monomers attached to the second label attachment region emit light at the same wavelength, which light constitutes the second signal; and the label monomers attached to the third label attachment region emit light at the same wavelength, which light constitutes the third signal. Thus, in such embodiments, the label monomers attached to a particular label attachment region emit light at the same wavelength (s),

[0043] In other specific embodiments, at least one of the first signal, second signal, and third signals comprises light at a plurality of different wavelengths. In such embodiments, two or more monomers attached to at least one label attachment region emit light at different wavelengths.

[0044] We also disclose herein a population often or more nanoreporter- target molecule complexes, each complex comprising a synthetic molecule bound to a target molecule, said synthetic molecule comprising, sequentially in the following order: (a) a first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a first signal; (b) a second labels attachment region to which is attached one or more label monomers that emit light constituting a second signal; (c) a third libel attachment region to which is attached one or more label monomers that emit light constituting a third signal; and (d) a target-specific sequence that binds to said target molecule; wherein the first and second signals are spectrally distinguishable; wherein the second and third signals are spectrally distinguishable; wherein the first and second signals are not spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein the second and third signals are not spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein the first and third signals are

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 203

spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein each target-specific sequence (or, in a dual nanoreporter, pair of target-specific sequences) binds to a different said target molecule; and wherein the identities of the first, second and third signals and the locations of the first and third signal relative to each other constitute at least part of a code that identifies each target molecule.

[0045] Optionally, the populations is immobilized to a solid surface, and is preferably stretched. For example the population can be subjected to flow-strelching, a receding meniscus technique, electrostretching, or constriction in the flow of a liquid containing said complexes in conjunction with an oscillating electric field.

[0046] The present invention yet further provides methods of detecting a target molecule, said methods comprising: (a) contacting a biomolecular sample with a probe pair according to the invention under conditions that permit binding of the first target - specific sequence and the second target-specific sequence to the target molecule and (b) detecting the code that identifies a target molecule.

[0047] Preferably, the probe and the target molecule are nucleic acids. In certain specific embodiments, the label monomers attached to the first label attachment region emit light at the same wavelength, the label monomers attached to the second label attachment region emit light at the same wavelength, and the label monomers attached to the third label attachment region emit light at the same wavelength. In such embodiment, each of the first signal, second signal, and third signal consists of light at a single wavelength.

[0048] In other embodiments, at least one of the first signal, second signal, and third signal comprises light at a plurality of different wavelengths.

[0049] We also disclose an artificial nucleic acid molecule comprising one or more synthetic regions (hereinafter, "Region"), each Region being at least 50 nucleotides and comprising a regularly repeated base, said regularly repeated base having a periodicity of four to twenty five, such that said type of base occurs at regular intervals of every fourth to every twenty fifth position in said region. In a specific embodiment, not more than 5% of said base in said Region appears at other than said regular intervals. The artificial nucleic acid can be a single-stranded nucleic acid (e.g., useful as a nanoreporter scaffold) or at least partially (or fully) double-stranded (e.g., a nanoreporter scaffold to which patches are hybridized). The artificial nucleic acid can be RNA, DNA, or a DNA/RNA hybrid. Preferably the regularly repeated base is a thymidine or a uracil. In specific embodiments, the periodicity of the regularly repeated base in said one or more Regions is about 6 to about 15; alternatively, the periodicity of the regularly repeated base in said one or more Regions is about 8 to about 12.

[0050] Preferably, the artificial nucleic acid comprises a plurality of Regions, for example about 2 to about 50 Regions, or about 3 to about 10 Regions.

[0051] The plurality of Regions can be immediately adjacent to one another, or separated by one or more spacer regions. Spacer regions can include, for example, one our more restriction endonuclease recognition sequences (or the complement thereof) and/or one or more RNA polymerase promoter sequence (or the complement thereof).

[0052] Optionally, the plurality of Regions differ from one another in sequence.

[0053] The artificial nucleic acid may further comprises an origin of replication.

[0054] Preferably, the GC content of the one or more Regions is about 25% to about 75%, more preferably about 40 to about 60%, and most preferably is about 50% (+/- 5%).

[0055] Each of said one or more Region in the artificial nucleic acids of the invention is preferably about 50 to about 50,000 bases in length, and is more preferably about 200 to about 2,000 bases in length.

[0056] The artificial nucleic acids disclosed herein are Preferably about 50 to about 100,000 bases in length, and are more preferably about 200 to about 10,0000 or about 500 to about 20,000 bases in length.

[0057] The artificial nucleic acids disclosed herein may comprise one or a plurality of Regions, which are useful as label attachment regions. Where an artificial nucleic acid comprise a plurality of Regions, the type of regularly repeated base can be the same in the plurality of Regions, or different. Also, the periodicity of the regularly repeated base can be the same in the plurality of Regions, or different. We therefore disclose an artificial nucleic acid comprising a plurality of Regions, in which both the type and the periodicity of the regularly repeated base are the same in the plurality of Regions.

[0058] The artificial nucleic acids disclosed herein may further comprise a target-specifie sequence.

[0059] Preferably, the Regions do not comprise a recognition site for at least one six-cutter restriction endonucleotides, such as one of the following restriction endonucleases: BamHI, EcoRI, PstI, SmaI, XbaI, NotI, SacI, EcoRV, HindIII, ClaI, SalI, XhoI, ApaI or KpnI. In a specific embodiment, the Regions do not comprise any recognition site for 6-cutter enzymes, or at least 6-cutter enzymes that are used for cloning within the artificial nucleic acid, for example 6-cutter enzymes present in the polylinker of a plasmid or in regions flanking the Regions.

[0060] The artificial nucleic acids or at least the Regions therein, preferably are not predicted by the MFOLD program to contain hairpin loops.

[0061] The artificial nucleic acids or at least the Regions therein, preferably do not have direct or inverted repeats that are greater than 12 bases in length. Alternatively, the artificial nucleic acids and/or Regions do not have direct or inverted repeats that are greater than about 11, about 10 or about 9 bases in length,

[0062] The artificial nucleic acids preferably comprises a promoter sequence (such as *T7,* T3 or SP6) or its complement.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 204

[0063] The artificial nucleic acids may further comprise a phage encapsidation sequence.

[0064] The Regions in the artificial nucleic acids may be flanked by restriction endonuclease recognition sites.

[0065] The artificial nucleic acids can be used as nanoreporter scaffolds, in which the Regions are useful as labels attachment regions. Accordingly, we disclose an artificial nucleic acid comprising one or more Regions, the regularly repeated base of said one or more Regions being attached to at least one light-emitting label monomer. At least one regulary-repeated base can be covalently attached to the label monomer, or it can be attached to the label monomer via hybridization to another nucleic acid to which the label monomer is covalently attached. Optionally, at least two different Regions are attached to a different light-emitting labeling monomer.

[0066] The artificial nucleic acids may further comprise an affinity tag and/or an origin of replication.

[0067] Optionally, an artificial nucleic acid is a partially or fully double- stranded molecules in which one strand is DNA to which is hybridized a plurality of RNAS, each RNA comprising a single region in which the regularly-repeated base is aminoallyl- modified uracil. Optionally, at leas 50% of occurrences of said aminoallyl- modified uracil is covalently attached to at least one light-emitting label monomer; in other embodiments, the percentage of occurrences is at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or at least 98%. This percentage of occurrences can be measured by any means known in the art.

[0068] As discussed elsewhere, the nanoreporters of the invention may comprise nucleic acid scaffold that are attached to patches and, optionally, flaps. Details of such structures are discussed in Section 5.3, infra, and exemplary patch and flap structures are depicted in Figure 7.

[0069] The present invention further provides methods of detecting a plurality of target molecules, said methods comprising: (a) contacting a biomolecular sample that may contain the target molecule with a population of probe pairs according to the invention under conditions that permit binding of the first target-specific sequence and the second target-specific sequence of each probe pair to their respective target molecule, wherein each probe pair in said population when bound to its respective target molecule is associated with a distinguishable code; and (b) detecting the codes that idenitify the plurality of the target molecules.

[0070] In certain aspects, the methods nanoreporter is a synthetic molecule comprising: (i) a first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a first signal, all said first signals being the same; (ii) a second label attachment region to which is attached one or more label monomers that emit light constituting a second signal, al said second signals being the same; (iii) a third label attachment region to which is attached one or more label monomers that emit light constituting a third signals, all said third signals being the same; wherein each attachment region comprises a plurality of patch pairs; wherein the first and second signals are spectrally distinguishable; wherein the second and third signals are spectrally distinguishable, and determining whether or not the nanoreporter code is present and/or produced comprises (b) detecting said first, second and third signals under conditions; wherein the first and second signals are not spatially resolvable; wherein the second and third Signals are not spatially resolvable; wherein the first and third signals are spatially resolvable under said conditions; and wherein the identities of the first, second and third signals and the locations of the first and third signal relative to each other constitute at least part of a code that identifies a target molecule. Thereby, the target molecule is detected.

4. BRIEF DESCRIPTION OF THE FIGURES

[0071] FIGURE 1A-1F: Figure IA illustrates a dual nanoreporter with a 16-position nanoreporter code, using two 8-position nanoreporter components. Figure 1B illustrates a dual nanoreporter with a 9-position nanoreporter code, using one 8-position nanoreporter component and one single-position nanoreporter component. Figure 1 C illustrates a dual nanoreporter with an 8-position nanoreporter code, using one ghost probe and one 8-position nanoreporter component. Figure ID illustrates a single nanoreporter with an 8-position nanoreporter code. In Figures 1A-11 D, the star shape (depicted with an arrow) is illustrative are discussed in Section 5.3, *infra,* and exemplary patch and flap structures are depicted in Figure 7.

[0072] The present invention further provides methods of detecting a target molecule, said methods comprising: (a) contacting a biomolecular sample that may contain the target molecule with any of the nanoreporter structures described herein under conditions that permit binding of the synthetic molecule to the target molecule, (b) subjecting the conditions that permit formation and/or detection of the nanoreporter code and (c) determining whether or not the nanoreporter code is produced or present, thereby detecting said target molecule.

[0073] In certain aspects of the invention, the methods nanoreporter is a synthetic molecule comprising: (i) a first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a first signal, all said first signals being the same; (ii) a second label attachment region to which is attached one or more label monomers that emit light constituting a second signal, all said second signals being the same; (iii) a third label attachment region to which is attached one or more label monomers that emit light constituting a third signal, all said third signals being the same; wherein each attachment region comprises a plurality of patch pairs; wherein the first and second signals are spectrally distinguishable; wherein the second and third signals are spectrally distinguishable, and

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 205

determining whether or not the nanoreporter code is present and/or produced comprises (b) detecting said first, second and third signals under conditions: wherein the first and second signals are not spatially resolvable; wherein the second and third signals are not spatially resolvable; wherein the first and third signals are spatially resolvable under said conditions; and wherein the identities of the first, second and third signals and the locations of the first and third signal relative to each other constitute at least part of a code that identifies a target molecule. Thereby, the target molecule is detected.

4. BRIEF DESCRIPTION OF THE FIGURES

[0074]

FIGURE 1A-1F: Figure 1A illustrates a dual nanoreporter with a 16-position nanoreporter code, using two 8-position nanoreporter components. Figure 1B illustrates a dual nanoreporter with a 9-position nanoreporter code, using one 8-position nanoreporter component and one single-position nanoreporter component. Figure 1 C illustrates a dual nanoreporter with an 8-position nanoreporter code, using one ghost probe and one 8-position nanoreporter component. Figure 1D illustrates a single nanoreporter with an 8-position nanoreporter code. In Figures 1A-1D, the star shape (depicted with an arrow) is illustrative of an affinity tag, which can be used to purify the nanoreporter or immobilize the nanoreporter (or nanoreporter-target molecule complex) for the purpose of imaging. The numbered region in Figure 1A-1D refer to separate label attachment regions. All except for position 12 of Figure 1A are labeled with one of four types of label monomers, depicted as grey, white, hatched or stripe "sun" diagrams. Position 12 of Figure 1A is an unlabeled "dark spot." Figures 1E and 1F represent variations on the nanoreporters of Figures 1B and 1D, respectively, in which the target molecule to which the nanoreporters are bound comprises biotin moieties (shown as small asterisks), for example biotin-modified nucleotides randomly incorporated into a target nucleic acid. The nanoreporters themselves further optionally comprise an affinity tag (not shown).

FIGURE 2A-2C: Figure 2A shows an illustration of a label unit of a nanoreporter, containing a scaffold with patch units and corresponding split flaps disposed along its length. Figure 2B illustrates the components of a single patch pair and its corresponding flap, containing: 1: a portion of a nanoreporter scaffold (e.g., M13 single-stranded DNA); 2: A patch pair; 3: a split flap pair; and 4: labeled oligonucleotides, each with a label monomer incorporated, hybridized to the split flap. Figure 2C shows a nanoreporter with 4 "spots," each spot designed to contain 9 patch pairs of 60-65 nucleotides, each attached to a split flap pair of 95-100 nucleotides. Each split flap pair had binding site for 12 oligonucleotides each attached to a single label monomer. Each spot therefore had binding sites for 108 label monomers.

FIGURE 3: A nanoreporter in which the patches are RNA segments can be used with (Figure 3A) and without registers (Figure 3B). Both Figure 3A and 3B depict a (1) nanoreporter scaffold (heavy black line) to which are attached (2) 8 RNA segments (heavy grey lines 1-8), (3) a target-specific sequence (dotted line "T") and (4) an oligonucleotide (checkered line "0") that is partly complementary to the nanoreporter scaffold and partly complementary to the target-specific sequence. This oligonucleotide is referred to as a "ligator" oligonucleotide. In Figure 3A, only one register, i.e., every alternate RNA segment is labeled. The second register positions serve as "spacers," making it possible to generate a nanoreporter code in which consecutive positions in the code are the same "color," or spectrally indistinguishable. In Figure 3B, both registers, *i.e.*, adjacent RNA segments with no intervening spacers, are labeled, with no nearest neighbor of the same "color."

FIGURE 4: Is an image of a dual nanoreporter hybridized to a target molecule. Here, both registers are labeled. The nanoreporters are labeled with three different colors, Alexa 488, Cy3 and Alexa 647 (labeled 1, 2 and 3, respectively). The left brackets show one probe of the dual nanoreporter and the right brackets show the other probe of the dual nanoreporter. Colors 1, 2 and 3 were each acquired in different channels and the first and second registers, seen as rows of spots, were shifted up by several pixels to be able to show each register individually.

FIGURE 5A-5D: This figure illustrates the various components of the dual nanoreporters shown in Figure 4. Figure 5A illustrates one color (here, Alexa 488, depicted in the left column as open circles), which is spectrally distinguishable from Cy3 (shown in Figure 5B, depicted in the left column as vertically striped circles) and Alexa 647 (shown in Figure 5C as diagonally striped circles). The images obtained from each were superimposed to generate Figure 5D.

FIGURE 6A-6E: Figure 6A is a schematic illustration of the experiment shown in Figures 6B and 6C. In this case, the star represents biotin that was used to attach the complex by one end to the surface prior to stretching. Figures 6B and 6C show images from experiments in which S2-A ghost probe, S2-B labeled nanoreporter and S2 target DNA (Figure 6B) or S2 target RNA (Figure 6C) were hybridized. Figure 6E shows a close-up of a nanoreporter complexes from Figure 6B, each containing S2-A ghost probe, S2-B labeled nanoreporter and S2 target DNA. Figure 6D shows an image of a negative control experiment, in which S2-A ghost probe, S2-B labeled nanoreporter and no S2 target RNA were hybridized.

FIGURE 7A-G. Figures 7A, 7B, 7C and 7D depict different permutations of patches on a nanoreporter scaffold,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 206

Figures 7E and 7F depict different permutations of split flaps on a nanoreporter scaffold, optionally hybridized to one or more oligonucleotides, as in Figure 7G. In Figure 7A-G, α refers to a 5' or 3' molecule or end of a molecule, and β refers to a corresponding 3' or 5' molecule or end of a molecule.

FIGURE 8: Figure 8 depicts a scheme in which single-stranded M13 phage is linearized for use as a nanoreporter scaffold. The circular M13 phage is annealed to a fivefold excess of BamHI cutter oligonucleotide (hatched lines) (1), and the resulting partially double-stranded M13 digested with the restriction endonuclease BamHI (2), resulting in a linearized M13 in which BamHI cutter oligonucleotide is still attached (3). This M13-oligonucleotide complex is heated in the presence of an excess oligonucleotide complementary to the BamH1 cutter oligonucleotide (an "anti-BamH1 oligonucleotide") (grey lines) (4). The BamH1 cutter oligonucleotide anneals to the excess of anti-BamH1 oligonucleotide, and the M13 molecule is purified from the oligonucleotide, for example by using size exclusion columns, to yield M13 scaffold.

FIGURE 9A-9B: Shows a labeled nanoreporter with an affinity tag at each end, A1 and A2. In Figure 9, the labeled nanoreporter is immobilized through the binding of A1 to an immobilized affinity partner. In the absence of an affinity binding partner for A2, the A2 end of the nanoreporter remains in solution (Figure 9A), but in the presence of an affinity binding partner (A2'), the A2 end of the nanoreporter is also immobilized (Figure 9B). Upon immobilization, the nanoreporter can be stretched, or "elongated" as depicted in Figure 9B, for example by electrostretching, for separation of the label attachment regions in a manner that permits detection of the nanoreporter code.

FIGURE 10A-10C: Figure 10A shows a labeled nanoreporter containing a single affinity tag, A1. Another affinity tag, A2, can be attached to the nanoreporter by direct binding of the nanoreporter to a molecule containing A2 (e.g., if the nanoreporter is or comprises a nucleic acid, it can hybridize directly with another nucleic acid to which A2 is attached), as depicted in Figure 10B. Alternatively, the second affinity tag, A2, can be attached to the labeled nanoreporter via a bridging molecule, such as the bridging nucleic acid ("X") depicted in Figure 10C.

FIGURE 11A-11B: Shows a labeled (nucleic acid-based) nanoreporter with an affinity tag, A1, at one end. In Figure 11, the labeled nanoreporter is immobilized through the binding of A1 to an immobilized affinity partner. The other end of the nanoreporter is in solution (Figure 11A), but can be immobilized by hybridization to a complementary oligonucleotide which contains another affinity tag (A2) used to immobilize the nanoreporter (Figure 11B). A and A2 can be the same, for example biotin, for immobilization on an avidin- or streptavidin-coated surface. Upon immobilization of A1, the nanoreporter can be stretched, or "elongated" as depicted in Figure 11, for example by electrostretching, for separation of the label attachment regions in a manner that permits detection of the nanoreporter code. Optionally, while the nanoreporter is in an elongated state, A2 is introduced and binds the end of the nanoreporter that is complementary to A2 down to the surface.

FIGURE 12A-12B. Figure 12A provides an illustration of a nanoreporter comprising an immobilized first portion F1; and Figure 12B provides an illustration of a nanoreporter extended in an electrical field and comprising immobilized first portion F1 and immobilized second portion F2, wherein F2 is immobilized via a complex with molecule F3.

FIGURE 13A- 3C. Figure 13A provides an illustration of a three-member complex for immobilization of an extended nanoreporter; FIG. 13B provides an illustration of a two- member complex for immobilization of an extended nanoreporter; and FIG. 13C provides an illustration of an incomplete complex for immobilization of an extended nanoreporter.

FIGURE 14A-14D. Figure 14A provides an illustration off a nanoreporter comprising an immobilized first portion F1; FIG. 14B provides an illustration of an extended nanoreporter immobilized at first portion F1 and at a second portion via complexes with F2; FIG. 14C provides an illustration of a nanoreporter comprising a first portion immobilized to an avidin surface via biotin: and FIG. 14D provides an illustration of an extended nanoreporter immobilized at a first portion and at a second portion via selective binding of biotin to an avidin surface.

FIGURE 15A-15C. Figure 15A illustrates immobilization of one terminus of a DNA molecule in a microfluidic device; Figure 15B illustrates extension of the DNA in an electric field; and Figure 15C illustrates selective immobilization of a second terminus of the extended DNA molecule.

FIGURE 16 provides an image of extended nanoreporters selectively immobilized by the methods of the present invention.

FIGURE 17 depicts the relationship between the number of label attachment regions to the calculated entanglement threshold for nanoreporters tor label attachment region sizes of 900 brand 1100 bp

FIGURE 18 is a scatter plot showing normalized and average log2 signal values from each positive sample (n=3) for all 509 genes whose expression was measured in a nanoreporter multiplex assay as described in Example 9 (Section 14) below.

5. DETAILED DESCRIPTION OF THE INVENTION

[0075] Disclosed herein are nanoreporters, and their manufacture and use. A fully assembled and labeled nanoreporter comprises two main portions, a target-specific sequence that is capable of binding to a target molecule, and a labeled

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 207

region which emits a "code" of signals (the "nanoreporter code") associated with the target-specific sequence. Upon binding of the nanoreporter to the target molecule, the nanoreporter code identifies the target molecule to which the nanoreporter is bound.

[0076] Nanoreporters are modular structures. Generally, a nanoreporter is a molecular entity containing three basic elements: a scaffold containing two or more label attachment regions, one or more patches attached to the scaffold, and a target-specific sequence, also attached to the scaffold. The elements of a nanoreporter can be found in a single molecular entity (a "singular" nanoreporter), or two distinct molecular entities (a "dual" nanoreporter). Each molecular entity may be composed of one molecule or more than one molecule attached to one another by covalent or non-covalent means. Generally, each component of a dual nanoreporter has a target-specific sequence that binds to a different site on the same target molecule. This allows for smaller nanoreporter components with more efficient kinetics of binding of the nanoreporter to the target molecule and better signal:noise ratios resulting from the greater binding specificity.

[0077] The patches attached to a nanoreporter scaffold serve to attach label monomers to a nanoreporter scaffold. Patches may be directly labeled, for example by covalent incorporation of one or more label monomers into nucleic acid patches. Alternatively, patches may be attached to flaps, which maybe labeled directly, for example by covalent incorporation of one or more label monomers into a nucleic acid flap, or indirectly, for example by hybridization of a nucleic acid flap to an oligonucleotide which is covalently attached to one or more label monomers. Where the label monomers attached to a label attachment region are not directly incorporated into a patch or flap, the patch or flap serves as a "bridge" between the label monomer and the label attachment region, and may be referred to as a "bridging molecule," e.g., a bridging nucleic acid.

[0078] Additionally, nanoreporters may have affinity tags for purification and/or for immobilization (for example to a solid surface). Nanoreporters, or nanoreporter-target molecule complexes, are preferably purified in two or more affinity selection steps. For example, in a dual nanoreporter, one probe can comprise a first affinity tag and the other probe can comprise a second (different) affinity tag. The probes are mixed with target molecules, and complexes comprising the two probes of the dual nanoreporter are separated from unbound materials (e.g., the target or the individual probes of the nanoreporter) by affinity purification against one or both individual affinity tags. In the first step, the fixture 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 probes comprising the first affinity tag. At this point only full complexes would be bound. The complexes are finally released from the affinity reagents for the second affinity tag and then preferably stretched and imaged. 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. Immobilizing and stretching nanoreporters using affinity reagents is fully described in US2010-0261026 A1.

[0079] Nanoreporter and nanoreporter-target complexes which are or comprise nucleic acids may be sffinity-punfisd or immobilized using a nucleic acid, such as an oligonucleotide, that is complementary to at least part of the nanoreporter or target. In a specific application where the target includes a poly A or poly dA stretch, the nanoreporter-target complex can be purified or immobilized by an affinity reagent coated with a poly dT oligonucleotide.

[0080] The sequence of signals emitted by the label monomers associated with the various label attachment regions of the scaffold of a given nanoreporter allows for the unique identification of the nanoreporter. A nanoreporter having a unique identity or unique spectral signature is associated with a target-specific sequence that recognizes a specific target molecule or a portion thereof. When a nanoreporter is exposed to a mixture containing the target molecule under conditions that permit binding of the target-specific sequence(s) of the nanoreporter to the target molecule, the target-specific sequence(s) preferentially bind(s) to the target molecule. Detection of the spectral code associated with the nanoreporter allows detection of the presence of the target molecule in the mixture (qualitative analysis). Counting all the label monomers associated with a given spectral code or signature allows the counting of all the molecules in the mixture associated with the target-specific sequence coupled to the nanoreporter (quantitative analysis). Nanoreporters are thus useful for the diagnosis or prognosis of different biological states (e.g., disease vs. healthy) by quantitative analysis of know biological markers. Moreover, the exquisite sensitivity of single molecule detection and quantification provides by the nanoreporters 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 detect a correlation with a particular biological state using traditional Molecular methods. The sensitivity of nanoreporter-based Molecular detection permits detailed pharmacokinetic analysis of therapeutic and diagnostic agents in small biological samples.

[0081] Many nanoreporters, referred to as singular nanoreporters. are composed of one molecular entity, as depicted in Figure 1D. However, to increase the specificity of a nanoreporter and/or to improve the kinetics of its binding to a target molecule, a preferred nanoreporter is a dual nanoreporter composed of two molecular entities, each containing, a different target-specific sequence that binds to a different region of the same target molecule. Various embodiments of dual nanoreporters are depicted in Figures 1A-1C, in a dual nanoreporter, at least one of the two molecular entities is labeled. The other molecular entity is not necessarily labeled. Such unlabeled components of dual nanoreporters are

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 208

referred to herein as "ghost probes" (see Figure 1C) and often have affinity tags attached, which are useful to immobilize and/or stretch the complex containing the dual nanoreporter and the target molecule to allow visualization and/or imaging of the complex.

[0082] Because of their modular structures, nanoreporters may be assembled and labeled in a variety of different ways. For example, a nanoreporter scaffold can be attached to a target-specific sequence (for example by hybridization and, optionally, ligation), and the structure comprising the scaffold and target-specific sequence attached to one or more patches and, where desired, flaps. Alternatively, the nanoreporter scaffold can first be attached to one or more patches (and, optionally, flaps), and the scaffold/patch structure then attached to a target specific sequence. Thus, unless stated otherwise, a discussion or listing of steps in nanoreporter assembly does not imply that a specific route of assembly must be followed.

[0083] Nanoreporter assembly and use is exemplified herein largely by way of description of a variety of nucleic acid-based nanoreporters; however, one of skill in the art would recognize that the methods described herein are applicable to an amino acid-based (or hybrid nucleic acid/amino acid-based) nanoreporter. Illustrative embodiments of partially and fully assembled nanoreporters are listed below.

[0084] At its simplest, a scaffold has at least two label attachment regions capable of being labeled and resolved. The scaffold can be any Molecular entity that allows the formation of label attachment regions on the scaffold that can be separately labeled and resolved. The number of labels attachment regions to be formed on a scaffold is based on the length and nature of the scaffold, the means of labeling the nanoreporter, as well as the type of label monomers emitting a signal to be attached to the label attachment regions of the scaffold. A nanoreporter may have a scaffold including two or more label attachment regions. Suitable scaffold structures include DNA- based scaffolds.

[0085] We also disclose labeled nanoreporter wherein one or more label attachment regions are attached to corresponding label monomers, each label monomer emitting a signal. For example a labeled nanoreporter according to the invention is obtained when at least two label monomers are attached to two corresponding label attachment regions of the scaffold such that these labeled label attachment regions, or "spots," are distinguishable. Label monomers emitting a signal associated with different label attachment regions of the scaffold can emit signals that are spectrally indistinguishable under the detections conditions ("like" signals), or can emit signals that are spectrally distinguishable, at least under the detection conditions (e.g., when the nanoreporter is immobilized, stretched and observed under a microscope).

[0086] We also disclose a nanoreporter wherein two or more label monomers are attached to a label attachment region. The signal emitted by the label monomers associated with said label attachment region produces an aggregate signal that is detected. The aggregate signal produced may be made up of like signals or made up of at least two spectrally distinguishable signals.

[0087] Additionally we disclose a nanoreporter wherein at least two label monomers emitting like signals are attached to two corresponding label attachment regions of the scaffold and said two label monomers are spatially distinguishable. We also disclose a nanoreporter wherein at least two label monomers emitting two distinguishable signals are attached to two neighboring label attachment regions, for example two adjacent label attachment regions, whereby said at least two label monomers are spectrally distinguishable.

[0088] We disclose a nanoreporter wherein two spots emitting like signals are separated by a spacer region, whereby interposing the spacer region allows resolution or better resolution of said like signals emitted by labels monomers attached to said two spots. Optionally, the spacer regions have a length determined by the resolution of an instrument employed in detecting the nanoreporter.

[0089] Also disclosed herein is a nanoreporter with one or more "double spots." Each double spot contains two or more (e.g., three, four or five) adjacent spots that emit like signals without being separated by a spacer region. Double spots can be identified by their sizes.

[0090] A label monomer emitting a signal may be attached covalently or non-covalently (e.g., via hybrization) to a patch that is attached to the label attachment region. The label monomers may also be attached covalently or non-covalently (e.g., via hybrization) to a flap attached to a patch that is in turn attached to the scaffold. The flap can be formed by one molecule or two or more molecules ("flap pieces") that form a split flap.

[0091] We also disclose a nanoreporter associated with spectral code determined by the sequence of signals emitted by the label monomers attached (e.g., indirectly via a patch) to label attachment regions on the scaffold of the nanoreporter, whereby detection of the spectral code allow identification of the nanoreporter.

[0092] Additionally, we disclose a nanoreporter further comprising an affinity tag attached to the nanoreporter scaffold, such that attachment of the affinity tag to a support allows scaffold stretching and resolution of signals emitted by label monomers corresponding to different label attachment regions on the scaffold. Nanoreporter stretching may involve any stretching means known in the art including but not limited to, means involving physical, hydrodynamic or electrical means.

[0093] We further disclose a nanoreporter further comprising flaps attached to label attachment regions of the scaffold, wherein a flap attached to a label attachment region of the scaffold attaches the label monomer corresponding to said label attachment region, thereby indirectly attaching label monomers to corresponding label attachment regions on said scaffold. Optionally, each label monomer comprises a signals emitting portion and an oligonucleotide portion of a pre-

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 209

determined sequence, and the flaps comprise repeats of a flap sequence complementary to the oligonucleotide portion of a corresponding label, whereby one or more label monomers attach to a corresponding label attachment region through hybridization of said oligonucleotide portions of said label monomers to said repeats of said flap sequence thereby producing a labeled nanoreporter.

[0094] A nanoreporter can further include a target-specific sequence coupled to the scaffold. The target-specific sequence is selected to allow the nanoreporter to recognize, bind or attach to a target molecule. The nanoreporters disclosed herein are suitable for identification of target molecules of all types. For example, appropriate target-specific sequences can be coupled to the scaffold of the nanoreporter to allow detection of a target molecule. Preferably the target molecule is DNA (including cDNA), RNA (including mRNA and cRNA), a peptide, a polypeptide, or a protein.

[0095] One embodiment of the invention provides increased flexibility in target molecule detection with label monomers. In this embodiment, a dual nanoreporter (or probe pair) comprising two different molecular entities, each with a separate target- specific region, at least one of which is labeled, bind to the same target molecule. Thus, the target-specific sequences of the two components of the dual nanoreporter bind to different portions of a selected target molecule, whereby detection of the spectral code associated with the dual nanoreporter provides detection of the selected target molecule in a biomolecular sample contacted with said dual nanoreporter.

[0096] The invention also provides a method of detecting the presence of a specific target molecule in a biomolecular sample comprising: (i) contacting said sample with a dual nanoreporter (or probe pair) according to the invention under conditions that allow binding of the target-specific sequences in the dual nanoreporter to the target molecule and (ii) detecting the code that identifies the target molecule. Depending on the nanoreporter architecture, the dual nanoreporter may be labeled before or after binding to the target molecule.

[0097] 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 preferred 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 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, from 500 to 5000 different target molecules, and so on and so forth.

[0098] In certain embodiments, the invention is directed to detecting different splice forms of the same RNA. The different splice forms can be detected using a plurality of nanoreporter probes, each with a different target-specific sequence complementary to a different exon of the same gene.

[0099] Structural stability of a nanoreporter can be increased through ligation of the patches and, optionally, ligation of the split flaps and/or the labeled oligonucleotides hybridized to the split flaps.

[0100] In addition to the qualitative analytical capabilities provided by the nanoreporters of the invention and the analytical techniques based thereon, the nanoreporters of the invention are uniquely suitable for conducting quantitative analyses. By providing a one to one binding between the nanoreporters (whether singular or dual nanoreporters) 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.

[0101] As will be appreciated from the description and examples provided below, the present invention provides numerous advantages. For example, the complex modularity in forming nanoreporters according to the invention allows for systematic creation of libraries of unique nanoreporters having a very high degree of diversity (*e.g.*, millions of uniquely recognizable nanoreporters). This modularity allows flexibility in customizing nanoreporter populations to specific applications which in turn provides significant manufacturing efficiencies. Another advantage that will be appreciated through the following description stems from the flexibility in assembling the nanoreporters of the invention. That is, due to their modular structure, the nanoreporters of the invention can be assembled prior to shipment to a point of use or assembled at the point of use.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 210

5.1 NANOREPORTER NOMENCLATURE

[0102] NANOREPORTER: The term "nanoreporter" refers to a molecular entity that has (i) a molecule ("scaffold") containing at least two label attachment regions; (ii) at least one patch attached to at least one label attachment region; and (iii) a target-specific sequence. As described in detail below, nanoreporters can be singular nanoreporters (all components being in a single molecular entity) or dual nanoreporters (all the components being in two separate molecular entities). Nanoreporters are preferably synthetic, *i.e.*, non-naturally-occurring molecules, for example are chimeric molecules made by joining two or more manmade and/or naturally occurring sequences that normally exist on more than one molecule (*e.g.*, plasmid, chromosome, viral genome, protein, etc.).

[0103] LABELED NANOREPORTER: A labeled nanoreporter is a nanoreporter in which at least one patch of the nanoreporter is attached to one or more label monomers that generate(s) a signal that forms at least part of the nanoreporter code.

[0104] LABEL UNIT: The term "label unit" refers to the non-target-specific portions of a labeled nanoreporter.

[0105] PROBE: This refers to a molecule that has a target-specific sequence. In the context of a singular nanoreporter, the term "probe" refers to the nanoreporter itself; in the context of a dual nanoreporter, the term "probe" refers to one or both of the two components of the nanoreporter.

[0106] PROBE PAIR: This refers to a dual nanoreporter.

[0107] PATCH: The term "patch" refers to a molecular entity attached to the label attachment region of the nanoreporter scaffold, generally for the purpose of labeling the nanoreporter. The patch can have one or more label monomers either directly (covalently or noncovalently) or indirectly attached to it, either prior to or after its attachment to the scaffold.

[0108] FLAP: The term "flap" as used herein refers to a molecular entity attached to a patch or patch pair attached to a label attachment region. The flap is one or more molecule containing label monomers or capable of binding one or more molecules containing label monomers. By providing indirect labeling of the regions, the flaps provide more flexibility in controlling the number of signal emitting monomers associated with a region as well as the nature of those monomers. Flaps may be formed by a single molecular piece or several molecular pieces (*e.g.*, two pieces) forming a "split flap" (see, *e.g.*, Figure 7)

[0109] TARGET-SPECIFIC SEQUENCE: The term "target-specific sequence" refers to a molecular entity that is capable of binding a target molecule. In the context of a nanoreporter, the target-specific sequence is attached to the nanoreporter scaffold. The target molecule is preferably (but not necessarily) a naturally occurring molecule or a cDNA of a naturally occurring molecule or the complement of said cDNA.

[0110] GHOST PROBE: A molecule comprising a target-specific sequence, but which is not labeled with a label monomer that emits a signal that contributes to the nanoreporter code.

[0111] REPORTER PROBE: A molecule comprising a target-specific sequence that is labeled with at least one label monomer that emits a signal that contributes to the nanoreporter code. A singular nanoreporter is a reporter probe, as is a labeled component of a dual nanoreporter.

[0112] F-HOOK and G-HOOK: In the context of a dual nanoreporter, F- and G-hooks are each an affinity tag that is capable of being selectively bound to one of the probes. In preferred embodiments, the F-hook and G-hook are biotinylated oligonucleotides that are hybridizable to respective complementary sequences present in (*e.g.* ,via ligation) or attached to (*e.g.*, via hybridization) the respective nanoreporter probes in a dual nanoreporter. Thus, the F-hooks and G-hooks can be used for purification, immobilization and stretching of the nanoreporter. Generally, where a dual nanoreporter contains one reporter probe and one ghost probe, the G-hook becomes attached to the reporter probe and the F-hook becomes attached to the ghost probe. F-hooks and G- hooks can be biotinylated on either end or internally. They can also be amine-modified to allow for attachment to a solid substrate for affinity purification.

[0113] F-TAG and G-TAG: Tandemly-repeated sequences of about 10 to about 25 nucleotides that are complementary to the F-hook and G-hook, respectively. G-tags and F-tags are attached to the nanoreporter probes. Generally, an F-tag is present in or attached to a ghost probe via a ligator sequence and a G-tag is present in or attached the reporter probe scaffold via a ligator sequence.

[0114] SPOT: A spot, in the context of nanoreporter detection, is the aggregate signal detected from the label monomers attached to a single label attachment site on a nanoreporter, and which, depending on the size of the label attachment region and the nature (*e.g.*, primary emission wavelength) of the label monomer, may appear as a single point source of light when visualized under a microscope. Spots from a nanoreporter may be overlapping or non-overlapping. The nanoreporter code that identifies that target molecule can comprise any permutation of the length of a spot, its position relative to other spots, and/or the nature (e.g., primary emission wavelength(s)) of its signal. Generally, for each probe or probe pair of the invention, adjacent label attachment regions are non-overlapping, and/or the spots from adjacent label attachment regions are spatially and/or spectrally distinguishable, at least under the detection conditions (e.g., when the nanoreporter is immobilized, stretched and observed under a microscope, as described herein).

[0115] Occasionally, reference is made to a spot "size" as a certain number of bases or nucleotides. As would be readily understood by one of skill in the art, this refers to the number of bases or nucleotide in the corresponding label

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 211

attachment region.

[0116] NANOREPORTER CODE: The order and nature (e.g., primary emission wavelength(s), optionally also length) of spots from a nanoreporter serve as a nanoreporter code that identifies the target molecule capable of being bound by the nanoreporter through the nanoreporter s target specific sequence(s). When the nanoreporter is bound to a target molecule, the nanoreporter code also identifies the target molecule. Optionally, the length of a spot can be a component of the nanoreporter code.

[0117] DARK SPOT: The term "dark spot" refers to a lack of signal, or "spot," from a label attachment site on a nanoreporter. Dark spots can be incorporated into the nanoreporter code to add more coding permutations and generate greater nanoreporter diversity in a nanoreporter population.

[0118] REGISTER: The term "register" refers to a set of alternating label attachment regions.

5.2 THE NANOREPORTER SCAFFOLD

[0119] The nanoreporter scaffold can be any molecular entity, more preferably a nucleic acid molecule, containing label attachment regions to which label monomers can be directly or indirectly attached. In one embodiment, the nanoreporter scaffold is a protein scaffold; in a preferred embodiment, the nanoreporter scaffold is a nucleic acid scaffold in which the label attachment regions are single-stranded regions to which other nucleic acids, such as oligonucleotide patches, RNA patches, or DNA patches, can attach by hybridization. In specific embodiments, the nanoreporter scaffold is a nucleic acid molecule.

[0120] There are no particular limitations on the types of scaffolds that are suitable for forming nanoreporters. A scaffold can essentially have any structure including, for example, single stranded linear scaffold, double stranded linear scaffold, single stranded circular scaffold or double stranded circular scaffold. Examples of scaffold structures include, for example, a scaffold made of one molecular entity such as polypeptides, nucleic acids or carbohydrates. A scaffold may also include a combination of structures, for example, a scaffold may be made of one or more polypeptide stretches coupled to one or more carbohydrate stretches.

[0121] Suitable molecular entities for scaffolds according to the invention include polymeric structures particularly nucleic acid based polymeric structures such as DNA. DNA based structures offer numerous advantages in the context of the present invention due at least in part to the vast universe of existing techniques and methodologies that allow manipulation of DNA constructs.

[0122] As indicated above, the scaffold may be single stranded or double stranded. Double stranded scaffold can be either conventional double stranded DNA or a double strand that is composed of a linear single stranded stretch of nucleic acid with patch units or flat-patches attached.

[0123] A scaffold can have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21-100 label attachment regions or more.

[0124] The label attachment regions of a nanoreporter scaffold will vary in size depending on the method of labeling. In various embodiments, a label attachment region can have a length anywhere from 10 nm to 10,000 nm, but is more preferably from 50 nm to 5,000 nm, and is more preferably from 100 nm to 1,000 nm. In various embodiments, the label attachment region is from about 100 nm to about 500 nm, from about 150 nm to about 450 nm, from about 200 nm to about 400 nm, or from 250 to about 350 nm. In a preferred embodiment, the label attachment region corresponds closely to the size of a diffraction-limited spot, *i.e.*, the smallest spot that can be detected with standard optics, which is about 300 nm.

[0125] Where the scaffold is a nucleic acid, 1 nm corresponds to approximately 3 nucleotides; thus, an approximately 300 nm-label attachment region corresponds to approximately 900 bases. In other preferred embodiments, the label attachment region is from about 300 nucleotides to about 1.5kb, from about 450 nucleotides to about 1.35 kb, from about 0.6 kb to about 1.2 kb, or from 0.75 kb to about 1.05 kb.

[0126] An illustrative example of a molecular entity for a nanoreporter scaffold according to the invention is M13 DNA, which is single-stranded. In one embodiment, the nanoreporter scaffold is circular at least partially single stranded DNA, such as circular M 13. In a more preferred embodiment, the nanoreporter scaffold is linear at least partially single stranded DNA, such as linear M13. In a specific embodiment, the M13 single-stranded DNA obtained by operating a cut at the BamHI site of circular M13 DNA. It should be noted that within the context of the present invention, linear DNA provides additional advantages compared to circular DNA. One advantage of using linear DNA in forming a scaffold according to the invention relates to the significantly reduced torsional stress associated with linear DNA. The added torsional stress associated with circular DNA may interfere with the structural integrity of the scaffold upon the addition to the scaffold of other components of the nanoreporter, such as patch units. Severe torsional stress may lead to the breaking of the structure of the scaffold.

[0127] It should be noted however that the nanoreporters where only a few, short label attachment sites are labeled, circular DNA may be suitable.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 212

5.2.1 NOVEL SYNTHETICNANOREPORTERSCAFFOLD

SEQUENCES

[0128] The present invention employs nanoreporter scaffold that are artificial nucleic acid molecules (DNA5 RNA, or DNA/RNA hybrids) designed to have features that optimize labeling and detection of the nanoreporter. In these aspects of the invention, a nanoreporter scaffold is an artificial nucleic acids comprising one or more synthetic sequences from 50 to 50,000 bases long. Accordingly, the nanoreporter scaffold, which is preferably a DNA, is designed to have one or more Regions, useful as label attachment regions, comprising a regular pattern of a particular base (the "regularly-repeated base"). In such regions, the regularly-repeated base occurs with a periodicity of every nth residue, where n is any number, and preferably from 4 to 25.

[0129] Preferably, not more than 25% of the regularly-repeated base in a Region appears at other than said regular intervals. For example, if in a Region of 100 nucleotides there are 12 thymidine bases, and thymidine is the regularly-repeated base, in this aspect of the invention not more than 25% of these, i.e., 3 thymidine bases, appear outside the regular pattern of thymidines. In specific embodiments, not more than 20%, not more than 15%, not more than 10%, not more than 9%, not more than 8%, not more than 7%, not more than 6%, not more than 5%, not more than 4%, not more than 3%, not more than 2% or not more than 1% of said base appears at other than said regular intervals in said region.

[0130] The regularly-repeated base in the Regions in a nanoreporter scaffold, or its complementary regularly-repeated base in an annealed patch (or segment) can be used to attach label monomers, preferably light emitting label monomers, to the nanoreporter in a regular, evenly spaced pattern for better distribution of the nanoreporter signal. Preferably, where a Region is labeled, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or at least 98% of occurrences of the regularly-repeated base is attached to at least one light-emitting label monomer, either by covalent attachment of a label monomer to a base, or by hybridization to a nucleic acid in which the complements of the regularly-repeated base are so-labeled.

[0131] This percentage of occurrences can be measured by any means known in the art. In one method, the amount of nucleic acid produced in a labeling reaction is purified (for example, RNA can be purified using a Qiagen RNeasy kit) and subjected to UV spectrophotometry. The absorbance ("A") at the appropriate wavelengths is measured for each of the nucleic acid (260 nm) and the label monomer whose occurrence is to be measured (e.g., 495 nm for Alexa Fluor 488; 590 nm for Alexa Fluor 594; 650 for Alexa Fluor 647; and 550 nm for Cy3). The absorbance of the nucleic acid is corrected by adjusting the value of the absorbance at 260 nm ("A260") to remove the "noise" contribution from the label monomerby subtracting the absorbance at the peak wavelength for the label monomer (A_{LM}) minus the correction factor for that label monomer. Where the nucleic acid is RNA, the number of label monomers per one thousand nucleotides is calculated according to the formula:

$$\frac{\text{no. of label monomers}}{\text{1000 nucleotides}} = \frac{A260}{A_{LM}} \times \frac{9010}{EC_{LM}} \times 1000$$

where EC_{LM} is the extinction coefficient for the label monomer. From this formula, the percentage of occurrences of the regularly-repeated base that are attached to a light-emitting label monomer can be calculated.

[0132] Generally, the preferred regularly-repeating base in a label attachment region is thymidine, so that the region can be labeled by hybridization to one or more complementary patches (e.g., RNA segments) in which the regularly-repeated base is uridine. This permits the use of amino-allyl-modified UTPs, which are readily commercially available, as label monomer attachment sites, in an otherwise random sequence. Preferably, in addition to the regular periodicity of the Regions, the regions (and the nucleic acid comprising them) contain minimal secondary structure. The overall GC-content is preferably maintained close to 50%, and is preferably consistent over relatively short stretches to make local Tm's similar.

[0133] The artificial nucleic acids of the invention, or at least the Regions therein, preferably do not have direct or inverted repeats that are greater than 12 bases in length. In other embodiments, the artificial nucleic acids and/or Regions do not have direct or inverted repeats that are greater than about 11, about 10 or about 9 bases in length.

[0134] In an exemplary Region in which the regularly-repeated nucleotide is a thymidine and a GC content of approximately 50%, excess adenines would make up the loss in abundance of T's. To generate the selected sequence, random sequences with fixed patterns of T's ranging from every 4th base to every 25th base are created and screened to minimize the presence of inverted and direct repeats.

[0135] Sequences are also screened preferably to avoid common six-base-cutter restriction enzyme recognition sites. Selected sequences are additionally subjected to predicted secondary structure analysis, and those with the least sec-

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 213

ondary structure are chosen for further evaluation. Any program known in the art can be used to predict secondary structure, such as the MFOLD program (Zuker, 2003, Nucleic Acids Res. 31 (13):3406-15; Mathews et al., 1999, J. Mol. Biol. 288:911-940).

[0136] An appropriate sequence is divided into label attachment regions ranging from 50 bases to 2 kilobases long (could be longer). Each label attachment region is a unique sequence, but contains a consistent number and spacing of T's in relation to the other label attachment regions in a given reporter sequence. These label attachment regions can interspersed with other regions whose sequence does not matter. The synthetic label attachment regions in a nanoreporter scaffold can be of different lengths and/or have different regularly-repeated bases. An optimized start sequence for transcription by RNA polymerase T7, T3, or SP6 (beginning at position +1 of the transcript) can be added to the 5' end of each label attachment region. Restriction sites are optionally added at the boundaries of each label attachment region to allow specific addition or deletion of individual label attachment regions to the sequence using conventional cloning techniques. The number of synthetic label attachment regions in a nanoreporter preferably ranges from 1 to 50. In yet other embodiments, the number of synthetic label attachment regions in a nanoreporter ranges from 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 synthetic label attachment regions to 15, 20, 30, 40, or 50 synthetic label attachment regions, or any range in between.

[0137] An example of such a novel synthetic label attachment region is given below. In this sequence, shown 5' to 3', the T's are placed in every 8th position and the region is bounded by a 5' Sac I restriction site and a 3' Kpn I restriction site. An optimized transcript start site for T7 polymerase (GGGAGA) is included at the 5' end of the region, downstream of the 5' restriction site. The complement of this sequence, when generated as a single-stranded forms the scaffold for the RNA molecule transcribed ftora this label attachment region.

GAGCTCGGGAGATGGCGAGCTGGAAGCATCAGAAAGTAGGAAGATGACAAAAT
AGGGCCATAGAAGCATGAAGAACTGAACGCATGAGACAATAGGAAGCTACGCC
ACTAGGGACCTGAGAAGCTGAGCGGCTCAGCGGGTCCGAGCGTCAAAAAATAA
AAGAGTGAAACAATAGACGAATGACGCGGTAAAACCATCCAGAAGTAAACGGG
TACAAACATACAGAGATAGCCACCTGGACCAATAGGCACGTACAAACGTACAA
GCCTGGCGCGATGAGGCAATCCACACGTGCAGAGCTGGAACAATGGAAAGATG
CAAGAATAAACCGATACCGGGATCGAGGGCTCAGCGAATAAAGCAGTCAACAA
CTGGAAAGATCCACACATACCGGCGTAACCGAGTCCAAACATACAGACCTGCA
AGACTCGCGACATGGGACGGTAAAACCATCCGACCGTAAACCGGTAACCAGGT
AGCCGGGTAAAAACATAGCAGGGTGGAGACCTCAGAACGTAAAGACGTCCAAG
GGTCGCCGGATAGCGAACTACGCGCATCGCCCAATGGGCCAATCAACAGATAA
ACGAGTAGAAAAGTCAGAAAATAAGAAACTAACGAAATACGAGGGTCCAAGG
ATGCAAGACTGAGGCCCTAAGGAGATAAGGAAATAGGCCGATGCAGACCTGAA
ACGATGCACCGATCCGACGGTAAAAGACTAGACACGTAGCCGGATCAGGGCCT
GGGAGGCTGGAACCGTGAGCACATAGCAAAGTCGCAGCGTCGGCAGATGCGCC
GGTAAAAAAGTAGAGGCATGACCGGATGGGCAAATAGCGACGTACAGCAGTGA
AGCACTAAAAGCATCCAAGGGTAGGAGACTAGGCGCCTCGACGGGTAGGTACC

[0138] The synthetic nucleic acids employed in the present invention can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotide designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the label attachment region and the annealed patches or segments, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the synthetic nucleic acid include 5- fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethy)uracil, 5-carboxymethylaminomethyl-2- thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcyto-

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 214

sine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D- mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil,2-methylthio-N6-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, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3- N-2-carboxypropyl)uracil, (acp3)w, and 2,6-diaminopurine.

[0139] Alternatively, the synthetic nucleic acid can be produced biologically using a vector into which a nucleic acid has been subcloned.

[0140] In various embodiments, the synthetic nucleic acid molecules which may be employed in the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al., 1996, Bioorganic & Medicinal Chemistry 4(1):5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al., 1996, Bioorganic & Medicinal Chemistry 4(1): 5-23; Perry-O'Keefe et al, 1996, Proc. Natl. Acad. Sci. USA 93: 14670- 675.

[0141] In an exemplary embodiment, the selected novel synthetic sequence can be constructed synthetically as double-stranded DNA by a commercial gene synthesis company and cloned in an oriented fashion into a "phagemid", a plasmid vector containing an MI3 or f1 phage intergenic (IG) region which contains the cis-acting sequences necessary for DNA replication and phage encapsidation, such as pUC119. The appropriate orientation of the cloned insert relative to the phage origin of replication allows for the generation of a single- stranded DNA scaffold which is the reverse complement of the RNA molecules generated by in vitro transcription for each label attachment region.

[0142] In order to generate the single-stranded DNA scaffold of the novel reporter, the phagemid is transformed into an E. coli strain containing an F' episome. Subsequent infection of the transformed bacteria with a helper phage such as the MI 3 mutant K07 results in the secretion of the phagemid carrying the novel reporter sequence as a single-stranded, packaged phage from which the circular, single-stranded DNA is prepared using a standard protocol. This DNA is linearized and the vector portion is excised by annealing short, complementary oligonucleotides to either end of the novel reporter sequence to generate double-stranded restriction sites, followed by treatment with the appropriate restriction enzymes.

[0143] To make the RNA molecules (patches or "segments") for each label attachment region, polymerase chain reaction ("PCR") primers are designed to generate a double-stranded template beginning with an RNA polymerase promoter (T7, T3, or SP6) directly upstream (5') of the transcription start site and ending following the 3' restriction enzyme site. Using this template, *in vitro* transcription of RNA molecules is performed in the presence of amino-allyl modified regularly-repeated base in the RNA (*e.g.*, UTP) and unmodified other bases (*e.g.*, ATP, CTP and GTP). This leads to an RNA product in which every regulalry-repeated base (*e.g.*, U) is modified to allow covalent coupling of a label monomer at that position in the RNA molecule.

[0144] Coupling of light-emitting label monomers to the RNA molecules and annealing of the labeled RNA molecules to the scaffold are carried out as described below.

[0145] Some design considerations for the de novo sequence are listed in Table 1 below.

Feature Of Synthetic Scaffold	Advantages
Novel synthetic sequence	Can be of any length and be designed to incorporate any desired sequence feature including but not limited to those listed in this table.
Minimal secondary structure (select against inverted repeats)	Allows for consistent transcription of full-length RNA molecules. Allows for consistent annealing of RNA molecules to scaffold at predictable temperatures. Minimizes self-annealing and/or cross-annealing between RNA molecules or scaffolds.
Minimal repeated sequences	Avoids mis-annealing between RNA molecules and inappropriate regions of the scaffold.
Unique restriction sites at borders of label attachment regions	Allows addition and deletion of individual label attachment regions using standard molecular cloning techniques.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 215

(continued)

Feature Of Synthetic Scaffold	Advantages
Defined, even spacing of T's and transcription with amino-allyl-modified UTP (no unmodified UTP)	Controls number of coupling sites for monomers in each label attachment region, allowing for consistent brightness of individual labeled RNA molecules. Controls distance between monomers: spacing can be optimized to avoid stearic hindrance and fluorescence quenching.
Optimized start sequence for transcription by RNA polymerase T7, T3, or SP6	Promotes efficient in vitro transcription of each label attachment region.

5.3 PATCHES

[0146] Label monomers that emit signals which constitute all or part of the nanoreporter code are attached to label attachment region(s) of the nanoreporter scaffold through a structure referred to herein as a "patch." The label monomers can be directly (*e.g.*, covalently or noncovalently) attached to a patch, or indirectly attached to a patch (*e.g.*, through hybridization).

[0147] Nucleic acid patches can by anywhere from 25 nucleotides to several kilobases (*e.g.*, 5kb) in length, and are preferably 50 nucleotides to 2kb in length. In specific embodiments, nucleic acid patches are approximately 25 to 250, 50 to 200, 50 to 150, or 50 to 100 nucleotides in length. In other embodiments, nucleic acid patches are approximately 500 to 2,000, 500 to 1,500, 500 to 1,000, 750 to 1,250, or 750 to 1,000 nucleotides in length. Nucleic acid patches can be RNA patches or DNA patches.

[0148] A label monomer can be covalently attached to a patch before or after the patch is attached to the label attachment region of a nanoreporter scaffold. For example, where the patch is a nucleic acid molecule, the label can be covalently attached by incorporation of a nucleotide containing a label monomer into the nucleic acid during its synthesis but before it is attached, e.g., via hybridization, to the label attachment region of the scaffold. Alternatively, during the synthesis of a nucleic acid patch, a nucleotide containing a label monomer acceptor group be included, and the label monomer added to the nucleic acid patch after its synthesis, either before or after it is attached to the label attachment region of the scaffold. Alternatively, the label monomer can be indirectly attached to the patch, for example by hybridization of the patch to a "flap" that serves as a basis for attachment of the label monomer to the nanoreporter.

[0149] Thus, where a patch is a nucleic acid, it can range anywhere from 20 nucleotides to more than 5 kb in length, depending on the method of assembly of the nanoreporter.

[0150] For example, where a patch has covalently incorporated into it one or more label monomers that emit signals that are part of the nanoreporter code in the context of the labeled nanoreporter, the patch is preferably about 100 to about 10,000 bases, more preferably 200 to about 2000 bases, and yet more preferably 700 to about 1200 nucleotides in length, and is generally referred to herein as a "segment," a "dark "segment being the patch prior to the incorporation of the label monomer (but, in a preferred embodiment, containing label monomer acceptor sites, such as amino allyl nucleotides), and a "colored" segment being one containing the desired label monomer or label monomers. The Tm of a segment when hybridized to its label attachment region preferably is >80°C, more preferably >90°C, in 825 mM Na^+ (5X SSC).

[0151] Where a patch merely serves as a template for flap attachment to the nanoreporter, then it is preferably smaller in size, for example about 25-250 nucleotides in length, an is most preferably about 50-100 nucleotides in length. Such patches are referred to herein as "oligonucleotide patches." As detailed in Section below, an oligonucleotide is preferably partially complimentary in sequence to a scaffold, such that when it is annealed to the scaffold, an overhang is generated that is complementary to all or portion of a flap.

[0152] The terms "segment" and "oligonucleotide patch" are used herein merely for convenience of description, however, there is no size cutoff to distinguish a "segment" from an "oligonucleotide patch." The purpose of both types of structures is to maximize the labeling --and thus signal intensity- from the nanoreporter, thereby allowing for single target molecule detection by a nanoreporter.

[0153] In certain aspects, the present invention employs a synthetic molecule, whose configuration is illustrated by reference to **Figure 7A**, comprising a strand of a nucleic acid **(scaffold)** and a plurality of patch pairs hybridized to said strand, wherein each patch pair comprises an "A" patch and a "B" patch, wherein, for each patch pair, (a) each "A" patch is an oligonucleotide comprising a first region (1P) and a second region **(2P),** said first region being (i) at the alpha end of said "A" patch, and (ii) hybridized to a first portion of said strand, said second region being (ii) at the beta end of said "A" patch; (b) each "B" patch is an oligonucleotide comprising a third region **(3P)** and a fourth region **(4P),** said third region being (i) at the alpha end of said "B" patch, and (ii) hybridized to said second region of said "A" patch, said fourth

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 216

region being (i) at the beta end of said "B" patch and (ii) hybridized to a second portion of said strand, said second portion of said strand being to the beta end of said first portion of said strand, wherein said second region or said third region further comprises at its beta end or alpha end, respectively, a hybridizable region that is not hybridized to said "B" patch or "A" patch, respectively.

[0154] In the synthetic molecule of **Figure 7A**, the second region may further comprise at its beta end a hybridizable region that is not hybridized to said "B" patch, as depicted in **Figure 7B**, or the third region further comprises at its alpha end a hybridizable region that is not hybridized to said "A" patch, as depicted in **Figure 7C.**

[0155] The present invention further employs a synthetic molecule, whose configuration is illustrated by reference to **Figure 7D**, comprising a strand of a nucleic acid **(scaffold)** and a plurality of patch pairs hybridized to said strand, wherein each patch pair comprises an "A" patch and a "B" patch, wherein, for each patch pair, (a) each "A" patch is an oligonucleotide comprising a first region **(1P)** and a second region **(2P),** said first region being (i) at the alpha end of said "A" patch, and (ii) hybridized to a first portion of said strand, said second region being (ii) at the beta end of said "A" patch; (b) each "B" patch is an oligonucleotide comprising a third region **(3P)** and a fourth region **(4P),** said third region being (i) at the alpha end of said "B" patch, and (ii) hybridized to said second region of said "A" patch, said fourth region being (i) at the beta end of said "B" patch and (ii) hybridized to a second portion of said strand, said second portion of said strand being to the first of said first portion of said strand, wherein said second region further comprises at its beta end a first hybridizable region that is not hybridized to said "B" patch, and wherein said third region further comprises at its alpha end a second hybridizable region that is not hybridized to said "A" patch.

[0156] In the synthetic molecule of **Figure 7B,** each patch pair can be attached to a flap pair, as depicted in **Figure 7F,** wherein each flap pair comprises an "A" flap and a "B" flap, wherein, for each flap pair, (a) each "A" flap is an oligonucleotide comprising a first flap region **(1F)** and a second flap region **(2F);** said first flap region being at the alpha end of said "A" flap; said second flap region (i) being at the beta end of said "A" flap and (ii) comprising at its beta end a hybridizable region that is not hybridized to said "A" patch, "B" patch or "B" flap; and (b) each "B" flap is an oligonucleotide comprising a third flap region **(3F)**, a fourth flap region **(4F),** and a fifth flap region **(5F);** said third flap region being (i) being at the alpha end of said "B" flap and (ii) comprising at its alpha end a hybridizable region that is not hybridized to said "A" patch, "B" patch or "A" flap; said fourth flap region (i) being between the third flap region and the fifth flap region and (ii) hybridized to said first flap region of said "A" flap; said fifth flap region being (i) at the beta end of said "B" flap, and (ii) hybridized to said hybridizable region of said second region of said "A" patch.

[0157] In the synthetic molecule of **Figure 7C,** each patch pair can be attached to a flap pair, as depicted in **Figure 7E**, wherein each flap pair comprises an "A" flap and a "B" flap, wherein, for each flap pair, (a) each "A" flap is an oligonucleotide comprising an first flap region **(1F),** a second flap region **(2F),** and a third flap region **(3F);** said "A" flap region being (i) at the alpha end of said "A" flap and (ii) hybridized to said hybridizable region of said third region of said "B" patch; said second flap region being between the first flap region and the third flap region; said third flap region (i) being at the beta end of said "A" flap and (ii) comprising at its beta end a hybridizable region that is not hybridized to said "A" patch, "B" patch or "B" flap, and (b) each "B" flap is an oligonucleotide comprising a fourth flap region **(4F)** and a fifth flap region **(5F);** said fourth flap region being (i) being at the alpha end of said "B" flap and (ii) comprising at its alpha end a hybridizable region that is not hybridized to said "A" patch, "B" patch or "A" flap; said fifth flap region being (i) at the beta end of said "B" flap, and (ii) hybridized to said second flap region of said "A" flap.

[0158] In the synthetic molecule of **Figures 7D** and **7E**, the split flaps can be attached one (*e.g.,* (10)), or more (*e.g.*, **(20)** and **(30))** oligonucleotides, as depicted in **Figure 7G.** Thus, the one or more oligonucleotides can be attached to the all or a portion of the "A" flap individually (*e.g.*, **(10)),** the "B" flap individually (*e.g.*, **(30)),** or span all or a portion of each of the "A" flap and "B" flap (*e.g.*, **(20)).** Such oligonucleotides are preferably covalently bound to one or more label monomers.

[0159] The hybridizable regions of said synthetic molecules may be hybridized to a plurality of oligonucleotides, each bound, preferably covalently bound, to at least one label monomer, more preferably to at least five label monomers. In certain embodiments, all the oligonucleotides attached to a single patch pair comprise the same label monomers, *e.g.*, comprise label monomers that emit light at the same wavelength(s); in specific embodiments, all the oligonucleotides attached to at least two, or at least four, adjacent patch pairs preferably comprise the same label monomers. One or more of the oligonucleotides may be bound to at least one affinity tag.

[0160] In certain preferred embodiments, the label monomers are fluorophores or quantum dots.

[0161] In the synthetic molecule described above, alpha can refers to either 5'or 3', and the corresponding beta to either 3' or 5', respectively.

[0162] The region of complementary in each patch pair, or between a given patch and corresponding flap, is preferably about 20 to 5,000 nucleotides. In certain embodiments, the region of complementary is about 20 to 100 nucleotides, or about 5 to 50 nucleotides.

[0163] In the synthetic molecules described above, each flap is preferably about 50 to 5,000 nucleotides in length. In certain embodiments, each flap is about 50 to 150 nucleotides.

[0164] The synthetic molecules described above may further comprise a target-specific region which binds to a target

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 217

molecule. The target-specific region can be attached to the beta or alpha end of said strand.

[0165] In certain embodiments, the synthetic molecule described above comprise at least ten patch pairs, or at least fifty patch pairs.

[0166] In the synthetic molecules described above, the strand, or scaffold, can be a linearized vector, such as linearized M 13.

[0167] The synthetic molecule described above may further comprise (a) a first label attachment region to which are attached (directly or indirectly) one or more label monomers that emit light constituting a first signal; (b) a second label attachment region, which is non- overlapping with the first label attachment region, to which is attached one or more label monomers that emit light constituting a second signal; (c) a third label attachment region, which is non-overlapping with the first and second label attachment regions, to which is attached one or more label monomers that emit light constituting a third signal; wherein each attachment region comprises a plurality of patch pairs; wherein the first and second signals are spectrally distinguishable; wherein the second and third signals are spectrally distinguishable; wherein the first and second signals are not spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein the second and third signals are not spatially resolvable under conditions that can be used to detect said first, second and third signals; wherein the first and third signals are spatially resolvable under conditions that can be used to detect said first, second and third signals; and wherein the identities of the first, second and third signals and the locations of the first and third signal relative to each other constitute at least part of a code that identifies the target molecule.

5.4 LABEL MONOMERS

[0168] The nanoreporters 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). At least the first and second label attachments regions in the nanoreporter are labeled with one or more label monomers, and the signals emitted by the label monomers attached to the label attachment regions of a nanoreporter constitute a code that identifies the target to which the target-specific region of the nanoreporter binds. In certain embodiments, the lack of a given signal from the label attachment region (i.e., a "dark" spot) can also constitute part of the nanoreporter code. An example of a dark spot is depicted at position 12 of the nanoreporter in Figure 1A.

[0169] Radioisotopes are an example of label monomers that can be utilized by the invention. Several radioisotopes can be used as label monomers for labeling nucleotides or proteins, including, for example, ^{32}P, ^{33}P, ^{35}S, ^{3}H, and ^{125}I. These radioisotopes have different half-lives, types of decay, and levels of energy which can be tailored to match the needs of a particular experiment. For example, ^{3}H is a low energy emitter which results in low background levels, however this low energy also results in long time periods for autoradiography. Radioactively labeled ribonucleotides, deoxyribonucleotides and amino acids are commercially available. Nucleotides are available that are radioactively labeled at the first, or α, phosphate group, or the third, or γ, phosphate group. For example, both [α - ^{32}P] dATP and [γ -^{32}P] dATP are commercially available. In addition, different specific activities for radioactively labeled nucleotides are also available commercially and can be tailored for different experiments.

[0170] Another example of label monomers that can be utilized by the invention are fluorophores. Several fluorophores can be used as label monomers for labeling nucleotides including, for example, fluorescein, tetramethylrhodamine, and Texas Red. Several different fluorophores are known, and more continue to be produced, that span the entire spectrum. Also, different formulations of the same fluorophore have been produced for different applications. For example, fluorescein, can be used in its isothiocynanate form (FITC), as mixed isomer or single isomer forms of carboxyfluorescein succinimidyl ester (FAM), or as isomeric dichlorotriazine forms of fluorescein (DTAF). These monomers are chemically distinct, but all emit light with a peak between 515-520 nm, thereby generating a similar signal. In addition to the chemical modifications of fluorescein, completely different fluorophores have been synthesized that have the same or very similar emission peaks as fluorescein. For example, the Oregon Green dye has virtually superimposable excitation and emission spectra compared to fluorescein. Other fluorophores such as Rhodol Green and Rhodamine Green are only slightly shifted in their emission peaks and so also serve functionally as substitutes for fluorescein. In addition, different formulations or related dyes have been developed around other fluorophores that emit light in other parts of the spectrum.

[0171] Non-radioactive and non-fluorescent label monomers are also available. For example, biotin can be attached directly to nucleotides and detected by specific and high affinity binding to avidin or streptavidin which has been chemically coupled to an enzyme catalyzing a colorimetric reaction (such as phosphatase, luciferase, or peroxidase). Digoxigenin labeled nucleotides can also similarly be used for non-isotopic detection of nucleic acids. Biotinylated and digoxigenin-labeled nucleotides are commercially available.

[0172] Very small particles, termed nanoparticles, also can be used as label monomers to label nucleic acids. These particles range from 1-1000 nm in size and include diverse chemical structures such as gold and silver particles and quantum dots.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 218

[0173] When irradiated with angled incident white light, silver or gold nanoparticles ranging from 40-120 nm will scatter monochromatic light with high intensity. The wavelength of the scattered light is dependent on the size of the particle. Four to five different particles in close proximity will each scatter monochromatic light which when superimposed will give a specific, unique color. The particles are being manufactured by companies such as Genicon Sciences. Derivatized silver or gold particles can be attached to a broad array of molecules including, proteins, antibodies, small molecules, receptor ligands, and nucleic acids. For example, the surface of the particle can be chemically derivitized to allow attachment to a nucleotide.

[0174] Another type of nanoparticle that can be used as a label monomer are quantum dots. Quantum dots are fluorescing crystals 1-5 nm in diameter that are excitable by a large range of wavelengths of light. These crystals emit light, such as monochromatic light, with a wavelength dependent on their chemical composition and size. Quantum dots such as CdSe, ZnSe, InP, or InAs possess unique optical properties.

[0175] Many dozens of classes of particles can be created according to the number of size classes of the quantum dot crystals. The size classes of the crystals are created either 1) by tight control of crystal formation parameters to create each desired size class of particle, or 2) by creation of batches of crystals under loosely controlled crystal formation parameters, followed by sorting according to desired size and/or emission wavelengths. Use of quantum dots for labeling particles, in the context of the present invention, is new, but is old in the art of semiconductors. Two examples of earlier references in which quantum dots are embedded within intrinsic silicon epitaxial layers of semiconductor light emitting/detecting devices are United States Patent Nos. 5,293,050 and 5,354,707 to Chapple Sokol, et al.

[0176] In specific embodiments, one or more of the label attachments regions in the nanoreporter is labeled with one or more light-emitting dyes, each label attachment region containing, directly or indirectly, one or more label monomers. The light emitted by the dyes can be visible light or invisible light, such as ultraviolet or infra red light. In exemplary embodiments, the dye is a fluorescence resonance energy transfer (FRET) dye; a xanthene dye, such as fluorescein and rhodamine; a dye that has an amino group in the alpha or beta position (such as a naphthylamine dye, 1-dimethyl-aminonaphthyl-5-sulfonate, 1-anilino-8-naphthalende sulfonate and 2-p-touidinyl-6-naphthalene sulfonate); a dye that has 3-phenyl-7-isocyanatocoumarin; an acridine, such as 9-isothiocyanatoacridine and acridine orange; a pyrene, a bensoxadiazole and a stilbene; a dye that has 3-(s-carboxypentyl)-3'-ethyl-5,5'-dimethyloxacarbocyanine (CYA); 6-carboxy fluorescein (FAM); 5&6-carboxyrhodamine-110 (R110); 6-carboxyrhodamine-6G (R6G); N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); 6-carboxy-X-rhodamine (ROX); 6-carboxy-4',5'-dichloro-2',7'-dimethoxyfluorescein (JOE); ALEXA Fluor™; Cy2; Texas Red and Rhodamine Red; 6-carboxy-2',4,7,7'-tetrachlorofluorescein (TET); 6-carboxy-2',4,4',5',7,7'-hexachlorofluorescein (HEX); 5-carboxy-2',4',5',7'-tetrachlorofluorescein (ZOE); NAN; NED; Cy3; Cy3.5; Cy5; Cy5.5; Cy7; and Cy7.5; Alexa Fluor 350; Alexa Fluor 488; Alexa Fluor 532; Alexa Fluor 546; Alexa Fluor 568; Alexa Fluor 594; or Alexa Fluor 647.

[0177] The label monomers can be incorporated into a nanoreporter at different stages of its assembly, or into a component (*e.g.*, a "flap" or of the nanoreporter prior to its assembly into the nanoreporter.

[0178] A label monomer can be directly attached to a nucleotide using methods well known in the art. Nucleotides can also be chemically modified or derivitized in order to attach a label monomer. For example, a fluorescent monomer such as a fluorescein molecule can be attached to dUTP (deoxyuridine-triphosphate) using a four-atom aminoalkynyl group. Each label monomer is attached to a nucleotide making a label monomer: nucleotide complex.

[0179] This label monomer: nucleotide complex can be incorporated into nucleic acids (for example, a DNA patch or a detection oligonucleotide) in a variety of ways. For example, a label monomer: nucleotide complex can be incorporated at only one location within a nucleic acid or at two or more locations within a nucleic acid.

[0180] Amine-reactive and thiol-reactive fluorophores are available and used for labeling nucleotides and biomolecules. Generally, nucleotides are fluorescently labeled during chemical synthesis, for example, incorporation of amines or thiols during nucleotide synthesis permit addition of fluorophores. Fluorescently labeled nucleotides are commercially available. For example, uridine and deoxyuridine triphosphates are available that are conjugated to ten different fluorophores that cover the spectrum.

[0181] A nucleotide can be attached to a label monomer first and then be incorporated into a nucleic acid. Alternatively, an existing nucleic acid can be labeled by attaching a label monomer to a nucleotide within the nucleic acid. For example aminoallyl- ("AA-") modified UTP nucleotides can be incorporated into the RNA product during transcription. In various embodiments, 20% or more of UTP nucleotides in a transcription reaction to generate RNA patches are AA modified. In various embodiments, about 20% to 100%, 20% to 80%, 30 to 80%, 40 to 60% or 50% to 75% of UTPs in a transcription reaction are AA-modified, in a preferred embodiment, approximately 50% of UTPs in a transcription reaction are AA-modified.

[0182] In addition, for example, different types of label monomer: nucleotide complexes can be incorporated into a single acid nucleic acid, where one component of the nanoreporter code comprises more than one type of signal.

[0183] Fluorescent dyes that can be bound directly to nucleotides can also be utilized as label monomers. For example, FAM, JOE, TAMRA, and ROX are amine reactive fluorescent dyes that have been attached to nucleotides and are used in automated DNA sequencing. These fluorecently labeled nucleotides, for example, ROX-ddATP, ROX-ddCTP, ROX-

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 219

ddGTP and ROX-ddUTP, are commercially available.

[0184] Other types of label monomers that may be used to label a nanoreporter are quantum dots. Due to their very small size the quantum dots can be coupled into oligonucleotides directly without affecting the solubility or use of the oligonucleotide. In a preferred embodiment, only one oligonucleotide molecule is coupled to each nanoparticle. To synthesize an oligonucleotide-nanoparticle complex in a 1:1 ratio by conventional batch chemistry, both the oligonucleotide and the nanoparticle require a single reactive group of different kinds that can be reacted with each other. For example, if an oligonucleotide has an amino group and a nanoparticle has an aldehyde group, these groups can react to form a Schiff base. An oligonucleotide can be derivitized to attach a single amino or other functional group using chemistry well known in the art. However, when a nanoparticle is derivatized, it is covered with a chemical reagent which results in coating the entire surface of the nanoparticle with several functional groups.

[0185] The invention provides a method of coupling one oligonucleotide to one nanoparticle by chemically coupling the oligonucleotide on a solid surface such as the glass support used for the oligonucleotide synthesis.

[0186] For example, commercially available resins for oligonucleotide synthesis such as long chain alkylamino controlled pore glass (lcaa CPG) can be used.

[0187] Alternatively, a flat surface such as a derivitized microscope slide can be used. The surface density of the nascent oligonucleotide chains should be lower than the diameter of the nanoparticle. This can be achieved by either choosing a glass support with low surface density of the reactive groups, or by using diluted reagent for the first step of the oligonucleotide synthesis so that the surface is not saturated. Another point of consideration when using the standard glass matrices for oligonucleotide synthesis is to use a pore diameter higher than the nanoparticle diameter to ensure the flow of the reagents. For example, an oligonucleotide can be synthesized on a diluted basis relative to the solid support, for example one tenth of a normal synthesis, to ensure good spacing of the oligonucleotides on the glass support. After the oligonucleotide is synthesized with a reactive functional group, for example, an amino group, derivitized nanoparticules are passed over the glass support to react with the oligonucleotides. A sufficiently large pore size of the glass support can be chosen to prevent slogging with nanoparticles. For example, a pore size of about 200 nm can be used. After the reaction is complete, un-reacted groups on the nanoparticle can be blocked and the complexes can be uncoupled from the glass support.

5.5 THE NANOREPORTERCODE

5.5.1 DUAL NANOREPORTERS

[0188] A nanoreporter whose components exist in two molecular entities is referred to as a dual nanoreporter (or probe pair). In a dual nanoreporter according to the present invention, each component contains a target- specific sequence, which improves the specificity of and binding kinetics of the nanoreporter to its target. The two different target-specific sequences are designed or selected such that each recognizes a different portion of a target molecule.

[0189] Figures 1A- 1C illustrates embodiments of the invention involving dual nanoreporters. In Figures 1A and 1B, each of the two component of the nanoreporter is labeled, such that the nanoreporter s spectral code is formed only when the two components of the nanoreporter come together upon binding of the dual nanoreporter to its target molecule. However, in a dual nanoreporter, it is not necessary that both components are labeled. For example, as depicted in Figure 1C, one component of a dual nanoreporter is labeled with the nanoreporter code; and the other component attached to an affinity tag (arrow) that is useful to immobilize the nanoreporter for stretching an visualization.

5.5.2 REGISTERS

[0190] The term "register" refers to a set of alternating (every other) label attachment regions. Registers are useful where it is desirable to label adjacent label attachment regions without a spacer region, and where the signal emanating from adjacent label attachment regions cannot be spatially resolved using the desired method of detection. Thus, the signals detected with use of a register is that form by the alternating, rather than adjacent, label attachment regions. Signals detected from a plurality of registers (*e.g.*, that together are all the label attachment regions) can be combined to form a nanoregister code. Generally when using registers, adjacent label attachment regions are labeled with spectrally distinguishable label monomers.

[0191] Examples of registers are depicted in Figures 3 and 5. For example, in Figure 3A-3B, there are 8 label attachment regions 1-8. Alternating label attachment regions 1, 3, 5 and 7 form one register, and label attachment regions 2, 4, 6 and 8 form another register. In Figure 3A, only one of the registers (1, 3, 5 and 7) is labeled; in Figure 3B, both registers are labeled.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 220

5.6 AFFINITY TAGS

[0192] A variety of affinity tags known in the art may be used to purify and/or immobilize nanoreporters.

[0193] Where an affinity tag is used to immobilize a nanoreporter for the purpose of detection or imaging, it may be referred to herein as an "anchor." In a preferred embodiment, a biotin anchor is attached to the nanoreporter, allowing immobilization of the nanoreporter on a streptavidin coated slide.

[0194] An affinity tag that can be used to used for attachment to beads or other matrixes for a variety of useful applications including but not limited to purification.

[0195] Non-limiting examples of suitable affinity tags are provided below. It should be understood that most affinity tags could serve dual purposes: both as anchors for immobilization of the nanoreporters and tags for purification of the nanoreporters (whether fully or only partially assembled) or their components.

[0196] In certain embodiments, the affinity tag is a protein monomer. Examples of protein monomers include, but are not limited to, the immunoglobulin constant regions (see Petty, 1996, Metal-chelate affinity chromatography, in Current Protocols in Molecular Biology, Vol. 2, Ed. Ausubel et al., Greene Publish. Assoc. & Wiley Interscience), glutathione S-transferase (GST; Smith, 1993, Methods Mol. Cell Bio. 4:220-229), the *E. coli* maltose binding protein (Guan et al., 1987, Gene 67:21-30), and various cellulose binding domains (U.S. Pat. Nos. 5,496,934; 5,202,247; 5,137,819; Tomme et al., 1994, Protein Eng. 7:117-123), etc. Other affinity tags are recognized by specific binding partners and thus facilitate isolation and immobilization by affinity binding to the binding partner, which can be immobilized onto a solid support. For example, the affinity tag can be an epitope, and the binding partner an antibody. Examples of such epitopes include, but are not limited to, the FLAG epitope, the myc epitope at amino acids 408-439, the influenza virus hemagglutinin (HA) epitope, or digoxigenin ("DIG"). In other embodiments, the affinity tag is a protein or amino acid sequence that is recognized by another protein or amino acid, for example the avidin/streptavidin and biotin.

[0197] In certain aspects of the invention, the affinity tag is a nucleotide sequence. A large variety of sequences of about 8 to about 30 bases, more preferably of about 10 to about 20 bases, can be used for purification and immobilization of nanoreporters, and the sequence can be tandemly repeated (*e.g.*, from 1 to 10 tandem repeats). Such a sequence is preferably not widely represented (that is, present in fewer than 5% of the genes, more preferably, present in fewer than 3% of the genes, and, most preferably, present in fewer than 1% of the genes) in the sample being assayed (for example, where the nanoreporter is used for detection of human cellular RNA, the sequence is preferably not widely represented in the human genome); have little or no secondary structure or self-complementarity either internally or with copies of itself when multimerized (that is, all secondary structures of the multimerized tag preferably have a Tm less than 25°C at 1 M NaCl); have no significant identity or complementarity with scaffold or segment sequences (that is, the Tm of complementary sequences is preferably less than 25°C at 0.2 M NaCl); and have a Tm of about 35-65°C, more preferably about 40-50°C, in 50 mM Na^+.

[0198] In certain embodiments, different sequences are used as purification and immobilization tags. In this case, for example, the purification tag can be as described above, but the immobilization tag can be in the range of 10 to 100 bases, with a Tm up to 95°C in 50 mM Na^+. An alternative embodiment would be to have the purification tag nested within the immobilization tag (*e.g.*, the affinity tag would comprise a 25-base sequence of which 15 bases are used as a purification tag and the entire 25 bases are used as the immobilization tag).

[0199] In certain instances, the affinity tag can be used for labeling a nanoreporter in addition to purifying or immobilizing the nanoreporter.

[0200] As will be appreciated by those skilled in the art, many methods can be used to obtain the coding region of the affinity tags, including but not limited to, DNA cloning, DNA amplification, and synthetic methods. Some of the affinity tags and reagents for their detection and isolation are available commercially.

5.7 TARGET-SPECIFIC SEQUENCES

[0201] The term target-specific sequence" refers to a molecular entity that is capable of binding a target molecule. In the context of a nanoreporter, the target-specific sequence is attached to the nanoreporter scaffold.

[0202] The target specific sequence is generally an amino acid sequence (*i.e.,* a polypeptide or peptide sequence) or a nucleic acid sequence.

[0203] In specific embodiments, where the target-specific sequence is an amino acid sequence, the target-specific sequence is an antibody fragment, such as an antibody Fab' fragment, a single chain Fv antibody.

[0204] The target-specific sequence is preferably a nucleic acid sequence, and is most preferably within an oligonucleotide that is either covalently attached (*e.g.*, by ligation) or noncovalently attached (*e.g.*, by hybridization) to the nanoreporter scaffold. A target-specific 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 approximately 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 approximately 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 221

[0205] A target-specific nucleotide sequence preferably has a Tm of about 65-90°C for each probe in 825 mM Na^+ (5X SSC), most preferably about 78-83°C.

[0206] In certain preferred embodiments, the target specific sequence of each probe of a dual nanoreporter is about 35 to 100 nucleotides (for a total target sequence of about 70 to 200 nucleotides, covered by 2 probes), most preferably about 40 to 50 nucleotides for each probe (for a total of about 80 to 100 nucleotides).

5.8 TARGET MOLECULES

[0207] The term "target molecule" is the molecule detected or measured by binding of a labeled nanoreporter whose target-specific sequence(s) recognize (are specific binding partners thereto). Preferably, an target molecule can be, but is not limited to, any of the following: DNA, cDNA, RNA, mRNA, peptide, a polypeptide/protein (*e.g.*, a bacterial or viral protein or an antibody), a lipid, a carbohydrate, a glycoprotein, a glycolipid, a small molecule, an organic monomer, or a drug. Generally, a target molecule is a naturally occurring molecule or a cDNA of a naturally occurring molecule or the complement of said cDNA.

[0208] 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.

[0209] In certain specific embodiments, that target molecule is not a chromosome. In other specific embodiments, the target molecule is no greater than 1,000 kb (or 1 mb) in size, no greater than 500 kb in size, no greater than 250 kb in size, no greater than 175 kb in size, no greater than 100 kb in size, no greater than 50 kb in size, no greater than 20 kb in size, or no greater than 10 kb in size. In yet other specific embodiments, the target molecule is isolated from its cellular milieu.

[0210] In specific, non-limiting embodiments, the target molecule is one of the following antibodies or an antigen recognized by one of the following antibodies: anti-estrogen receptor antibody, an anti-progesterone receptor antibody, an anti-p53 antibody, an anti-Her-2/neu antibody, an anti-EGFR antibody, an anti-cathepsin D antibody, an anti-Bcl-2 antibody, an anti-E-cadherin antibody, an anti-CA125 antibody, an anti-CA15-3 antibody, an anti-CA19-9 antibody, an anti-c-erbB-2 antibody, an anti-P-glycoprotein antibody, an anti-CEA antibody, an anti-retinoblastoma protein antibody, an anti-ras oncoprotein antibody, an anti-Lewis X antibody, an anti-Ki-67 antibody, an anti-PCNA antibody, an anti-CD3 antibody, an anti-CD4 antibody, an anti-CD5 antibody, an anti-CD7 antibody, an anti-CD8 antibody, an anti-CD9/p24 antibody, an anti-CD 10 antibody, an anti-CD11c antibody, an anti-CD13 antibody, an anti-CD14 antibody, an anti-CD 15 antibody, an anti-CD19 antibody, an anti-CD20 antibody, an anti-CD22 antibody, an anti-CD23 antibody, an anti-CD30 antibody, an anti-CD31 antibody, an anti-CD33 antibody, an anti-CD34 antibody, an anti-CD35 antibody, an anti-CD38 antibody, an anti-CD41 antibody, an anti-LCA/CD45 antibody, an anti-CD45RO antibody, an anti-CD45RA antibody, an anti-CD39 antibody, an anti-CD100 antibody, an anti-CD95/Fas antibody, an anti-CD99 antibody, an anti-CD106 antibody, an anti-ubiquitin antibody, an anti-CD71 antibody, an anti-c-myc antibody, an anti-cytokeratins antibody, an anti-vimentins antibody, an anti-HPV proteins antibody, an anti-kappa light chains antibody, an anti-lambda light chain antibody, an anti-melanosome antibody, an anti-prostate specific antigen antibody, an anti-S-100 antibody, an anti-tau antigen antibody, an anti-fibrin antibody, an anti-keratins antibody, an anti-Tn-antigen antibody receptor protein, a lymphokine, an enzyme, a hormone, a growth factor, or a nucleic acid binding protein, a ligand for a cell adhesion receptor; a ligand for a signal transduction receptor; a hormone; a molecule that binds to a death domain family molecule; an antigen; a viral particle, a viral coating protein or fragment thereof, a toxic polypeptide selected from the group consisting of: (a) ricin, (b) Pseudomonas exotoxin (PE); (c) bryodin; (d) gelonin; (e) α-sarcin; (f) aspergillin; (g) restrictocin; (h) angiogenin; (i) saporin; (j) abrin; (k) pokeweed antiviral protein (PAP); and (1) a functional fragment of any of (a)-(k); a cytokine, or a soluble cytokine selected from the group consisting of erythropoietin, interleukins, interferons, fibroblast growth factors, transforming growth factors, tumor necrosis factors, colony stimulating factors and epidermal growth factor, Class I MHC antigens, class II MHC antigens, internalizing cell-surface receptors and/or viral receptors.

[0211] In specific, non-limiting embodiments, the target molecule is an antigen such as alpha fetoprotein, alpha-1 antitrypsin, α-2 macroglobulin, adiponectin, apoliprotein-A-1, apoliprotein-CIII, apoliprotein-H, BDNF, β-2 microglobulin, C reactive protein, calcitonin, cancer antigen 19-9, cancer antigen 125, CEA, CD 40, CD 40 ligand, complement 3, GK-MB, EGF, ENA-78, endothelin-1, enrage, eotaxin, erythropoietin, Factor VII, FABP, ferritin, FGF-basic, fibrinogen, G-CSF, GST, GM-CSF, growth hormone, haptoglobin, ICAM-1, IFN-gamma, IgA, IgE, IGF-1, IgM, IL-1α, IL-1β, IL-1ra, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12 p40, IL-12 p70, IL-13, IL-15, IL-16, insulin, leptin, lipoprotein (a), lymphotactin, MCP-1, MDC, MIP-1α, MIP-1β, MMP-2, MMP-3, MMP-9, myeloperoxidase, myoglobin, PAI-1, PAP, PAPP-A, SGOT, SHBG, PSA (free), RANTES, serum amyloid P, stem cell factor, TBG, thrombopoietin, TIMP-1, tissue factor, TNF-α, TNF-β, TNF RII, TSH, VCAM-1, VEGF, or vWF.

[0212] In some embodiments, the target molecule is an autoimmune related molecule such as ASCA, β-2 glycoprotein,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 222

Clq, centromere Prot. B, collagen type 1, collagen type 2, collagen type 4, collagen type 6, Cyto P450, ds DNA, histone, histone H1, histone H2A, histone H2B, histone H3, histone H4, HSC-70, HSP-32, HSP-65, HSP-71, HSP-90α, HSP-90β, insulin, JO-I, mitochondrial, myeloperoxidase, pancreatic islet cells, PCNA, PM-I, PR3, ribosomal P, RNP-A, RNP-C5 RNP, Sel-70, Smith, SSA, SSB, T3, T4, thyroglobutin, tTG, (celiac disease), or thyroid microsomal.

[0213] In some embodiments, the target molecule is a component isolated from an infectious agent, such as Cholera Toxin, Cholera Toxinβ, Campylobacter jejuni, cytomegalovirus, Diptheria toxin, Epstein-Barr NA, Epstein-Barr EA, Epstein-Barr VCA, Heliobacter pylori, HBV core, HBV envelope, HBV surface (Ad), HBV surface (Ay), HCV core, HCV NS3, HCV NS4, HCV NS5, hepatitis A, hepatitis D, HEV or f2. 3KD, HEV or f2 6 KD, HEV orf 3KD, HIV-1 p24, HIV-1 gp41, HIV-I gp120, HPV, HSV-1/2, HSV-1 gD, HSV-2 gD, HTLV- 1/2, influenza A, influenza A H3N2, influenza B, Leishmania donorani, Lyme disease, mumps, M. pneumonia, M. tuberculosis, parainfluenza 1, parainfluenza 2, parainfluenza 3, polio virus, RSV. Rubella, Rubeola, Streptolysin O, Tetanus Toxin, T. pallidum 15kD, T. pallidum p47, T. cruzi, Toxoplasma, Varicella roster.

5.9 NANOREPORTER POPULATIONS

[0214] Disclosed herein are nanoreporter or nanoreporter label unit populations, for example nanoreporter or nanoreporter label unit libraries, that contain at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 75, at least 100, leash 200, at least 300, at least 400, at least 500, at least 750, or at least 1,000 unique nanoreporters or nanoreporter label units, respectively. As used herein, "unique" when used in reference to a nanoreporter or nanoreporter label units within a population is intended to mean a nanoreporter or label unit that has a code that distinguishes it from other nanoreporters or label units in the same population.

[0215] In particular we disclose nanoreporter populations with at least 5,000, at least 10,000, at least 20,000 or at least 50,000 unique nanoreporters or nancreporter label units.

[0216] The nanoreporters in a population of nanoreporters can be singular nanoreporters, dual nanoreporters, or a combination thereof. The nanoreporters can be labeled or unlabeled.

[0217] The size of a nanoreporter population and the nature of the target-specific sequences of the nanoreporters within it will depend on the intended use of the nanoreporter. Nanoreporter populations can be made in which the target-specific sequences correspond to markers of a given cell type, including a diseased cell type. Nanoreporters populations may be generated in which the target-specific sequences represent at least 0.1%, at least 0.25%, at least 0.5%, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 10%, at least 5%, as least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, or at least ; 70% of the different type of transcripts in a cell. Optionally, nanoreporters populations are generated in which the target-specific sequences represent at least 0.1%, at least 0.25%, at least 0.5%, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, or at least 70% of the different genes in a cell. Optionally, nanoreporter populations are generated in which at least some of the target-specific sequences represent rare transcripts in a cell or tissue. Such nanoreporter populations preferably represent at least 5 rare transcripts. In specific embodiments, such nanoreporter populations represent at least 10, at least 20, at least 30, at least 40 or at least 50 rare transcripts.

[0218] The cell or tissue may be a mammalian cell or tissue, and more Preferably is a human cell or tissue.

[0219] Optionally, the nanoreporter population is a diagnostic or prognostic nanoreporter populations. For example, a diagnostic nanoreporter population can be generated that is useful for screening blood products, in which the target-specific sequences bind to the nucleic acids of contaminating viruses such the hepatitis B, hepatitis C, and the human immunodeficiency virus. Alternatively, the diagnostic nanoreporter population may contain target-specific sequences corresponding to cellular disease markers, such as tumor antigens. Prognostic nanoreporter populations generally include target-specific markers that represent different stages of a given disease such as cancer. By selecting appropriate target- specific sequences, a nanoreporter population can be used both to diagnose and prognose disease.

5.10 BIOMOLECULAR SAMPLES

[0220] The probe pairs of the invention can be used to detect target molecule 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: cells (including both primary cells and cultured cell lines), cell lysates or extracts (including but not limited to RNA extracts; purified mRNA), tissues and tissue extracts (incuding but not limited to RNA extracts; purified mRNA); 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, natal and vaginal secretions, perspiration and semen, a translate, 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;

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 223

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.

[0221] The biomolecular samples can be indirectly derived from biological specimens. For example, where the target molecule of interest is a cellular transcript, e.g., a messenger RNA, the biomolecular sample of the invention can be a sample containing cDNA produced by a reverse transcription of messenger RNA. 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.

[0222] The biomolecular samples of the invention may be either "native," i.e., 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.

5.11 REPARATION OF LABEL MONOMERS

[0223] In addition to detecting an overall signal generated from a labeled nanoreporter, the invention provides for the determination of the spatial location of signals emanating from the label monomers (i.e., spots) on a nanoreporter, each spot representing the aggregate signal from label monomers attached to a given label attachment region. A spot may contain signals of the same wavelength or of different wavelengths. Thus, the nature of the spots on a nanoreporter and their location constitutes the nanoreporter code.

[0224] Any of a variety of means can be used to "stretch" the nanoreporter to separate the individual spots. For example, a nanoreporter can be stretched using a flowstretch technique (Henegariu et al, 2001, Biotechniques 31: 246-50), a receding meniscus technique (Yokota et al., 1997, Nuc. Acids Res. 25:1064-1070) or an electrostretching technique (Matsuura et al, 2001, Nuc. Acids Res. 29: E79).

[0225] The use of now-stretching, receding meniscus, or electro-stretching techniques allows for the separation of the label attachment regions within a nanoreporter so that one can determine spatially where a particular signal is positioned in the nanoreporter. Therefore, unique nanoreporters that have the same combination of label monomers and the same overall signal can be differentiated from one another based on the location of those label monomers within the nanoreporter.

[0226] This ability to locate the position of a label attachment region or spot within a nanoreporter allows for the position of the signal(s) emitted by the label monomers in each label attachment region to be used as a distinguishing characteristic when generating a set of unique nanoreporters. Hence, a complex set of nanoreporters can be generated using the same combination of starting label monomers by varying the positions of the label monomers within a nanoreporter.

[0227] Prior to stretching a nanoreporter, it is preferable to immobilize the nanoreporter to a solid surface using an affinity tag, as described in Section 5.6 above.

[0228] In certain aspects of the invention, one end of a nanoreporter is immobilized, either through specific or non-specific binding to a solid surface, the nanoreporter a stretched, and then the other end of the reporter is immobilized, also either through specific or non-specific binding to a solid surface. Accordingly, the nanoreporter is "frozen" in its stretched, or extended, state, to facilitate resolution of the nanoreporters code by detecting and/or imaging the signals emitted by the label monomers attached to a nanoreporter and their locations relative to one another. These aspects of the invention are described below in Section 5.12.

5.12 IMMOBILIZATION OF STRETCHED NANOREPORTERS

[0229] We disclose herein methods and compositions that facilitate the identification of primary structures of a variety of nanoreporters. In particular we disclose methods for the selective immobilization of nanoreporters in an extended state. Within the present invention, a nanoreporter can be selectively immobilized while fully extended under whatever force is used for the extension. In addition, the methods facilitate the selective immobilization of extended nanoreporters that are oriented with respect to each other. In other words, a plurality of nanoreporter can readily be immobilized in the same orientation with respect to each other.

[0230] Firstly we disclose methods for selectively immobilizing a nanoreporter in an extended state. For the methods, generally, a first portion of the nanoreporter is immobilized by any technique known to those of skill in the art. Indeed, the technique for immobilizing the first portion of the nanoreporter is not critical. In certain methods, the first portion of the nanoreporter can be immobilized selectively or non-selectively. In certain methods the first portion is immobilized by one or more covalent bonds. In certain methods, the first portion is immobilized by one or more non-covalent bonds. Exemplary immobilized first portions are described in the sections below.

[0231] With an immobilized first portion, the nanoreporter can be extended by any technique for extending a nanoreporter apparent to those of skill in the art. In certain methods, the technique for extending the nanoreporter is not critical for the methods of the invention. In certain methods, the technique for extending the nanoreporter appropriate for the

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 224

class of nanoreporter according to the judgment of one of skill in the art. In certain methods, the nanoreporter is extended by application of a force capable of extending the nanoreporter. The force can be any force apparent to one of skill in the art for extending the nanoreporter. Exemplary forces include gravity, hydrodynamic force, electromagnetic force and combinations thereof. Specific techniques for extending the nanoreporter are described in the sections below.

[0232] The nanoreporter is in an extended state if it would be recognized as extended by one of skill in the art. In certain methods, the nanoreporter is in an extended state when it is in the field of a force capable of extending the nanoreporter. In certain methods, the nanoreporter is in an extended state when its average hydrodynamic radius is more than double the average hydrodynamic radius of the nanoreporter in its native state as recognized by those of skill in the art.

[0233] The methods generally comprise the step of selectively immobilizing a second portion of the nanoreporter while it is in an extended state. This can result in an immobilized nanoreporter that is extended between the first and the second portion. Remarkably, since the nanoreporter is selectively immobilized white extended, that extension can be preserved in the immobilized nanoreporter. Generally, the first portion and the second portion of the nanoreporter are not the same.

[0234] The selective immobilization can be according to any technique for selective immobilization of a portion of a nanoreporter apparent to those of skill in the art. The selective immobilization can be through, for example, the formation of one or more covalent bonds or one or more non-covalent bonds, or both. Particular examples of selective immobilization techniques are described in the sections below. In particular methods, one or more binding pairs are used to immobilize the second portion of the nanoreporter.

[0235] The second portion can be immobilized onto any substrate apparent to those of skill in the art. The substrate can be any substrate judged to be useful for immobilization known to those of skill in the art. In certain methods, the second portion can be immobilized to another molecule. Further useful substrates include surfaces, membranes, beads, porous materials, electrodes, arrays and any other substrate apparent to those of skill in the art.

[0236] Secondly we disclose a composition comprising a selectively immobilized, extended nanoreporter. The compositions generally comprise a substrate and an extended nanoreporter selectively immobilized onto the substrate. The substrate can be any substrate known to those of skill in the art. Exemplary substrates include those described in the sections below. At least two portions of the nanoreporter are immobilized onto the substrate, and the nanoreporter is in an extended state between the two portions. In certain compositions, at least one portion of the nanoreporter is selectively immobilized onto the substrate. In certain compositions, two or more portions of the nanoreporter are selectively immobilized onto the substrate. The nanoreporter can be extended and/or immobilized by any technique apparent to those of skill, including particularly the methods disclosed herein.

[0237] Thirdly we disclose methods for selectively immobilizing a nanoreporter in an oriented state. The nanoreporter can be any nanoreporter described above. In certain methods, the nanoreporter can be flexible, or in certain embodiments the nanoreporter can be rigid or semirigid. For these methods, generally, a first portion of the nanoreporter is immobilized as described above. With an immobilized first portion, the nanoreporter can be oriented by any technique for extending a nanoreporter apparent to those of skill in the art. In certain methods, the technique for orienting the nanoreporter is not critical. In certain methods, the technique for orienting the nanoreporter appropriate for the class off nanoreporter according to the judgment of one of skill in the art. In certain methods, the nanoreporter is oriented by application of force capable of orienting the nanoreporter. The force can be any force apparent to one of skill in the art for orienting the nanoreporter. Exemplary forces include gravity, hydrodynamic force, electromagnetic force and combinations thereof. Specific techniques for extending the nanoreporter are described in the subsections below.

[0238] The nanoreporter is in an oriented state if it would be recognized as oriented by one of skill in the part. For example, the nanoreporter is in an oriented state when it is in the field of a force capable of orienting the nanoreporter. Alternatively, the nanoreporter is in an oriented state when its termini are arranged in parallel, as recognized by those of skill in the art, with the field of a force capable of orienting the nanoreporter. Additionally, a plurality of nanoreporters is in an oriented state when the termini of the nanoreporters are arranged in parallel, as recognized by those of skill in the art.

[0239] The methods generally comprise the step of selectively immobilizing a second portion of the nanoreporter while it is in an oriented state. This can result in an immobilized nanoreporter that is oriented between the first and the second portion. Remarkably, since the nanoreporter is selectively immobilized while extended, that orientation can be preserved in the immobilized nanoreporter. The selective immobilization can according to the methods described above.

[0240] Also disclosed herein is a composition comprising a selectively immobilized, oriented nanoreporter. The compositions generally comprise a substrate and an oriented nanoreporter selectively immobilized onto the substrate. The substrate can be any substrate known to those of skill in the art. Exemplary substrates include those described in the sections below. At least two portions of the nanoreporter are immobilized onto the substrate, and the nanoreporter is in an oriented state between the two portions. In certain compositions, at least, one portion of the nanoreporter is selectively immobilized onto the substrate. In certain compositions, both portions of the nanoreporter are selectively immobilized onto the substrate. The nanoreporter can be oriented and/or immobilized by any technique apparent to those of skill,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 225

including particularly the methods taught herein.

[0241] The methods and compositions can be used for any purpose apparent to those of skill in the art. For instance, the immobilized and extended and/or oriented nanoreporter can be used as a label for a substrate on which the nanoreporter is immobilized. The primary sequence of the immobilized and extended and/or oriented nanoreporter can be identified by any technique apparent to those of skill. Advantageously, immobilization of the extended and/or oriented nanoreporter can facilitate such techniques. Optionally, the immobilized and extended and/or oriented nanoreporter can be used to guide the manufacture of nanopaths, for example to create nanowires or nanocircuits. Further uses for the immobilized and extended and/or oriented nanoreporters are described in the sections below.

[0242] All terms used herein have their ordinary meanings to those of skill in the art unless indicated otherwise. The following terms shall have the following meanings.

[0243] As used herein, the term "binding pair" refers to first and second molecules or moieties that are capable of selectively binding to each other, i.e. binding to each other with greater affinity than to other components in a composition. The binding between the members of the binding pair can be covalent or non-covalent. In certain embodiments, the binding is noncovalent. Exemplary binding pairs included immunological binding pairs (e.g., any haptenic or antigenic compound in combination with a corresponding antibody or binding portion or fragment thereof, for example digoxigenin and anti-digoxigenin, fluorescein and anti-fluorescein, dinitrophenol and anti-dinitrophenol, bromodeoxyuridine and anti-bromodeoxyuridine, mouse immunoglobulin and goat anti-mouse immunoglobulin) and nonimmunological binding pairs (e.g., biotin-avidin, biotin-streptavidin, hormone- hormone binding protein, receptor-receptor ligand (e.g., acetylcholine receptor-acetylcholine or an analog thereof), IgG-protein A, lectin-carbohydrate, enzyme-enzyme cofactor, enzyme-enzyme inhibitor, complementary polynucleotide pairs capable of forming nucleic acid duplexes, and the like). For instance, immunoreactive binding members may include antigens, happens, aptamers, antibodies (primary or secondary), and complexes thereof, including those formed by recombinant DNA methods or peptide synthesis. An antibody may be a monoclonal or polyclonal antibody, a recombinant protein or a mixture(s) or fragment(s) thereof as well as a mixture of an antibody and other binding members. Other common binding pairs include but are not limited to, biotin and avidin (or derivatives thereof), biotin and streptavidin, carbohydrates and lectins, complementary nucleotides sequences (including probe and capture nucleic acid sequences), complementary peptide sequences including those formed by recombinant methods, effector and receptor molecules, hormone and hormone binding protein, enzyme cofactors and enzymes, enzyme inhibitors and enzymes, and so forth.

[0244] "Selective binding" refers to the any preferential binding of a pair of molecules or moieties for each other with respect to other molecules or moieties in a composition that would be recognized by one of skill in the art. Optionally, a pair of molecules or moieties selectively binds when they preferentially bind each other compared to other molecules or moieties. Selective binding can include affinity or avidity, or both, of one molecule or moiety for another molecule or moiety. Optionally, selective binding requires a dissociation constant (K_D) of less than about $1x10^{-5}$ M or less than about $1x10^{-6}$ M, $1x10^{-7}$ M, $1x10^{-8}$ M, $1x10^{-9}$ M, or $1x10^{-10}$ M. In contrast, non-selective binding optionally has significantly less affinity, for example, a KD greater than $1x10^{-3}$ M.

[0245] "Extended state" refers to a nanoreporter in a state that would be recognized as extended by one of skill in the art. A nanoreporter may be in an extended state when it is extended relative to its native conformation in solution. A nanoreporter may be in an extended state when it is in the field of a force capable of extending the nanoreporter. An extended state of a nanoreporter can be determined quantitatively, such that, those of skill in the art will recognize R as the end-to-end vector of the nanoreporter, i.e. the distance between two termini of the nanoreporter, and <R> as the average end-to-end vector such that 95% of R will be within 2<R> in a solution deemed appropriate to one of skill in the art. Exemplary solutions include, for example, a dilute solution of the nanoreporter in water or in a pH buffer. Optionally, a nanoreporter is in an extended state when R is greater than 2.0<R>.

[0246] "Oriented state" refers to a nanoreporter in a state that would be recognized as oriented by one of skill in the art. A nanoreporter may be in an oriented state when it is oriented relative to its native conformation in solution. The nanoreporter may be oriented when it is arranged in parallel with the field of a force capable of orienting the nanoreporter. The nanoreporter may be oriented when it is one of a plurality of nanoreporters that are arranged in parallel, as recognized by those of skill in the art.

5.12.1 METHODS OF SELECTIVE IMMOBILIZATION

[0247] As described above, we disclose methods for the selective immobilization of a nanoreporter in an extended state. The nanoreporter, once selectively immobilized, can be used for any purpose apparent to those of skill in the art.

[0248] The nanoreporter may be any polymer known to those of skill in the art. For instance, the nanoreporter can be a polysaccharide, a polypeptide or a polynucleotide. Useful polynucleotides include ribonucleic acids, deoxyribonucleic acids and other polynucleotides known to those of Skill in the art.

[0249] The nanoreporter can be of any size that is sufficient to allow extension and immobilization of the nanoreporter according to the methods of the invention. In certain embodiments when the nanoreporter is a polynucleotide, the

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 226

nanoreporter can have a length of greater than 500 bp, greater than 750 bp, greater than 1 kb, greater than 1.5 kb, greater than 2..0 kb, greater than 2.5 kb, greater than 3.0 kb, greater than 4.0 kb or greater than 5.0 kb. In certain embodiments, when the nanoreporter is a polypeptide, the nanoreporter can have a size of greater than 50 amino acids, greater than 100 amino acids, greater than 200 amino acids, greater than 300 amino acids, greater than 400 amino acids, greater than 500 amino acids, greater than 750 amino acids, greater than 1000 amino acids, greater than 1500 amino acids, greater than 2000 amino acids, greater than 2500 amino acids, greater than 3000 amino acids, greater than 4000 amino acids or greater than 5000 amino acids. In certain embodiments, when the nanoreporter is a polysaccharide, the nanoreporter can have a size of greater than 50 saccharides, greater than 100 saccharides, greater than 200 saccharides, greater than 300 saccharides, greater than 400 saccharides, greater than 500 saccharides, greater than 750 saccharides, greater than 1000 saccharides, greater than 1500 saccharides, greater than 2000 saccharides, greater than 2500 saccharides, greater than 3000 saccharides, greater than 4000 saccharides or greater than 5000 saccharides.

[0250] The nanoreporter can be a native nanoreporter as understood by those of skill in the art, or the nanoreporter can be a non-native nanoreporter. When the nanoreporter is a polypeptide, the nanoreporter can comprise only naturally occurring amino acids, or the nanoreporter can comprise naturally occurring amino acids and non-naturally occurring amino acids. The other amino acids can be any amino acids, or derivatives or analogs thereof, known to those of skill in the art. When the nanoreporter is a polynucleotide, the polynucleotide can comprise only naturally occurring nucleotides, or the polynucleotide can comprise naturally occurring nucleotides and non-naturally occurring nucleotides. When the nanoreporter is a polysaccharide, the polysaccharide can comprise only naturally occurring saccharides, or the polysaccharide can comprise naturally occurring saccharides and non-naturally occurring saccharides. The polymers can comprise only non-natural monomers. The nanoreporter can comprise a plurality of classes of monomers, such as amino acids, nucleotides and/or saccharides.

[0251] Optionally, the nanoreporter comprises only one primary, covalently linked chain of monomers. For instance, when the nanoreporter is a polypeptide, in certain embodiments, the nanoreporter comprises only one primary amino acid chain. When the nanoreporter is a polynucleotide, the nanoreporter may be single stranded. The nanoreporter may comprise two primary, covalently linked chains of monomers. For instance, when the nanoreporter is a polypeptide, the nanoreporter may comprise two primary amino acid chains. When the nanoreporter is a polynucleotide, the nanoreporter may comprise two polynucleotide strands; the nanoreporter can be double stranded, in part or in whole. Further, the nanoreporter may comprise three or more primary, covalently linked chains of monomers. For instance, when the nanoreporter is a polypeptide, the nanoreporter optionally comprises three primary amino acid chains. When the nanoreporter is a polynucleotide, the nanoreporter optionally comprises three polynucleotide strands. For instance, the nanoreporter can comprise three strands F1. X and F2 where a portion of strand X is complementary to strand F1 and a portion of strand X is complementary to strand F2. An example is illustrated in Figure 13A. Optionally, the nanoreporter comprises more than three primary, covalently linked chains of monomers.

[0252] Advantageously, a nanoreporter can comprise one or more label that facilitate the detection, imaging or identification of the nanoreporter by techniques known to those of skill in the art. The label can be any detectable moiety known to those of skill in the art. Exemplary labels for nanoreporters include detectable isotopes, radioisotopes, fluors, dyes, enzymes, ligands, receptors, antigens, antibodies, lectins, carbohydrates, nucleotide sequences, and any other detectable label apparent to those of skill in the art.

[0253] A polynucleotide is otionally a polymer of natural (e.g. A, G, C, T, U) or synthetic nucleobases, or a combination of both. The backbone of the polynucleotide can be composed entirely of "native" phosphodiester linkages, or it may contain one or more modified linkages, such as one or more phosphorothioate, phosphorodithioate, phosphoramidate or other modified linkages. As a specific example, a polynucleotide may be a peptide nucleic acid (PNA), which contains amide interlinkages. Additional examples of synthetic bases and backbones that can be used in conjunction with the invention, as well as methods for their synthesis can be found, for example, in U.S. Patent No. 6,001,983; Uhlman & Peyman, 1990, Chemical Review 90(4):544 584; Goodchild, 1990, Bioconjugate Chem. 1(3):165 186; Egholm et al., 1992, J. Am. Chem. Soc. 114:1895 1897; Gryaznov et al., J. Am. Chem. Soc. 116:3143 3144, as well as the references cited in all of the above. Common synthetic nucleobases of which polynucleotides may be composed include 3-methlyuracil, 5,6-dihydrouracil, 4 thiouracil, 5 bromouracil, 5-thorouracil, 5-iodouracil, 6-dimethyl aminopurine, 6-methyl aminopurine, 2-aminopurine, 2,6-diamino purine, 6-amino-8-bromopurine, inosine, 5-methylcytosine, 7-deazaadenine, and 7-deazaguanosine. Additional non-limiting examples of synthetic nucleobases of which the target nucleic acid may be composed can be found in Fasman, CRC Practical Handbook of Biochemistry and Molecular Biology, 1985, pp. 385-392; Beilstein's Handbuch der Organischen Chemie, Springer Verlag, Berlin and Chemical Abstracts, all of which provide references to publications describing the structures, properties and preparation of such nucleobases.

[0254] The nanoreporter can be prepared according to any technique apparent to those of skill in the art. Advantageously, nanoreporters according to the invention can comprise labels and/or members of binding pairs, as described in the sections below, that can be used to facilitate preparation and/or purification of the nanoreporter. In addition, certain nanoreporters of the invention are capable of forming complexes with molecules that comprise members of binding

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 227

pairs, as described below. These complexes can be used to facilitate preparation and/or purification of the nanoreporter or complex.

5.12.2 IMMOBILIZATION OF FIRST PORTION

[0255] In the methods of the invention, a first portion of the nanoreporter is immobilized. Generally, the first portion is immobilized if it would be recognized as immobilized by one of skill in the art. The first portion can be immobilized by any technique apparent to those of skill in the art. In certain embodiments, the technique for immobilization of the first portion of the nanoreporter is not critical for the methods of the invention.

[0256] The first portion of the nanoreporter can be at any location in the nanoreporter. In certain embodiments, the first portion is at a terminus of the nanoreporter. For the purposes of the invention and this disclosure in general, a portion of a nanoreporter can be "at a terminus" when it is less than five, four, three, two, one or zero monomers from a terminus of the nanoreporter. Of course, although many nanoreporters have two termini, the methods disclosed herein are applicable to nanoreporters have more than two termini and to nanoreporters having one or zero termini, e.g. circular nanoreporters. The first portion is not necessarily at a terminus of the nanoreporter.

[0257] The nanoreporter can be immobilized onto any substrate apparent to those of skill in the art. The substrate can be any moiety to which the nanoreporter can be immobilized without limitation. Optionally, the substrate is a surface, membrane, bead, porous material, electrode or array.

[0258] In some examples, the first portion of the nanoreporter can be immobilized non-selectively. In further examples, the first portion of the nanoreporter can be immobilized selectively. In advantageous examples, after the first portion of the nanoreporter is immobilized, some portion of the nanoreporter should be free to move sufficiently so that the nanoreporter can be extended in the following steps of the method. In particular, in certain embodiments, when the first portion of the nanoreporter is immobilized non-selectively, it is important that the entire nanoreporter not be immobilized non-selectively to an extent that prevents extension of any portion of the nanoreporter.

[0259] The immobilization can be by any interaction with the substrate apparent to those of skill in the art. The immobilization can be via electrostatic or ionic interaction, via one or more covalent bonds, via one or more non-covalent bonds or combinations thereof. In certain cases, the immobilization can be via electrostatic interaction with an electrode. In further cases, the immobilization is via electrostatic interaction with a substrate other than the electrode.

[0260] Optionally, the first portion of the nanoreporter comprises a first member of a binding pair. The first member of the binding pair can be covalently bound to the first portion of the nanoreporter, or they can be non-covalently bound. Useful covalent bonds and non-covalent bonds will be apparent to those of skill in the art. Usefully, the substrate onto which the first portion of the nanoreporter is bound will comprise a second member of the binding pair. The substrate can be covalently bound to the second member, or they can be non-covalently bound. Figure 12A illustrates a nanoreporter that comprises a moiety F1 that is capable of selectively binding a moiety of the substrate. Moiety F1 can be, for example, biotin, capable of binding, for example, a substrate coated with avidin.

[0261] Optionally, the first portion of the nanoreporter can comprise a member of a binding pair that is capable of binding with a member of a binding pair on the substrate to form one or more non-covalent bonds. Exemplary useful substrates include those that comprise a binding moiety selected from the group consisting of ligands, antigens, carbohydrates, nucleic acids, receptors, lectins, and antibodies. The first portion of the nanoreporter would comprise a binding moiety capable of binding with the binding moiety of the substrate. Exemplary useful substrates comprising reactive moieties include, but are not limited to, surfaces comprising epoxy, aldehyde, gold, hydrazide, sulfhydryl, NHS-ester, amine, thiol, carboxylate, maleimide, hydroxymethyl phosphine, imidoester, isocyanate, hydroxyl, pentafluorophenyl-ester, psoralen, pyridyl disulfide or vinyl sulfone, or mixtures thereof. Such surfaces can be obtained from commercial sources or prepared according to standard techniques.

[0262] Advantageously, the first portion of the nanoreporter can be immobilized to the substrate via an avidin-biotin binding pair. Optionally, the nanoreporter can comprise a biotin moiety in its first portion. For instance, a polynucleotide nanoreporter can comprise a biotinylated nucleotide residue. Similarly, a polypeptide nanorsporter can comprise a biotinylated amino acid residue. The substrate comprising avidin can be any substrate comprising avidin known to those of skill in the art. Useful substrates comprising avidin are commercially available including TB0200 (Accelr8), SAD6, SAD20, SADIOO, SAD500, SAD2000 (Xantec), SuperAvidin (Array-It), streptavidin slide (catalog #MPC 000, Xenopore) and STREPT AVIDINnslide (catalog #439003, Greiner Bio-one).

[0263] The first portion of the nanoreporter may comprise a nucleotide sequence that is capable of selectively binding a nucleotide sequence on the substrate.

[0264] The first portion of the nanoreporter can comprise avidin, and the substrate can comprise biotin. Useful substrates comprising biotin are commercially available including Optiarray-biotin (Accler8), BD6, BD20, BDIOO, BD.500 and BD2000 (Xantec).

[0265] The first portion of the nanoreporter may be capable of forming a complex with one or more other molecules that, in turn, are capable of binding, covalently or non-covalently, a binding moiety of the substrate. For instance, a first

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 228

portion of the nanoreporter can be capable of selectively binding another molecule that comprises, for instance, a biotin moiety that is capable of selectively binding, for instance, an avidin moiety of the substrate. Figure 13A illustrates a nanoreporter that is capable of selectively binding a second molecule X that is capable of selectively binding a third molecule that comprises F1. F1 is capable of selectively binding a moiety on a substrate. Figure 13B illustrates a nanoreporter that is capable of selectively binding a second molecule that comprises F1, and F1 is capable of selectively binding a moiety on a substrate.

[0266] The first portion of the nanoreporter can comprise a member of a binding pair that is capable of reacting with a member of binding pair on the substrate to form one or more covalent bonds. Exemplary useful substrates comprising reactive groups include those that comprise a reactive moiety selected from the group consisting of succinamides, amines, aldehydes, epoxies and thiols. The first portion of the nanoreporter would comprise a reactive moiety capable of reacting with the reactive moiety of the substrate. Exemplary useful substrates comprising reactive moieties include, but are not limited to, OptArray-DNA NHS group (Accler8), Nexterion Slide AL (Schott) and Nexterion Slide E (Schott).

[0267] The first portion of the nanoreporter can comprise a reactive moiety that is capable of being bound to the substrate by photoactivation. The substrate could comprise the photoreactive moiety, or the first portion of the nanoreporter could comprise the photoreactive moiety. Some examples of photoreactive moieties include aryl azides, such as N-((2-pyridyldithio)ethyl)-4-azidosalicylamide; fluorinated aryl azides, such as 4-azido-2,3,5,6-tetrafluorobenzoic acid; benzophenone-based reagents, such as the succinimidyl ester of 4-benzoylbenzoic acid; and 5-Bromo-deoxyuridine.

[0268] The first portion of the nanoreporter can be immobilized to the substrate via other binding pairs apparent to those of skill in the art.

5.12.3 EXTENSION OF THE NANOREPORTER

[0269] In certain methods of the invention, the nanoreporters comprising the probe pair are in an extended state. Generally, any nanoreporter is in an extended state if it would be recognized as such by one of skill in the art.

[0270] In certain embodiments, the nanoreporter is in an extended state when it is in the field of a force capable of extending the nanoreporter under conditions suitable for extending the nanoreporter. Such forces and conditions should be apparent to those of skill in the art. For instance, many nanoreporters can be extended by hydrodynamic force or by gravity, and many charged nanoreporters can be extended by electromagnetic force. In certain embodiments, the force can be applied to the nanoreporter indirectly. For instance, the nanoreporter can comprise or can be linked, covalently or noncovalently, to a moiety capable of being moved by a force. In certain embodiments, the nanoreporter can be linked to a moiety

[0271] In certain embodiments, the force is an electromagnetic force. For instance, when the nanoreporter is charged, such as a polynucleotide, the nanoreporter can be extended in an electric or magnetic field. The field should be strong enough to extend the nanoreporter according to the judgment of one of skill in the art. Exemplary techniques for extending a nanoreporter in an electric or magnetic field are described in Matsuura et al., 2002, JBiomol Struct Dyn. 20(3):429-36; Ferree & Blanch, 2003, Biophys J. 85(4):2539-46; Stigter & Bustamante, 1998, Biophys J. 1998 75(3):1197-210; Matsuura et al., 2001, Nucleic Acids Res. 29(16); Ferree & Blanch, 2004, Biophys J. 87(1):468-75; the contents of which are hereby incorporated by reference in their entirety.

[0272] In certain embodiments, the force is a hydrodynamic force. For instance, many nanoreporters, including polysaccharides, polypeptides, and polynucleotides, can be extended in the field of a moving fluid. The hydrodynamic force should be strong enough to extend the nanoreporter according to the judgment of one of skill in the art. Exemplary techniques for extending a nanoreporter in hydrodynamic field are described in Bensimon et al., 1994, Science 265: 2096-2098; Henegariu et al., 2001, BioTechniques 31: 246-250; Kraus et al., 1997, Human Genetics 99:374 - 380; Michalet et al., 1997, Science 277:1518-1523; Yokota et al., 1997, Nucleic Acids Res. 25(5):1064-70; Otobe et al., 2001, Nucleic Acids Research 29:109; Zimmerman & Cox, 1994, Nucleic Acids Res. 22(3):492-7, and U.S. Patent Nos. 6,548,255; 6,344,319; 6,303,296; 6,265,153; 6,225,055; 6,054,327; and 5,840,862, the contents of which are hereby incorporated by reference in their entirety.

[0273] In certain embodiments, the force is gravity. In advantageous embodiments, the force of gravity can be combined with, for example, hydrodynamic force to extend the nanoreporter. In certain embodiments, The force should be strong enough to extend the nanoreporter according to the judgment of one of skill in the art. Exemplary techniques for extending a nanoreporter with gravity are described in Michalet et al, 1997, Science 277:1518-1523; Yokota et al., 1997', Nucleic Acids Res. 25(5); 1064-70; Kraus et al., 1997, Human Genetics 99:374 - 380.

[0274] In particular embodiments, the force is applied through a moving meniscus. Those of skill in the art will recognize that a moving meniscus can apply various forces to a nanoreporter including hydrodynamic force, surface tension and any other force recognized by those of skill in the art. The meniscus can be moved by any technique apparent to those of skill in the art including evaporation and gravity. Exemplary techniques for extending a nanoreporter with a moving meniscus are described in, for example, U.S. Patent Nos. 6,548,255; 6,344,319; 6,303,296; 6,265,153; 6,225,055; 6,054,327; and 5,840,862.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 229

[0275] In particular, embodiments, the nanoreporter can be extended by an optical trap or optical tweezers. For instance, the nanoreporter can comprise or can be linked, covalently or noncovalently, to a particle capable of being trapped or moved by an appropriate source of optical force. Useful techniques for moving articles with optical traps or optical tweezers are described in Ashkin et ah, 1986, Optics Letters 11 :288-290; Ashkin et al, 1987, Science 235: 1517-1520; Ashkin et al, Nature 330:769-771; Perkins et al, 1994, Science 264:822-826; Simmons et al, 1996, Biophysical Journal 70:1813-1S22; Block et al., 1990, Nature 348:348 - 352; and Grier, 2003, Nature 424: 810-816.

[0276] In certain embodiments, the nanoreporter can be extended by combinations of the above forces that are apparent to those of skill in the art. In the examples, below, certain nanoreporters are extended by a combination of an electric field and hydrodynamic force.

[0277] The nanoreporter is extended when it would be recognized as extended by one of skill in the art according to standard criteria for extension of a nanoreporter. In certain embodiments, the nanoreporter is extended when it loses most of its tertiary structural features as recognized by those of skill in the art. In certain embodiments, the nanoreporter is extended when it loses most of its secondary structural features as recognized by those of skill in the art. In certain embodiments, the nanoreporter is extended when its primary structural features are detectable in sequence when imaged according to standard techniques. Exemplary imaging techniques are described in the examples below.

[0278] In certain embodiments, an extended state of a nanoreporter can be recognized by comparing its hydrodynamic radius to its average hydrodynamic radius when free in dilute solution. For instance, in certain embodiments, a nanoreporter, or portion thereof, is extended when its hydrodynamic radius is more than about double its average hydrodynamic radius in dilute solution. More quantitatively, R represents the hydrodynamic radius of the nanoreporter, or portion thereof, and <R> represents the average hydrodynamic radius of the nanoreporter, or portion thereof, in dilute solution. The average <R> should be calculated such that R for the nanoreporter, or portion thereof, when unbound in dilute solution is less than 2<R> 95% of the time. In certain embodiments, a nanoreporter, or portion thereof, is in an extended state when R is greater than 1.5<R>, greater than 1.6<R>, greater than 1.7 <R>, greater than 1.8<R>, greater than 1.9<R>, greater than 2.0<R>, greater than 2.1<R>, greater than 2.2<R>, greater than 2.3<R>, greater than 2.4<R>, greater than 2.5<R> or greater than 3.0<R>. In particular embodiments, a nanoreporter, or portion thereof, is in an extended state when R is greater than 2.0<R>.

5.12.4 ORIENTATION OF THE NANOREPORTER

[0279] Optionally, the nanoreporter is in an oriented state. Generally, any nanoreporter is in an oriented state if it would be recognized as such by one ofskill in the art.

[0280] For example, the nanoreporter is in an oriented state when it is in the field of a force capable of orienting the nanoreporter under conditions suitable for orienting the nanoreporter. Such forces and conditions should be apparent to those of skill in the art.

[0281] The force may be an electromagnetic force. For instance, when the nanoreporter is charged, such as a polynucleotide, the nanoreporter can be oriented in an electric or magnetic field. The field should be strong enough to orient the nanoreporter according to the judgment of one of skill in the art. Exemplary techniques for orienting a nanoreporter in an electric or magnetic field are described above.

[0282] Alternatively, the force is a hydrodynamic force. For instance, many nanoreporters, including polysaccharides, polypeptide, and polynucleotides, can be oriented in the field of a moving fluid. The hydrodynamic force should be strong enough to orient the nanoreporter according to the judgment of one of skill in the art. Exemplary techniques for orienting a nanoreporter in hydrodynamic field are described above.

[0283] Alternatively, the force is gravity. Advantaegously, the force of gravity can be combined with, for example, hydrodynamic force to orient the nanoreporter. In certain examples, the force should be strong enough to orient the nanoreporter according to the judgment of one of skill in the art. Exemplary techniques for orienting a nanoreporter with gravity are described above.

[0284] Optionally, the nanoreporter can be oriented by combinations of the above forces that are apparent to those of skill in the art. In the examples, below, certain nanoreporters are oriented by a combination of an electric field and hydrodynamic force.

[0285] The nanoreporter is oriented when it would be recognized as oriented by one of skill in the art according to standard criteria for orientation of a nanoreporter. For example, the nanoreporter is oriented when it is arranged in parallel, as recognized by those of skill in the art, with the field of a force capable of orienting the nanoreporter. In certain examples, the nanorsporter is oriented when it is one of a plurality of nanoreporters that are arranged in parallel, as recognized by those of skill in the art. For instance, a plurality of nanoreporters can be oriented when the vector from a first terminus to a second terminus of a nanoreporter is parallel, as recognized by those of skill in the art, to the vectors between corresponding termini of other nanoreporters in the plurality.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 230

5.12.5 SELECTIVE IMMOBILIZATION OF SECOND PORTION OF NANOREPORTER

[0286] As discussed above, in the methods disclosed herein, a second portion of the nanoreporter is selectively immobilized. The second portion of the nanoreporter can be any portion of the nanoreporter that is not identical to the first portion of the nanoreporter. Optionally, the second portion of the nanoreporter does not overlap any part of the first portion of the nanoreporter.

[0287] We disclose methods that comprise the single step of selectively immobilizing a second portion of a nanoreporter while the nanoreporter is in an extended or oriented state, and while a first portion of the nanoreporter is immobilized. Exemplary methods for immobilization of the first portion of the nanoreporter, and for extension or orientation of the nanoreporter are described in detail in the sections above.

[0288] We also discloses methods that comprise the step of extending a nanoreporter, while a first portion of the nanoreporter is immobilized, and the step of selectively immobilizing a second portion of a nanoreporter while the nanoreporter is in an extended state. Exemplary methods for immobilization of the first portion of the nanoreporter, and for extension of the nanoreporter are described in detail in the sections above.

[0289] We also disclose methods that comprise the step of immobilizing a first portion of a nanoreporter, the step of extending the nanoreporter while the first portion is immobilized and the step of selectively immobilizing a second portion of a nanoreporter while the nanoreporter is in an extended state. Exemplary methods for immobilization of the first portion of the nanoreporter, and for extension of the nanoreporter are described in detail above. In certain embodiments, the present invention provides methods that comprise the step of orienting a nanoreporter, while a first portion of the nanoreporter is immobilized, and the step of selectively immobilizing a second portion of a nanoreporter while the nanoreporter is in an oriented state. Exemplary methods for immobilization of the first portion of the nanoreporter, and for orienting the nanoreporter are described in detail in the sections above.

[0290] We also disclose methods that comprise the step of immobilizing a first portion of a nanoreporter, the step of orienting the nanoreporter while the first portion is immobilized and the step of selectively immobilizing a second portion of a nanoreporter while the nanoreporter is in an oriented state. Exemplary methods for immobilization of the first portion of the nanoreporter, and for orienting the nanoreporter are described in detail above.

[0291] The selective immobilization of the second portion of the nanoreporter can follow any technique for selective immobilization of a nanoreporter apparent to those of skill in the art. Significantly, in advantageous embodiments of the invention, the second portion of the nanoreporter is not immobilized non-selectively. Selective immobilization can allow the nanoreporter to be immobilized while in a fully extended state or nearly fully extended state. Selective immobilization can also allow the nanoreporter to be immobilized in an oriented manner. In other words, the first portion and second portion of the nanoreporter can be immobilized along the direction of the field or fields used to extend the nanoreporter, with the first portion preceding the second portion in the field. When a plurality of nanoreporters are immobilized, the can be uniformly oriented along the field.

[0292] The second portion of the nanoreporter can be at any location in the nanoreporter. Optionally, the second portion is at a terminus of the nanoreporter. Optionally, the second portion is not at a terminus of the nanoreporter. Optionally, the first portion, described in the sections above, is at one terminus of the nanoreporter, and the second portion is at another terminus of the nanoreporter.

[0293] As discussed above, the second portion of the nanoreporter is immobilized selectively. The immobilization can be by any selective interaction with the substrate apparent to those of skill in the art. The immobilization can be via electrostatic or ionic interaction, via one or more covalent bonds, via one or more non-covalent bonds or combinations thereof. The immobilization can be via electrostatic interaction with an electrode. The immobilization may be via electrostatic interaction with a substrate other than the electrode.

[0294] If the first portion and the second portion of the nanoreporter are selectively immobilized to the same substrate, the techniques of selective immobilization should of course be compatible with the substrate. In particular, the techniques of immobilization are the same. For instance, on a substrate coated with avidin, both the first and second portion of the nanoreporter can be immobilized selectively via biotin - avidin interactions. However, as will be apparent to those of skill in the art, the same interaction need not be used at both the first and second portions for immobilization on the same substrate. For instance, the substrate can comprise multiple moieties capable of selective binding, or the first portion can be immobilized non-selectively, or other techniques apparent to those of skill in the art.

[0295] Optionally, the second portion of the nanoreporter comprises a first member of a binding pair. The second member of the binding pair can be covalently bound to the second portion of the nanoreporter, or they can be non-covalently bound. Useful covalent bonds and non-covaient bonds will be apparent to those of skill in the art. Usefully, the substrate onto which the second portion of the nanoreporter is bound will comprise a second member of the binding pair. The substrate can be covalently bound to the second member, or they can be non-covatently bound.

[0296] The second portion of the nanoreporter can comprise a member of a binding pair that is capable of binding with a member of a binding pair on the substrate to form one or more non-covalent bonds. Exemplary useful substrates include those that comprise a binding moiety selected from the group consisting of ligands, antigens, carbohydrates,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 231

nucleic acids, receptors, lectins, and antibodies such as those described in the sections above.

[0297] Advantageously, the second portion of the nanoreporter can be immobilized to the substrate via an avidin-biotin binding pair. The nanoreporter can comprise a biotin moiety in its first portion. For instance, a polynucleotide nanoreporter can comprise a biotinylated nucleotide residue. Similarly, a polypeptide nanoreporter can comprise a biotinylated amino acid residue. Useful substrates comprising avidin are described in the sections above.

[0298] Optionally, the second portion of the nanoreporter can comprise avidin, and the substrate can comprise biotin. Useful substrates comprising biotin are described in the sections above.

[0299] Optionally, the second portion of the nanoreporter can comprise a member of a binding pair that is capable of reacting with a member of a binding pair on the substrate to form one or more covalent bonds. Exemplary useful substrates comprising reactive groups are described in the sections above.

[0300] The second portion of the nanoreporter can comprise a reactive moiety that is capable of being bound to the substrate by photoactivation. The substrate could comprise the photoreactive moiety, or the second portion of the nanoreporter could comprise the photoreactive moiety. Some examples of photoreactive moieties include aryl azides, such as *N*-((2-pyridyldithio)ethyl)-4-azidosalicylamide; fluorinated aryl azides, such as 4-azido-2,3,5,6-tetrafiuorobenzoic acid; benzophenone-based reagents, such as the succinimidyl ester of 4-benzoylbenzoic acid; and 5-Bromo-deoxyuridine.

[0301] Alternatively, the second portion of the nanoreporter can be immobilized to the substrate via other binding pairs described in the sections above.

[0302] Alternatively, the second portion of the nanoreporter is capable of forming a complex with one or more other molecules that, in turn, are capable of binding, covalently or non-covalently, a binding moiety of the substrate. For instance, the second portion of the nanoreporter can be capable of selectively binding another molecule that comprises, for instance, a biotin moiety that is capable of selectively binding, for instance, an avidin moiety of the substrate. Figure 12B illustrates a nanoreporter of selectively binding a second molecule that comprises F3 that is, in turn, capable of selectively binding a moiety on a substrate. The interaction between the second portion of the nanoreporter and the molecule that comprises F3 can be mediated, for example, by an antigen-antibody interaction.

[0303] FIGS. 14A and 14B illustrate the selective immobilization of a nanoreporter according to methods of the present invention. In Figure 14 A, a first portion of the nanoreporter comprises binding moiety F1 that is capable of selectively binding a moiety on the illustrated substrate S. Binding moiety F1 can be, for instance, biotin, and substrate S can be coated with, for instance, avidin. The nanoreporter of Figure 14A is extended by a force as described in the sections above. In Figure 14B, the force is an electrical potential. While extended, the nanoreporter is contacted with molecules comprising binding moiety F2 that is capable of selectively binding a moiety on the illustrate substrate S. Binding moiety F2 can be, for instance, biotin, and substrate S can be coated with, for instance, avidin, Significantly, up to three molecules comprising F2 are capable of selectively binding to a second portion of the nanoreporter to selectively immobilize it in its extended state. As illustrated, the molecules comprise a second binding moiety that electively binds a repeated binding moiety of the nanoreporter. The binding moieties can be, for instance, complementary nucleic acid sequences, as illustrated in Figure 14B. The resulting nanoreporter is selectively immobilized in an extended state and should remain extended even when the force is removed. The selectively immobilized, extended nanoreporter can be used for any purpose apparent to those of skill in the art.

5.12.6 IMMOBILIZATION OF TWO PORTIONS OF AN EXTENDED OR ORIENTED NANOREPORTER

[0304] Also disclosed herein are methods for selective immobilization of a first portion and a second portion of a nanoreporter that is in an extended or oriented state. Significantly, according to these methods, the nanoreporter need not be immobilized prior to application of a force capable of extending or orienting the nanoreporter.

[0305] In these methods, the nanoreporter is extended or oriented, or both, by a force capable of extending or orienting the nanoreporter. Such forces are described in detail in the sections above. The force is a force capable of extending or orienting the nanoreporter while maintaining the nanoreporter in one location, i.e. a force capable of extending or orienting without substantially moving the nanoreporter. Exemplary forces include oscillating electromagnetic fields and oscillating hydrodynamic fields. In a particular method, the force is an oscillating electrical field. Exemplary techniques for extending or orienting a nanoreporter in an oscillating electric field are described in Asbury et al, 2002, Electrophoresis 23(16):2658-66; Kabata et al, 1993,Science 262(5139):1561-3; and Asbury and van den Engh, 1998, Biophys J. 74: 1024-30.

[0306] In the methods, the nanoreporter is immobilized at a first portion and at a second portion while extended or oriented. Both the first portion and the second portion can be immobilized non-selectively, both can be immobilized selectively, or one can be immobilized selectively and the other non-selectively. Techniques for immobilization of the first portion and second portion are described in detail in the sections above.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 232

5.12.7 SUBSTRATE FOR IMMOBILIZATION

[0307] In the methods discussed herein, the substrate for immobilization can be any substrate capable of selectively binding the nanoreporter apparent to those of skill in the art. Further, we disclose compositions comprising a selectively immobilized nanoreporter in an extended state. The compositions comprise a substrate, as described herein, having immobilized thereto a nanoreporter in an extended state. The nanoreporter can be, of course, immobilized according to a method of the invention.

[0308] The only requirement of the substrate is that it be capable of selectively binding the second portion of the nanoreporter as described above Thus, the substrate can be a filter or a membrane, such as a nitrocellulose or nylon, glass, a polymer such as polyacrylamide, a gel such as agarose, dextran, cellulose, polystyrene, latex, or any other material known to those of skill in the art to which capture compounds can be immobilized. The substrate can be composed of a porous material such as acrylic, styrene methyl methacrylate copolymer and ethytene/acrylic acid.

[0309] The substrate can take on any form so long as the form does not prevent selective immobilization of the second portion of the nanoreporter. For instance, the substrate can have the form of a disk, slab, strip, bead, submicron particle, coated magnetic bead, gel pad, microtiter well, slide, membrane, frit or other form known to those of skill in the art. The substrate is optionally disposed within a housing, such as a chromatography column, spin column, syringe barrel, pipette, pipette tip, 96 or 384 well plate, microchannel, capillary, etc., that aids the flow of liquid over or through the substrate.

[0310] The nanoreporter can be immobilized on a single substrate or on a plurality of substrates. For instance, the first and second portions of nanoreporter may be immobilized on the same substrate, as recognized by those of skill in the art. The first portion of the nanoreporter can be immobilized on a first substrate while the second portion of the nanoreporter can be immobilized on a second substrate, distinct from the first.

[0311] The substrate can be prepared according to any method apparent to those of skill in the art. For a review of the myriad techniques that can be used to activate exemplary substrates of the invention with a sufficient density of reactive groups, see, the Wiley Encyclopedia of Packaging Technology, 2d Ed., Brody & Marsh, Ed., "Surface Treatment," pp. 867 874, John Wiley & Sons (1997), and the references cited therein. Chemical methods suitable for generating amino groups on silicon oxide substrates are described in Atkinson & Smith, "Solid Phase Synthesis of Oligodeoxyribonucleotides by the Phosphite Triester Method," In: Oligonucleotide Synthesis: A Practical Approach, M J Gait, Ed., 1984, IRL Press, Oxford, particularly at pp. 45 49 (and the references cited therein); chemical methods suitable for generating hydroxyl groups on silicon oxide substrates are described in Pease et ah, 1994, Proc. Natl. Acad. Sci. USA 91:5022 5026 (and the references cited therein); chemical methods for generating functional groups on polymers such as polystyrene, polyamides and grafted polystyrenes are described in Lloyd Williams et al., 1997, Chemical Approaches to the Synthesis of Peptides and Proteins, Chapter 2, CRC Press, Boca Raton, FL (and references cited therein).

[0312] Exemplary useful substrates include surfaces coated with streptavidin, e.g. Accelr8 TB0200. Further useful substrates include surfaces coated with N-hydroxysuccinarnide that are capable of reacting with a portion of a nanoreporter that comprises an amine. One such surface is OptArray-DNA (Accelr8). Additional useful surfaces are coated with aldehyde (e.g. Nexterion Slide AL., Schott) and surfaces coated with epoxy (e.g. Nexterion Slide E, Schott). Another useful surface is a biotinylated BSA coated surface useful for selective immobilization of a portion of a nanoreporter that comprises avidin or streptavidin.

5.12.8 METHODS OF USING SELECTIVELY IMMOBILIZED.EXTENDED OR ORIENTED NANOKEPORTERS

[0313] The selectively immobilized, elongated nanoreporters can be used to create macromolecular barcodes for the purposes of separation and sequential detection of labels. These labels spaced along the molecule provide a unique code that can be read when the nanoreporter is extended and immobilized. Extension and selective immobilization can facilitate the decoding of the macromolecular barcode.

[0314] The selectively immobilized, elongated nanoreporters can further be used for can be used in any context where detection or imaging of a nanoreporter might be useful. They can be used for diagnostic, prognostic therapeutic and screening purposes. For instance, they 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 and/or to stage the disease. They 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. The compositions and methods of the invention can be used to quantitate target molecules 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. In addition, the compositions and methods of the invention can be used to provide prognostic information that assists in determining a course of treatment for a patient.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 233

5.12.9 KITS COMPRISING SELECTIVELY IMMOBILIZED, EXTENDED OR ORIENTED NANOREPORTERS

[0315] We also disclose kits comprising one or more components of the invention. The kits can comprises, for example, a substrate according to the invention and one or more extended or oriented, or both, nanoreporters selectively immobilized on the substrate. The kits can be used for any purpose apparent to those of skill in the art, including, those described above.

[0316] Furthermore, we disclose kits useful for the extension and selective immobilization of nanoreporters. The kits can comprise a substrate for immobilization and one or more binding partners to facilitate extension or immobilization of a nanoreporter. The binding partners could comprise a moiety useful for extension of the nanoreporter in an appropriate force. In certain embodiments, the binding partners could facilitate immobilization or selective immobilization of the nanoreporter to the surface. The kit could comprise a nanoreporter for extension and immobilization. The kit could comprise a device capable of extending the nanoreporter.

5.13 DETECTION OF NANOREPORTERS

[0317] Nanoreporters are detected by any means available in the art that is capable of detecting the specific signals on a given nanoreporter. Where the nanoreporter is fluorescently labeled, suitable consideration of appropriate excitation sources may be investigated. Possible sources may include but are not limited to are 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 nanoreporter.

5.13.1 MICROSCOPY AND OBJECTIVE LENS SELECTION.

[0318] The major consideration regarding the microscope objective lens is with the optical resolution, which is determined by its numerical aperture (NA). Generally, the larger the NA, the better the optical resolution. The required NA is preferably at least 1.07 based on the relationship of [delta] = 0.617JNA ([delta] = optical resolution and [lambda] = wavelength). The amount of light that is collected by an objective is determined by NA<4>/Mag<2> (Mag = magnification of the objective). Therefore, in order to collect as much light as possible, objectives with high NA and low magnifications should be selected.

5.13.2 CCD CAMERA SELECTION AND IMAGE CAPTURE TECHNIQUES.

[0319] When selecting a CCD camera, the first consideration is the pixel size, which partially determines the resolution of the imaging system. Optimally the optical resolution should not be compromised by the CCD camera. For example, if the optical resolution is 210-300 non, which corresponds to 12.6-18 [mu]m on a CCD chip after a 60 x magnification, in order to resolve and maintain the optical resolution there should be at least two pixels to sample each spot. Or the pixel size of the CCD chip should be at most 6.3- 9 [mu]m.

[0320] The second consideration is detection sensitivity which can be determined by many factors that include but are not limited to pixel size, quantum efficiency, readout noise and dark noise. To achieve high sensitivity, select a qualitative camera with big pixel size
(which can give big collection area), high quantum efficiency and low noise. An exemplary camera with these criteria is the Orca-Ag camera from Hamamatsu Inc. The chip size is 1344 x 1024 pixels; when using the 6Ox objective, the field of view is 144 x 110 [mu]m<2>.

5.14 COMPUTER SYSTEMS

[0321] We Also disclose computer systems that may be used to computerize nanoreporter image collection, nanoreporter identification and/or decoding of the nanoreporter code. Specifically, we disclose various computer systems comprising a processor and a memory coupled to the processor and encoding one or more programs. The computer systems can be connected to the microscopes employed in imaging the nanoreporter, allowing imaging, identification and decoding the nanoreporter, as well as storing the nanoreporter image and associated information, by a single apparatus. The one or more programs encoded by the memory cause the processor to perform the methods of the invention.

[0322] We also disclose computer program products for use in conjunction with a computer system (e.g., one of the above-described computer systems of the invention) having a processor and a memory connected to the processor. The computer program products of the invention comprise a computer readable storage medium having a computer

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 234

program mechanism encoded or embedded thereon. The computer program mechanism can be loaded into the memory of the computer and cause the processor to execute the steps of the methods of the invention.

[0323] The methods described in the previous subsections can preferably be implemented by use of following computer systems, and according to the following methods. An exemplary computer system suitable for implementation of the methods of this invention comprises internal components and being linked to external components. The internal components of this computer system include a processor element interconnected with main memory. For example, the computer system can be an Intel Pentium-based processor of 200 MHz or greater clock rate and with 32 MB or more of main memory.

[0324] The external components include mass storage. This mass storage can be one or more hard disks which are typically packaged together with the processor and memory. Such hard disks are typically of 1 GB or greater storage capacity. Other external components include user interface device, which can be a monitor and a keyboard, together with pointing device, which can be a "mouse", or other graphical input devices (not illustrated). Typically, the computer system is also linked to a network link, which can be part of an Ethernet link to other local computer systems, remote computer systems, or wide area communication networks, such as the internet. This network link allows the computer system to share data and processing tasks with other computer systems.

[0325] Loaded into memory during operation of this system are several software components, which are both standard in the art and special to the instant disclosure. These software components collectively cause the computer system to function according to the methods discussed herein. The software components are typically stored on mass storage. A first software component is an operating system, which is responsible for managing the computer system and its network interconnections. This operating system can be, for example, of the Microsoft Windows(R) family, such as Windows 95, Windows 2000, or Windows XP or, alternatively, a Macintosh operating system, a Linux operating system or a Unix operating system. A second software component may include common languages and functions conveniently present in the system to assist programs implementing the methods specific to this invention. Languages that can be used to program the analytic methods of the invention include, for example, C, C++, JAVA, and, less preferably, FORTRAN, PASCAL5 and BASIC. Another software component of the present invention comprises the analytic methods of this invention as programmed in a procedural language or symbolic package.

[0326] In an exemplary implementation, a nanoreporter code (i.e., a correlation between the order and nature of spots on a nanoreporter and the identity of a target molecule to which such a nanoreporter binds) is first loaded in the computer system. Next the user causes execution of analysis software which performs the steps of determining the presence and, optionally, quantity of nanoreporters with a given nanoreporter code.

[0327] The analytical systems also include computer program products that contain one or more of the above-described software components such that the software components may be loaded into the memory of a computer system. Specifically, a computer program product of the invention includes a computer readable storage medium having one or more computer program mechanisms embedded or encoded thereon in a computer readable format. The computer program mechanism encoded, e.g., one or more of the analytical software components described above which can be loaded into the memory of computer system and cause the processor of the computer system to execute the analytical methods of the present invention.

[0328] The computer program mechanism or mechanism are preferably stored or encoded on a computer readable storage medium. Exemplary computer readable storage media are discussed above and include, but are not limited to: a hard drive, which may be, e.g., an external or an internal hard drive of a computer system of the invention, or a removable hard drive; a floppy disk; a CD-ROM; or a tape such as a DAT tape. Other computer readable storage media will also be apparent to those skilled in the art that can be used in the computer program mechanisms disclosed herein.

[0329] We also disclose herein databases useful for practicing the methods of the present invention. The databases may include reference nanoreporter codes for a large variety of target molecules. Preferably, such a database will be in an electronic form that can be loaded into a computer system. Such electronic forms include databases loaded into the main memory of a computer system used to implement the methods of this invention, or in the main memory of other computers linked by network connection, or embedded or encoded on mass storage media, or on removable storage media such as a CD-ROM or floppy disk.

5.15 APPLICATIONS OF NANOREPORTER TECHNOLOGY

[0330] The compositions and methods or the invention are for use in diagnostic, prognostic therapeutic 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.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 235

5.15.1 DIAGNOSTIC/PROGNOSTIC METHODS

[0331] 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 and/or to stage the disease.

[0332] 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 cancerous cell type in the sample, thereby diagnosing or staging the cancer.

[0333] Alternatively, the methods descried 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.

[0334] 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.

[0335] Because of the quantitative nature of nanoreporters, the compositions and methods of the invention can be used to quantitate target molecules 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.

[0336] In addition, the compositions and methods of the invention can be used to provide prognostic information that assists in determining a course of treatment for a patient. For example, the amount of a particular marker for a tumor can be accurately quantified from even a small sample from a patient. For certain diseases like breast cancer, overexpression of certain genes, such as Her2-neu, indicate a more aggressive course of treatment will be needed.

5.15.2 ANALYSIS OF PATHOLOGY SAMPLES

[0337] RNA extracted from formaldehyde- or paraformaldehyde-fixed paraffin-embedded tissue samples is typically poor in quality (fragmented) and low in yield. This makes gene expression analysis of low-expressing genes in histology samples or archival pathology tissues extremely difficult and often completely infeasible. The nanoreporter technology can fill this unmet need by allowing the analysis of very small quantities of low-quality total RNA.

[0338] To use nanoreporter technology in such an application, total RNA can be extracted from formaldehyde- or paraformaldehyde-fixed paraffin-embedded tissue samples (or similar) using commercially available kits such as RecoverAll Total Nucleic Acid Isolation Kit (Ambion) following manufacturer's protocols. RNA in such samples is frequently degraded to small fragments (200 to 500 nucleotides in length), and many paraffin-embedded histology samples only yield tens of nanograms of total RNA. Small amounts (5 to 100 ng) of this fragmented total RNA can be used directly as target material in a nanoreporter hybridization following the assay conditions described herein. As described in Example 6 in Section 11 below, nanoreporter analysis of approximately 3.3 ng cellular RNA permitted detection of transcripts present at approximately 0.5 copy/cell.

5.15.3 SCREENING METHODS

[0339] The methods of the present invention may be for use, inter alia, in determining the effect of a perturbation, including chemical compounds, mutations, temperature changes, growth hormones, growth factors, disease, or a change in culture conditions, on various target molecules, thereby identifying target molecules whose presence, absence or levels are indicative of a particular biological states. Optionally, the methods are for use to elucidate and discover components and pathways of disease states. For example, the comparison of quantities of target molecules present in a disease tissue with "normal" tissue allows the elucidation of important target molecules involved in the disease, thereby identifying targets for the discovery/screening of new drug candidates that can be used to treat disease.

5.16 KITS

[0340] We disclose kits comprising one or more components of the invention. The kits can contained pre-labeled nanoreporters, or unlabeled nanoreporter with one or more components for labeling the nanoreporters. Moreover, the nanoreporters provided in a kit may or may not have target-specific sequences pre-attached. Optionally, the target sequences are provided in the kit unattached to the nanoreporter scaffold.

[0341] The kit can include other reagents as well, for example, buffers for performing hybridization reactions, linkers, restriction endonucleases, and DNA ligases,

[0342] The kit also will include instructions for using the components of the kit, and/or for making and/or using the labeled nanoreporters.

6. EXAMPLE 1: NANOREPORTER MANUFACTURING AND PROTOCOL

[0343] Herein is a step-by-step example of a method construction of a nanoreporter from various components.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 236

[0344] It can be appreciated that various components can be constructed or added either at the same time, before or after other components. For example, annealing patch units or flaps to a scaffold can be done simultaneously or one after the other.

[0345] In this example the starting material is a circular M13mpl8 viral vector. Using a single Linear strand M13Mp18, patch units are annealed to it to form a double stranded scaffold. Next, flaps are added then a target-specific sequence is ligated. Meanwhile purification steps aid to filter out excess, unattached patch units and flaps. Construction of labeled nucleic acids (patches and/or flaps and/or other labeled oligonucleotides) that bind the nanoreporter are also described.

[0346] Upon attachment (e.g., via hybridization) of a target molecule, the nanoreporter is attached to a surface and stretched. Finally the nanoreporters are imaged by a camera.

[0347] Nanoreporters were generated and successfully employed to detect target molecules using methods substantially as described in this example. An example of target detection using such this method is shown in Figure 4.

6.1 SCAFFOLD CONSTRUCTION

[0348] The oligonucleotide scaffold sequence selected was analyzed using Vector NTI® software. First, a single stranded nucleic acid was made from linearizing a circular M13mp18 single stranded DNA, which was commercially purchased from New England Biolabs. The circular M13mp 18 was digested with BamH1 enzyme to linearize it. Materials used consisted of M13mp18 vector (250 ng/μl), Patch_1L_BamH1.02 (10 μM dilution of a 100 μM stock), 10X BamH1 Buffer, BamH1 enzyme. Protocol for making 0.8 pmol total of linear M13mp18 involve the following steps. 1) preheat heating block to 37 °C; 2) in a 0.65ml ependorff tube combine 40 μl of 250 ng/μl M13mp18 vector, 2 μl of 10 μM Patch_1L_BamH1.02, and 5 μl of 10X BamH1 Buffer; 3) place the ependorff tube in the 37°C heating block with foil over the top. Incubate the tube at 37°C for 15 minutes to allow the patch unit to hybridize to the M13mp18 scaffold; 4) after 15 minutes add 2 μl of BamH1 enzyme and let the reaction digest at 37 °C for 30 minutes, after which add an additional 2 μl of BamH 1 enzyme and let the reaction continue to digest for another 30 minutes at 37°C (final volume of BamHI enzyme is 8%); and 5) aliquot 10 μl into 0.65ml ependorff tubes and store in freezer (final concentration of linear M13mp18 is 200 ng/μl).

6.2 PATCH UNIT PREPARATION OF THE BASE PATCH POOLS (BPP).

[0349] Second, patch units are prepared in pools. Patch oligonucleotide sequences were selected for optimal length and desired homology/non-homology to M13mp18 strand and the human genomic sequence. Patches were commercially manufactured oligonucleotides (purchased from Integrated DNA technologies) either 60 or 65 nucleotide bases in length. 50 nucleotide bases of each patch oligonucleotide are complementary to the M13mp18 single stranded DNA, 10 nucleotide bases are complementary to an adjacent patch, and 5 nucleotides base pairs are complementary to a corresponding flap. The 10 nucleotide base match between patches forms a stem structure which stabilizes the structure and helps lift the flaps off the covered scaffold so they are more available to bind labeled oligonucleotides. Synthetic binding sites, the 5 nucleotide bases, on the patches for binding to the flaps make leveraging the power of a modular system possible.

[0350] The base patch pools contain nine patch units all corresponding to a specific letter grouping and position on the nanoreporter. For this example, there are four different fluorescent dyes (color) labeled A, B, C, and D and 8 different positions or regions where labeled nucleic acids can bind on a nanoreporter. For example, BPP A3 corresponds to all of the A patch units at position 3 (patch units 19-27) on the nanoreporter.

[0351] The nanoreporter positions are as follows:

Position 1: Patch units 1-9 (A or C)
Position 2: Patch units 10-18 (B or D)
Position 3: Patch units 19-27 (A or C)
Position 4: Patch units 28-36 (B or D)
Position 5: Patch units 37-45 (A or C)
Position 6: Patch units 46-54 (B or D)
Position 7: Patch units 55-63 (A or C)
Position 8: Patch units 64-72 (B or D)

[0352] Materials: right and left patches, pre-annealed to each other (each oligonucleotide is at a concentration of 10 μM). Materials for making 100 pmol of BPP 1: (In position 1, patch coordinate 1L is used for the BamHI digest -this patch is not included in BPP 1): 10 μl each pre-annealed (10 μM/each) patch unit (coordinates 2-9), 5 μl [20 μM] Patch_1R (A or C). Final concentration of each patch is 1.18 pmol/ μl. Materials for making 100 pmol of BPP 2-8: 10 μl each pre-annealed (10 μM/each) appropriate patch unit. There are 9 patch units added to each, or 90 μl total. Final concentraion of each patch is 1.11 pmol/ μl.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 237

[0353] Below is a table of all the patch unit pools made for this example, with 8 positions or regions for dye-labeled nucleic acids to bind on the nanoreporter. Positions 1, 3, 5, and 7 can bind to nucleic acid labeled with dye A or dye C, and a positions 2, 4, 6, and 8 can bind to nucleic acid labeled with dye B or dye D.

Table 2 of resulting Basic Patch Pools (correspond to labels on tubes)

BPP-A1 [Pre-Paired, Color=A, Coordinates 1-9] Patch_(1-9)R.A Patch_(2-9)L
BPP-B2 [Pre-Paired, Color=B, Coordinates 10-18] Patch_(10-18)R.B Patch (10-18)L
BPP-A3 [Pre-Paired, Color=A, Coordinates 19-27] Patch_(19-27)R.A Patch (19-27)L
BPP-B4 [Pre-Paired, Color=B, Coordinates 28-36] Patch_(28-36)R.B Patch_(28-36)L
BPP-A5 [Pre-Paired, Color=A, Coordinates 37-45] Patch_(37-45)R.A Patch (37-45)L
BPP-B6 [Pre-Paired, Color=B, Coordinates 46-54] Patch_(46-54)R.B Patch (46-54)L
BPP-A7 [Pre-Paired, Color=A, Coordinates 55-63] Patch_(55-63)R.A Patch_(55-63)L
BPP-B8 [Pre-Paired, Color=B, Coordinates 64-72] Patch_(64-72)R.B Patch (64-72)L
BPP-C1 [Pre-Paired, Color=C, Coordinates 1-9] Patch_(1-9)R.C Patch (2-9)L
BPP-D2 [Pre-Paired, Color=D, Coordinates 10-18] Patch_(10-18)R.D Patch (10-18)L
BPP-C3 [Pre-Paired, Color=C, Coordinates 19-27] Patch_(19-27)R.C Patch (19-27)L
BPP-D4 [Pre-Paired, Color=D, Coordinates 28-36] Patch_(28-36)R.D Patch (28-36)L
BPP-C5 [Pre-Paired, Color=C, Coordinates 37-45] Patch_(37-45)R.C Patch (37-45)L
BPP-D6 [Pre-Paired, Color=D, Coordinates 46-54] Patch_(46-54)R.D Patch (46-54)L
BPP-C7 [Pre-Paired, Color=C, Coordinates 55-63]

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 238

(continued)

Patch_(55-63)R.C Patch (55-63)L
BPP-D8 [Pre-Paired, Color=D, Coordinates 64-72] Patch_(64-72)R.D Patch_(64-72)L

6.3 MATERIALS AND PREPARATION FOR ANNEALING THE SINGLE STRANDED OLIGONUCLEOTIDE WITH PATCH UNITS FOR A DOUBLE STRANDED SCAFFOLD.

[0354] Third, patch units are prepared to be annealed to the single stranded linear M13mp18, covering the strand in order to make a double stranded oligonucleotide scaffold. Conditions for annealing 60 and 65 nucleotide base patches to the M13mp18 need to occur at high salt concentrations so that binding will be very specific and patches will not anneal to an incorrect coordinate on the M13mp18 strand. For the annealing step, each patch unit is added at a 2:1 to 4:1 ratio with the single stranded M13mp18 sequence at 0.5 pmol total volume. Excess patches are removed before annealing flaps.

[0355] Materials used consisted of 20x SSC, linear M13mp18 (BamH1 digested at 0.08 pmol/μl or 200 ng/μl), appropriate base patch pools (BPP) (need 8 total at 1.11 pmol/μl - see above) and digital heat block set at 45°C. Annealing reaction make up is as follows. General guidelines: 2X each patch unit per M13mp18 molecule, pre-ligated flaps/patches (in position 1 or 8) added for purification later, and 5X SSC. Example (0.5 pmol of scaffold with F8 hook flaps) reaction consists of: 7.1 μl BamH1 Digested M13mp18 strand at 0.071 μM, 0.9 μl each new Base Patch Pools at 1.11 μM for first 7 positions: A1, B2, A3, B4, C5, B6 and A7:

1.7 μl A1 BPP (Pre-Annealed, 12/15; at 1.18 μM/each patch)
1.8 μl B2 BPP (Pre-Annealed, 12/15; at 1.11 μM/each patch)
1.8 μl A3 BPP (Pre-Annealed, 12/15; at 1.11 μM/each patch)
1.8 μl B4 BPP (Pre-Annealed, 12/15; at 1.11 μM/each patch)
1.8 μl C5 BPP (Pre-Annealed, 12/15; at 1.11 μM/each patch)
1.8 μl B6 BPP (Pre-Annealed, 12/15; at 1.11 μM/each patch)
1.8 μl A7 BPP (Pre-Annealed, 12/15; at 1.11 μM/each patch),
2.4 μl BPP-D8 (pool of the first seven patch units - coordinates 64, 65, 66, 67, 68, 69, and 70 at position 8 - "D" specificity) with purification tags- F8 (FHF, which anneal to patch coordinates 7 1 L, 71R, 72L, 72R, 73L making full split-flap/patch units that have "F" specificity for use as biotin linkers, at position F8) at 0.83 μM, and 7.3 μl 20X SSC. The final reaction volume will be 29.3 μl at 0.027 pmol/ μl.

[0356] Anti-Bam oligonucleotide is also added to anneal to region in M13 that is complementary to the (missing) 1L patch unit and to prevent recircularization of the M13 scaffold during ligation.

6.4 ANNEALING PATCH UNITS TO SINGLE STRANDED M13MP18 TO FORM A DOUBLE STRANDED SCAFFOLD.

[0357] The fourth step involves the protocol to anneal the patch units to the single stranded linear M13mp18, covering the strand in order to make a double stranded oligonucleotide scaffold, is performed in the following steps: 1) preheat heating block to 42°C, heat above reaction solution to 45°C in small PCR (or strip) tube(s) with foil over top for 15 minutes, turn heat block to 65°C and incubate for an additional 1 hour and 45 minutes and remove tubes, place on ice or freeze.

6.5 PURIFICATION OF NANOREPORTER SCAFFOLD USING BIOTIN AND MAGNETIC BEADS WITH STREPTAVIDIN.

[0358] The fifth step occurs before attaching the flaps, where excess patch units that have not annealed to the M13mp18 strand are separated from the double stranded oligonucleotide scaffold. A purification tag with a 5 nucleotide base homologous region to some of the patch units' complementary 5 nucleotide base overhang is annealed to 'hook' the scaffold. Biotinylated oligonucleotides are annealed to the 'purification tag' and magnetic beads with streptavidin are used to capture the scaffold using the biotinylated oligonucleotides. Excess patch units are removed with the supernatant. The scaffold melts off of the magnetic beads into solution for recovery.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 239

6.6 ANNEAL THE D-BIOTIN CATCHERS TO THE PURIFICATION TAGS

[0359] Anneal the D-Biotin catchers to the purification tags on the nanoreporter (making 2X to amount of D8-flap positions available in solution, which is 2X to M13, or 4X final): 0.5 pmol X 25 hook oligonucleotide positions (5 multiplied by 5), 4X makes 50 pmols translates to 0.50 μl of 100 pmol/ μl D-biotin, add 0.5 μl (D, E, F) - Biotin (at 100 μM) to sample, mix and incubate at room temperature for 30 minutes.

6.7 PURIFICATION PROTOCOL TO WASH OFF UNATTACHED PATCH UNITS FROM DOUBLE STRANDED SCAFFOLD.

[0360] Anneal F-hook oligonucleotides in a 25 fold excess to nanoreporters in 5X SSC for 30 min at room temperature. Pipet 200 μl DynaBead MyOne Streptavidin™ bead solution into 1.5 ml tubes, place on magnet and remove supernatant. Wash twice with 5X SSC by resuspending and clearing with magnet as in step above. Add 80 μl of sample in 5X SSC (80 fmoles of sample in this example). Resuspend well, by placing on vortex for 15 minutes. Clear solution with magnet and transfer supernatant to fresh tubes for later gel analysis. While on magnet, wash pellets (do not resuspend) with 80 μl TE by pipeting over pellet three times with the same 80 μl volume originally added. Remove wash, place in freshy "wash" tubes for analysis. Heat up TE buffer to 45°C, add 80 μl to each pellet and resuspend. Place tubes on 45°C heat block for 15 minutes, pipetting up/down once to insure beads remain suspended. Immediately clear product with magnet while warm and save. The majority of purified nanoreporters should be present in this product eluted at 45°C.

6.8 ANNEALING AND LIGATION OF FLAPS TO SCAFFOLD.

[0361] The sixth step involves split flap oligonucleotides which are annealed to the scaffold to make a 'covered scaffold.' Purification with magnetic beads is performed afterwards to remove excess split flaps. Ligation of the covered scaffold is done using T4 ligase to increase the stability of the structure. Only one type of flap is needed per fluorescent dye. Flaps are either 95 or 100 bases in length and have regions complementary to the patches, to labeled oligonucleotides and to each other. Each flap has 15 base repeating sequences for binding to labeled oligonucleotides. The repeat sequences are based on Lambda sequences that have been analyzed to remove any palindromes and hairpin structures.

[0362] Conditions for annealing the flaps are as follows. The sequence on the flaps that corresponds to the patch is 5 nucleotide base pairs long, and therefore the flaps anneal specifically to the patches even at high salt concentrations. The ratio of flaps to patches is 2:1. In order to increase stability at high temperatures, ligation of patches to each other and the flap to the patches may be carried out in the same reaction.

[0363] 1) Quantify the purified scaffold using a spectrometer at A260 nm. Calculate the volume needed for appropriate amount of nanoreporter to prepare. For this example we used 110 ng or 0.023 pmol, reading at A260 nm shows 7.7 ng/μl, or 14.3 μl for 110 ng. 2) Setup ligation reaction as follows (volume will vary, depending on the purification and scale). Currently using 1.5X flaps to patches, calculate accordingly. For this example, there are four different fluorescent dyes (color) labeled A, B, C, and D and 8 different positions or regions where dye-labeled nucleic acids can bind on a nanoreporter. The number of positions for each color (in this case 1-4) multiply by 9 multiply by 1.5 moles of scaffold = moles of flaps to use.

[0364] For the nanoreporter with fluorescent dye in the sequence/positions [ABABCBAD] :

ABABCBAD =

A:40.5 x .023 = .93 pmol; vol: .93 μl of SF (split flap)-AL at 1 μM

.93 μl of SF-AR at 1 μM

B:40.5 x .023 = .93 pmol; vol: .93 μl of SF-BL at 1 μM

.93 μl of SF-BR at 1 μM

C:13.5 x .023 = .31 pmol; vol: .31 μl of SF-CL at 1 μM

.31 μl of SF-CR at 1 μM

D:13.5 x .023 = .31 pmol; vol: .31 μl of SF-DL at 1 μM

.31 μl of SF-DR at 1 μM

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 240

[0365] Ligation reaction (25 μl total) consists of : Split Flaps (see above; 4.96 μl, or ~5 μl total), 14.3 μl of MODB-Scaffold at 0.0016 pmol/μl, 2.5 μl 10X T4 ligation Buffer, 2.2 μl NanoPure H2O and 1 μl T4 ligase. Incubate tubes 5 minutes at 45°C. Move to 37°C water bath, inc. for 5 minutes. Add 1 μl T4 ligase to samples. Incubate for additional 1hour at 37°C. Freeze immediately, or heat at 75°C for 5 minutes to kill T4 ligase.

6.9 LIGATION OF TARGET-SPECIFIC SEQUENCES TO NANOREPORTERS

[0366] The seventh step involves ligation of a target-specific sequence to the nanoreporter. A DNA target-specific sequence is designed to be complementary to the target molecule, which can be RNA (*e.g.*, mRNA) or DNA (*e.g.*, cDNA or genomic DNA). The target-specific sequence can be from 35, 60 or 70 nucleotide bases in length. The target-specific sequence can be ligated to the scaffold using a single stranded overhanging region on the covered scaffold. The scaffold with a single type of target-specific sequence can be manufactured separately and then mixed to form libraries.

6.10 NANOREPORTER CONSTRUCTION

[0367] Addition of oligonucleotides to a nanoreporter can be done at any point during the construction of a nanoreporter. In certain aspects of the present invention, a labeled oligonucleotide is 15 nucleotide bases long. On the 5' end, a single fluorophore dye is attached. Oligonucleotides with a particular fluorophore dye will generally have the same sequence. These labeled oligonucleotides bind to the repeat sequences of the split flaps. Fluorophores best suited for this example include but are not limited to Alexa 488, cy3, Alexa 594, and Alexa 647. The 15 nucleotide base length holds the fluorophores far enough apart so that they cannot quench each other and ensure that the labeled nucleic acids will be stable (will not melt off complementary strand) at conditions in the visualization process. Labeled oligonucleotides are stable at 40°C. This short length also allows for packing a large number of fluorescent dyes onto the flaps. In certain aspects of the invention, labeled oligonucleotides are introduced during the target sample processing.

6.11 ATTACHMENT OF NANOREPORTERS TO TARGET MOLECULES

[0368] Nanoreporters can be attached to target molecules using any means known to one of skill in the art. In an exemplary embodiment, dual nanoreporters are hybridized to target molecules by mixing 250 pmols each of both the first probe and the second probe with 125 pmols of target. The total volume is adjusted to 4 μl and a final concentration of buffer of 5X SSC. This mixture is incubated in a covered PCR tube overnight at 42 degrees to allow hybridization to occur.

6.12 SURFACE ATTACHMENT

[0369] Once the nanoreporters are attached to both target molecule and corresponding labeled nucleic acids, *i.e.*, nucleic acids attached to label monomers, they are attached to a surface and stretched in resolve the order of signals emitted by the label monomers and thus identify the target molecule. In this example, the nanoreporters are stretched to spatially resolve their fluorescent dye codes which correspond to a particular target molecule. The nanoreporters are stretched by attaching one end to a surface (in this example - a coverslip, see preparations below). Two methods for surface attachment may be used: A) streptavidin coated slides from Accelr8 Corporation with the nanoreporters being biotinylated and B) biotin coated slides with the nanoreporters having streptavidin. In buffer, the nanoreporters are brought into contact with the active surface and allowed to incubate for a period of time. The reaction is performed in flow cells which were made from PDMS molded in etched silicon wafers to make the channels. Metal tubing is used to core wells at the ends of the channels for buffer and sample insertion. Channel dimensions are 0.5 mm or 1 mm wide and 54 μm high. Once the sample has been loaded into the flow cell lane and incubated, the nanoreporters should be attached. Nanoreporters can be stretched either by applying a voltage or by removing the liquid with a receding meniscus leaving the strings stretched and dry.

6.13 PREPARATION OF SURFACE AND ASSEMBLY OF DEVICE

[0370] The binding surfaces (Accclr8 brand Streptavidin-OptiChem, coated coverslips) are shipped in units of 5 surfaces per slide container, and each container is enclosed with a package of silica dessicant in a foil pouch. The pouches are stored at -20 °C until use.

[0371] To prepare the surface for binding, a pouch is first pulled from the freezer and allowed to come to room temperature over several minutes. If previously unopened, the pouch is then sliced along one edge to form a slit, and the container of surfaces is removed. Upon removal of the required surface, the container is replaced in the pouch with its dessicant, the slit is sealed closed with a strip of packaging tape, and the pouch is replaced in the freezer.

[0372] The surface is then lightly rinsed with a stream of Nanopure water (Barnstead Nanopure Diamond) and soaked

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 241

for 10 minutes in 0.2 μm-filtered 1X PBS in a clean, slotted Coplin jar. After soaking, the surface is dipped in Nanopure water and dried by blowing filtered nitrogen across the surface edge.

[0373] The PDMS device used to mate with the surface and provide localization of the sample is cleaned just before use by applying cellophane tape to the PDMS surface and then peeling away dust or other particles which may have become attached during storage. The binding side of the AccelR8 surface is laid face-up, and the clean PDMS structure is centered, channel side down, on the surface. PDMS adheres readily to coated glass, and no further attachment mechanism is necessary.

6.14 SAMPLE BINDING AND WASHING

[0374] The sample is bound to the surface by first applying a 5 μL drop of the sample (currently diluted in 100 mM sodium borate buffer, pH 9.8) in one well of the chosen lane. The drop should just touch the point at which the channel joins the well (some sample may wick into the channel at this point). The channel is filled, and binding is equalized throughout the channel, by pulling the droplet through the channel to the opposite well using a very weak vacuum (< 2 kPa). The process is repeated for the other samples in their respective lanes. Excess fluid is then removed from the wells, the wells are taped to reduce evaporation, and the device is incubated at room temperature in the dark for 20 minutes.

[0375] After binding, the tape is removed, and the top well of each lane is filled with 100 μL of the borate buffer described above. About 20 μL of that buffer is pulled through the channels to the other wells using the vacuum, and the process is repeated once. All borate buffer is then removed from all wells, and the top well is filled with 1X TAE, pH 8.3. About 50 μL TAE is pulled through the channel, then all TAE is removed and the well is refilled. The process is repeated three times, for a total of about 150 μL of TAE rinse. Finally, all wells are filled with 100 μL 1X TAE.

6.15 ELECTROSTRETCHING

[0376] The bottom of the coverslip/PDMS device is spotted with immersion oil and placed on the microscope. Electrodes are inserted into the wells on opposite ends of the first PDMS channel (negative electrode in top well, positive in bottom). The first image of the channel will be taken close to the bottom well; the microscope stage is adjusted so that the area of interest is in focus.

[0377] Voltage (200 V) is then applied across the channel. Voltage is supplied by a DC power supply (Agilent E3630A) and amplified 100X through a home-built amplifier. After the current is applied, focus is readjusted, and the imaging process begins.

[0378] The electrostretching and imaging process is then repeated with the remaining channels. Image the bindings.

6.16 LIGHT SOURCE FOR THE FLUORESCENT DYES ON THE NANOREPORTER

[0379] In using an arc lamp as a light source, then best fluorophore selection is the brightest types without leading to fluorescent overlap such as Alexa 488, Cy3, and Alexa 594. Weaker fluorescent dyes such as Alexa 647 and Cy5.5 may also be used.

6.17 FILTERS TO IMAGE THE FLUORESCENT DYES ON THE NANOREPORTER

[0380] For the selected fluorophores Alexa 488, Cy3, Alexa 594 and Alexa 647 there maybe an overlap between the Cy3 and Alexa 594. However, custom ordering an emission filter with a bandwidth of 572-600 nm minimizes the overlap.

6.18 MICROSCOPE AND OBJECTIVE LENS TO IMAGE THE NANOREPORTERS

[0381] The microscope model used was the Nikon Eclipse TE2000E from Nikon Incorporation using the inverted fluorescence imaging station which has 6 filter cassettes that allow the selection of fluorescent emission from multiple fluorescent dye candidates. For the selected dyes, the optical resolution required is about 400 nm for all the wavelengths (500-700 nm). The selected objective lens is the Nikon Plan Apo TIRF lens which has a NA of 1.45 and magnification of 60. The optical resolution is ~210-300 nm for different wavelengths.

7. EXAMPLE 2: PATCH/FLAP NANOREPORTER MANUFACTURING PROTOCOL

[0382] This example demonstrates another way of making a nanoreporter which consists of a single stranded linear M13mp18 viral DNA, oligonucleotide patch units and long flaps.

[0383] Nanoreporter label units were successfully generated using methods substantially as described in this example.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 242

[0384] Pre-phosphorylated patch units and flaps are added together with the M13mp18 DNA vector and ligated together. After the ligation of the flaps to the patch units which are ligated to the M13mp18 DNA, the BamH1 enzyme is introduced to linearize the vector.

[0385] Prepare a batch of nanoreporters starting with 5 μg of M13mp18 as a scaffold. The hybridization may be scaled up accordingly to the desired amount. This process will take about 1-2 days to complete.

Materials:

Qty	Item	Vendor
20	250 ug/μl M13mp18 viral ssDNA	New England Biolabs
27 μl	0.74 pmol/μl Oligonucleotide Patch Unit Mix	IDT
8 μl	Long Flap Oligonucleotide A 100 pmol/μl	IDT
8 μl	Long Flap Oligonucleotide B 100 pmol/μl	IDT
0.5 μl	Flap patch Oligos at 100 pmol/μl from plates #529916 and #610591	IDT
31 μl	T4 Ligase 10x buffer	Fermentas
19 μl	T4 Ligase	Fermentas
15 μl	Optikinase 10x buffer	USB
4.2 μl	100 mM ATP	ANY
5 μl	Optikinase Enzyme 10 units/μl	USB
1 μl	BamH1 oligonucleotide 10 pmol/μl	IDT
20 μl	BamH1 10x buffer	Fermentas
3 μl	BamH1 Enzyme 10 units/μl	Fermentas

[0386] Preheat water bath to 37°C and 55°C °C before beginning protocol. Make sure buffers are all well mixed and thawed before using. A work plate should be available and labeled with the ordered oligos from IDT in plates #529916 and #610591. Take these two plates out and thaw at room temperature for 0.5-1 hours and spin down contents before removing the tape that covers the wells. Four separate reactions will be set up in 1.5 ml eppendorf tubes using specific oligonucleotides from these plates. To begin label these four separate tubes with roman numerals on their caps. Columns 5 and 6 A through H are for reaction I, Columns 7 and 8 A through H are for reaction ii are all found in plate #529916. Columns 1 and 2 are for reaction iv, and Columns 3 and 4 are for reaction iii.

[0387] Flap Ligations (Step A): Label four separate 1.5 ml tubes with roman numerals i through iv (mentioned above). Add the reagents below accordingly to each 50 μl reaction containing: 5 μl 10x ligase buffer, 0.5 μl/oligonucleotide from designated wells from plates #529916 and #610591, 4 μl Long Flap Oligo/reaction (A or B) for reactions I, ii and iv. 3 μl of LF for area iii, 29 H_2O for reactions I,ii and iv. 32 μl H_2O for reaction iii, and 4 μl T4 ligase. Preanneal oligos in this mix without the ligase at 37 °C for half an hour. Add ligase as last reagent and allow to ligate at room temperature for at least four hours. Product concentration is 1 pmol/flap/μl.

[0388] Flap Ligation Phosphorylation (Step B) Label four separate 1.5 ml tubes with roman numerals again, one through four with a P inside a circle to designate that the products are phosphorylated. Add the following reagents to the corresponding tube: 10 μl/Flap ligation. reaction (take 10 μl/flap ligation reaction above), 2.5 μl Optikinase buffer, 0.5 μl 100 mM ATP, 11.5 μl H_20, and 0.5 μl Optikinase enzyme. Incubate at 37°C for 1 hour. Product concentration 0.4 pmol/flap/μl.

[0389] Oligonucleotide Patch Unit Phosphorylation (Step C) 27 μl Oligonucleotide Patch Unit mix 0.74 pmol/μl, 5 μl 10x buffer, 1 μl 100 mM ATP, 3 μl Optikinase enzyme, and 14 μl H_20. Once reagents are all together gently mix the solution by flicking the tube a few times and spin down. Incubate at 37 °C for 1 hour.

[0390] Hybridization to M13mp18 scaffold (Step D) In a new 1.5 ml tube add the following reagents: 20 μl M13mp18 at 250 ng/μl, 27 μl Phosphorylated Oligonucleotide Patch Units 0.4 pmol/μl (Step C), 12.5 μl/ Phosph. Flap Ligation (Step B) preheat at 55 °C for 5 minutes and put on ice, 11 μl 10x ligase buffer and heat entire mixture at 55 °C for 1 minute. Hybridize mixture at 37 °C for at least 4 hours.

[0391] Ligation (Step E) Spin down eppendorf contents. Add 1.2 μl 100 mM ATP and 3 μl T4 ligase. Gently mix contents by flicking the tube, then spin down.

[0392] BamH1 Digest (Step F): I μl of 10 pmol BamH1 oligo, 20 μl 10x BamH1 buffer and hybride at 37 ~ 1 hour. Adjust volume to 200 μl. Add 3 μl BamH1 enzyme. Incubate at 37 °C for 1 hour.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 243

[0393] First step: start by adding 20 μl of M13mp18 (NEB 250 μg/ml) to a clean 1.7 ml eppendorf tube. Take 5 μl of Phosphorylated Flap ligation reaction and preheat it at 70 for 2 minutes and immediately put on ice. Add the 5 μl of each Phosphorylated Flap Ligation reaction (1 pmol/flap/μl) to the tube and gently mix by pipetting a few times. Incubate the eppendorf tube at 37 °C for 1 hour.

[0394] Second step: put 13.5 μl Oligonucleotide Patch Unit Mix (0.74 pmol/μl) and 1 μl of Acrydite Mix (10 pmol/μl) in a new eppendorf 1.7 ml eppendorf tube. Add 5 μl 10x Optikinase buffer, 1 μl 100 mM ATP and 27.5 μl H_20. Mix gently by pipetting the solution. Add 2 μl Optikinase enzyme, gently mix by pipetting and incubate at 37°C for 1 hr.

[0395] Third step: take the phosporylated oligos rxn and add it entirely to the contents of the M13mp18+Flaps Hybridization. The reaction is mixed gently by pipetting and it is allowed to incubate at 30 °C for 1 hour. After the hybridization is complete adjust the ATP by adding 1 μl (100 ATP) to the reaction.

[0396] Fourth step: spin down contents in eppendorf tube and add 4 μl T4 Ligase enzyme (5 units/μl), mix gently by pipetting. Incubate at room temperature for at least four hours. Add 1 μl BamH 1 oligonucleotide (10 pmol/μl) to hybridize at room temperature while ligation is taking place.

[0397] Fifth step: digest ligation reaction by adding 4 μl BamH1 enzyme (5 units/μl), mix gently by pipetting and incubate at 37°C for 1 hour. Once the incubation period is over. Take an aliquot of 500 ng for QC.

[0398] Sixth step: treat with Psoralen, UV or DMPA light for 15 minutes.

[0399] Calculations include:

$$5\ \mu g\ of\ M13 = 20\ \mu l\ stock\ from\ New\ England\ Biolabs = 2\ pmols$$

$$Oligonucleotide\ mix:\ 180\text{-}34\ flap\ areas - 10\ Acrydite\ modified\ Oligos = 0.74\ pmol/oligo$$

$$10\ pmols/oligonucleotide = 13.5\ \mu l = 1350\ pmols$$

[0400] Optikinase 1 unit converts 1 nmol of phosphate to ends - use excess. 4 μl of Optikinase was used.

[0401] SEQ ID NO: 1 = M13mp18.

8. EXAMPLE 3: PROTOCOL FOR PRODUCTION OF RNA NANOREPORTERS

[0402] Nanoreporters were generated and successfully employed to detect target molecules using methods substantially as described in this example. An example of target detection using such this method is shown in Figure 6.

8.1 SCAFFOLD PRODUCTION

[0403] Single-stranded circular M13mp18 DNA (USB Corporation) is annealed to a 10-fold molar excess of an oligonucleotide complementary to the Bam HI recognition site (Bam Cutter oligo) and cut with Bam HI restriction enzyme to yield a linear single-stranded DNA backbone. An oligonucleotide complementary to the Bam Cutter oligonucleotide (anti-Bam oligonucleotide) is subsequently added in 50-fold excess to the Bam Cutter oligonucleotide to sequester free Bam Cutter oligonucleotide and thus prevent recircularization of the M13 during later steps.

[0404] The linear M13 molecule serves as a scaffold onto which RNA patches, or RNA segments, with incorporated fluorophores can be annealed.

8.2 PCR TO FORM DOUBLE-STRANDED POSITIONS ON THE M13 SCAFFOLD

[0405] Ten sets of oligonucleotide primer pairs were designed to create 10 different regions along the M13 scaffold. Each pair contains one primer which has a T7 RNA polymerase promoter at the 5' end. Regions 2-7 are designed to be 900 bases (approximately 300 nm) long, as this is the approximate size of a diffraction-limited spot (the smallest spot that can be achieved with standard optics). Regions 1 and 8 have both long and short versions: the long versions cover the whole 900-base region, while the short versions cover only a portion of the 900-base region to allow a target-specific sequence to be ligated. Thus a target-specific sequence can be attached to either end. The ends can also be used for attachment of anchors or tags.

[0406] PCR is performed using Taq polymerase and 0.5 ng of double-stranded M13mp18 (USB Corporation) as a

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 244

template. Reactions are cleaned up using a Qiaquick purification kit from Qiagen. Each PCR reaction yields a double-stranded fragment corresponding to one specific segment as illustrated below. These fragments are used as templates for the *in vitro* transcription of the RNA segments.

8.3 *IN VITRO* TRANSCRIPTION TO PRODUCE DARK RNA SEGMENTS

[0407] Using the PCR products described above as double-stranded templates, RNA segments are generated using an *in vitro* transcription kit from Ambion (Megascript T7 kit). The products of the transcription reactions are purified (including treatment with DNAse I to remove template) using a RNeasy Kit from Qiagen.

8.4 *IN VITRO* TRANSCRIPTION TO PRODUCE RNA SEGMENTS MODIFIED WITH AMINOALLYL GROUPS

[0408] Using the PCR products described above as double-stranded templates, RNA segments for later dye-coupling are generated using an *in vitro* transcription kit from Ambion (MessageAmp aRNA kit). Aminoallyl-modified UTP nucleotides are incorporated into the RNA segments during transcription. The products of the transcription reactions are purified (including treatment with DNAse I to remove template) using a RNeasy Kit from Qiagen.

8.5 DYE COUPLING OF AMINOALLYL RNA SEGMENTS TO PRODUCE COLORED RNA SEGMENTS

[0409] 20-100 μg of aminoallyl-modified RNA segment is coupled with NHS-ester dyes using Ambion Aminoallyl Labeling Kit. Dyes used include Alexa 488, Alexa 594 and Alexa 647 (Invitrogen/Molecular Probes) as well as Cy3 (Amersham).

[0410] Each segment is made separately in 4 colors so that each position on the scaffold can be filled with a segment in any of the four colors; thus different colors can be added at different positions to create many unique color combinations.

[0411] In this particular embodiment, adjacent segments must be of different colors or there may be dark segments interspersed so that each segment is detected as an individual 'spot'. Dark segments may be used as part of the nanoreporter code.

8.6 ASSEMBLY OF THE LABEL MOLECULE

[0412] Segments for each position are annealed in a 2:1 ratio of segment to M13 scaffold in 1X SSPE buffer at 70°C for 2 hours.

[0413] An assembled nanoreporter with labeled RNA segments is depicted in Figure 3A-3B. Figure 3A depicts a nanoreporter in which only alternate "spots" (1, 3, 5 and 7) are labeled, and Figure 3B depicts a nanoreporter in which every spot is labeled.

9. EXAMPLE 4: DETECTION OF TARGET (S2) RNA AND DNA MOLECULES USING AN RNA NANOREPORTER/ GHOST PROBE COMBINATION

9.1 SYNTHESIS OF PROBE AND TARGET OLIGONUCLEOTIDES

[0414] S2 DNA target oligonucleotide was synthesized and purified by polyacrylamide gel electrophoresis (Integrated DNA Technologies). S2 RNA target molecules were generated by *in vitro* transcription of PCR products corresponding to region of cloned SARS coronavirus gene (Invitrogen) using an Ambion Megascript™ kit per manufacturer's instructions. The S2 ghost probe (Figure 6A (i)) was complementary to a specific 50-base region of the S2 target sequence (S2-*a*) and was synthesized with a biotin-TEG monomer at the 5' end and purified by high performance liquid chromatograpy (Integrated DNA Technologies). A second oligonucleotide with 50 bps complementary to the S2 target (S2-b) plus 9 bp of a additional sequence used for ligation to the M13 scaffold (59bp total) was synthesized and purified by HPLC (Integrated DNA Technologies). Note that S2-a and S2-b target regions were not overlapping.

9.2 NANOREPORTER SYNTHESIS

[0415] Oligonucleotide S2-b was ligated to the 5' end of linearized M13 [Figure 6A (iii)], and the resulting product was purified away from residual unligated oligonucleotide by size-exclusion filtration through a YM100 filter (Millipore) per manufacturer's instructions. Amino-allyl-modified RNA segments complementary to M13 is positions 2, 4, 6, and 8 (SEQ ID NOs:) (Figure 1C) were generated from *in vitro*-transcription of DNA templates (PCR products) via the Ambion Megascript™ kit per manufacturer's instructions. The segments were then coupled to NHS-ester-modified Alexa 647 dye (Invitrogen) per Ambion's instructions (amino allyl MessageAmp™ II aRNA kit). RNA segments corresponding to

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 245

positions 1, 3, 5, and 7 of the M13 scaffold (Figure 1C) were generated as unmodified *in vitro*-transcribed RNAs from DNA templates as described above. Assembly of the nanoreporter was carried out by annealing 10 fmol/μl of each of the eight segments to 5 fmol/μl of the M13-S1-*b* scaffold for 2 hours at 70°C in 1X SSPE buffer (150 mM sodium chloride, 10 mM sodium phosphate, 1mM EDTA). The final product was a nanoreporter with 4 segments labeled with A647 (red) interspersed with dark segments.

9.3 HYBRIDIZATION CONDITIONS

[0416] Hybridization of nanoreporters and ghost probes to target were carried out under the following conditions: 5X SSPE (750 mM sodium chloride, 50 mM sodium phosphate, 5 mM disodium EDTA), 40 pM ghost probe (attachment oligonucleotide S2-a), 40 pM Nanoreporter S2-*b*, 100 ng/μl sheared salmon sperm DNA, 5X Denhardt's solution and 0.1% Tween. Final target concentrations were 20 pM S2 DNA target (Figure 6B) and 1pM S2 RNA target (Figure 6C). No target was added to the negative control (Figure 6D). The hybridization reaction was incubated at 65°C for at least 16h.

[0417] Hybridization reactions were diluted 1:2 with 100 mM Borate buffer solution (pH 9.8) and introduced into a flow cell channel and bound to a streptavidin-coated coverslip forming the bottom of the channel (Streptavidin-OptiChem coverslips from AcceIr8). Attachment to the slide by one end of the nanorerporter/target/ghost probe complex was achieved via interaction of the biotinylated ghost probe with the streptavidin surface. After rinsing the channel with additional borate buffer to remove excess reporters not bound to the surface, the buffer was exchanged with 1X TAE (40 mM Tris-acetate, 1 mM EDTA) and a current of 200V was applied to stretch out the nanoreporter/target complexes during image capture.

[0418] Images were obtained using a Leica DMI 6000B microscope with a 63X oil immersion objective (1.4 NA), Xcite-120 light source (Exfo), customized filter sets (Chroma Technologies), an Orca-ER CCD camera (Hamamatsu) and Metamorph data acquisition software (Molecular Devices).

[0419] As predicted, when the correct target molecule S2 hybridizes [Figure 6A (ii)] to both ghost probe [Figure 6A (i), S2-*a*] and S2-*b*-target-specific nanoreporter [Figure 6A (iii)], the ghost probe/target/nanoreporter complex forms a single species that attaches to the slide and was visualized as 4 spots when exposed to 647 nm wavelength light (Figure 6B, 6C, and 6E). The amount of binding was dependent on the target concentration. There was no significant binding in absence of S2 target sequence (Figure 6D).

10. EXAMPLE 5: NANOREPORTER COMPRISING A MONOVALENT OR BIVALENT ANTIBODY FRAGMENT

[0420] Where a target molecule is a protein or polypeptide, a nanoreporter can be generated in which the nanoreporter scaffold is a nucleic acid and the target-specific sequence is a monovalent or bivalent antibody fragment.

[0421] Using routine methods, an antibody that recognizes a target molecule of interest is optinally digested with pepsin to generate F(ab')2 fragments. The two parts of the antibody or the two F(ab')2 fragments generated by the pepsin digestion are separated by mild reduction, for example with 2-mercaptoethylamine. This reduction separates either the antibody or the two F(ab')2 fragments into two monovalent fragments with two sulfhydryl groups that can be functionalized.

[0422] A heterobifunctional crosslinking reagent (*e.g.*, m-Maleimidobenzoyl-N-hydroxysuccinimide ester from Pierce Biotechnology Inc.) is used to attach a maleimide to an oligonucleotide with an amine modification (which can be ordered from many sources, such as Integrated DNA Technologies). The NHS on the cross-linking reagent is reacted with the amine on the oligonucleotides to produce a maleimide-conjugated oligonucleotide.

[0423] This maleimide conjugated oligonucleotide is then reacted with one of the sulfhydryl groups on the antibody fragment. Due to steric limitations, it is preferable that only one oligonucleotide be attached to each fragment.

[0424] This monovalent or bivalent antibody fragment attached to an oligonucleotide can then be hybridized to a complementary sequence on a nanoreporter scaffold, to generate a reporter probe in which the target-specific sequence is an antibody sequence. Such a reporter probe can be used alone to detect the target molecule, or in conjunction with a ghost probe or another reporter probe whose target-specific sequence is a monovalent or bivalent antibody or antibody fragment that binds to a different portion of the same target molecule.

11. EXAMPLE 6: HYBRIDIZATION OF 25 CELLULAR GENES TO 100 NG OF PLACENTAL TOTAL RNA USING NANOSTRING REPORTER SYSTEM

[0425] Detection and quantitation of 25 endogenous cellular genes was carried out in a single multiplexed hybridization reaction. In addition, three non-human control sequences were spiked into each reaction that corresponded to approximately 10, 100 and 300 copies per cell, respectively. A negative control hybridization was also performed in the absence of cellular RNA.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 246

11.1 HYBRIDIZATION REACTION

[0426] Each sample was hybridized in triplicate. Final concentrations of the hybridization reagents were as follows: 1.12nM total Nanoreporters (28 individual Nanoreporters at 40 pM each), 1.12nM total ghost probe (28 individual ghost probes), 5X SSPE (pH 7.5), 5X Denhardt's reagent, 100 ng/μl sheared salmon sperm DNA, 0.1% tween 20, 150 fM S3 spike DNA, 50 fM S4 spike, and 5 fM S6 spike. The final concentration of total placental RNA was 33 ng/μl. No total placental RNA was added to the negative control hybridizations. The final volume of the reaction was 30 μl. Reagents were mixed and incubated at 65°C in thermocycler block with heated lid for 20 hours.

Master mix	(1 Reaction)	(6 reactions)
1.8X hybridization mix*	16.7 μl	100 μl
25 endogenous gene reporters (0.6 nM each)	2 μl	12 μl
25 endogenous gene ghost probes (0.6 nM each)	2 μl	12 μl
Control reporters (0.6 nM each)	2 μl	12 μl
Control ghost probes (0.6 nM each)	2 μl	12 μl
10X control target mix	3 μl	18 μl
H_20	1.3 μl	8 μl
Total	29 μl	174 μl
*Hybridization mix (9X SSPE, 9X Denhardt's reagent, 180 ng salmon sperm DNA, 0.18% tween 20)		

Reactions	1	2	3	4	5	6
Master mix	29 μl	29 μl	29 μl	29 μl	29 μl	29 μl
100 ng/μl placental RNA	1 μl	1 μl	1 μl	0 μl	0 μl	0 μl
H_20	0 μl	0 μl	0 μl	1 μl	1 μl	1 μl
Total Rxn volume	30 μl	30 μl	30 μl	30 μl	30 μl	30 μl
Incubate reactions in thermocycler with heated lid overnight (18hours).						

11.2 POST-HYBRIDIZATION PURIFICATION

[0427] Hybridization reactions were purified to remove unhybridized reporters using an oligonucleotide complimentary to ghost probe attached to magnetic beads (F-bead). Hybridization reactions were diluted 5 fold in 0.1 % tween 20 to bring the final salt concentration to 1X SSPE and the solution added to 30 μl of F-beads (prewashed 2 times in 150 μl of 1X SSPE/0.1% tween 20). Hybridized complexes were allowed to bind to the beads at room temperature for 15 minutes with continuous rotation, washed once in 150 μl of 0.5X SSPE, and eluted in 25 μl of 0.1X SSPE for 15 minutes at 45°C.

11.3 BINDING, STRETCHING, AND IMMOBILIZATION

[0428] The samples was prepared for binding by addition of 1 μl of 1/1000 dilution of 0.1 uM Tetraspec™ fluorescent microspheres (product # T7279, Molecular Probes) and 3 μl of 1 M bis-tris propane (pH 9.0). Samples were loaded into a Nanostring fluidic device for attachment to Accer18 Optichem® slide coated with streptavidin (product #TB0200). After loading, slide surface was washed once with 1X TAE and prepared for electrostretching by addition of 40 μl of TAE to each well. Attached complexes were stretch by applying 200V across the fluidic channel. After 1 minute the samples were immobilized in the stretched position by adding 60 μl of 500 mM of G-hook oligo solution to the well containing the negatively charged electrode while continuing to apply voltage for 5 minutes. After immobilization the TAE solution is removed and replaced with anti-photobleaching reagent for imaging.

11.4 IMAGING

[0429] Slides were imaged on Nikon Eclipse TE2000E equipped with a metal halide light source (X-cite 120, Exfo Corporation) and a 60X oil immersion lens (1.4 NA Plan Apo VC, Nikon). For each field of view, 4 images at different excitation wavelengths (480, 545, 580 and 622) were acquired with an Orca Ag CCD camera (Hamamatsu) under control

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 247

of Metamorph software (Universal Imaging Corporation). Images were processed with custom image processing software.

11.5 DATA ANALYSIS

[0430] Raw data were extracted from processed images using custom software. Data where normalized to the average counts for control spikes in each sample. To determine if a gene was "detected" by the system, the counts obtained for each gene from hybridizations containing RNA were compared to counts obtained in hybridizations without RNA using a Student's ttest. Genes with p values <0,05 were determined to be detected. After background subtraction, the concentrations of cellular mRNA were estimated from the linear regression of the spike controls. These concentrations were converted to copies per cell using the following assumptions: 1 cell contains 10 pg total RNA; each cell contains 300,000 mRNA molecules; final volume of the reaction is 30 μl.

11.6 RESULTS AND CONCLUSION

[0431] Table 3 below shows the results of the data analysis described above. These results show that using the nanoreporter technology described herein, it was possible to detect transcripts, such as CASP3, that are present at a concentration of less than 1 transcript/cell. Thus, the nanoreporter technology provides an exquisitely sensitive means of detecting and quantifying gene expression.

Table 3: Transcript Concentration and Abundances

Gene	Avg counts*	error (counts)	concentration (fM)	error (conc.)	calculated copies/cell	error (copy/ cell)	Detected/ Not Detected (p<0.05)
GM2A	149	17	3.39	0.39	6.12	0.07	D
ATF4	68	2	1.55	0.06	2.80	0.01	D
CTNNB1	792	50	17.95	1.19	32.44	0.22	D
IRF1	221	20	5,01	0.47	9.05	0.09	D
STAT5A	120	11	2.72	0.25	4.91	0.05	D
CREG1	409	17	9.28	0.44	16.76	0.08	D
CASP3	13	1	0.30	0.03	0.54	0.00	D
CCL20	2	1	0.04	0.03	0.07	0.01	ND
NMI	115	2	2.61	0.07	4.72	0.01	D
XBP1	719	46	16.30	1.10	29.45	0.20	D
PCGF4	75	18	1.70	0.40	3.08	0.07	D
IFI27	747	41	16.94	1.00	30.61	0.18	D
TAF7	185	11	4.19	0.26	7.57	0.05	D
OAS3	74	9	1.68	0.20	3.03	0.04	D
C2	850	49	19.28	1.19	34.83	0.21	D
IL6	8	3	0.19	0.07	0.34	0.01	D
MyD88	94	6	2.13	0.14	3.85	0.03	D
HIF1A	130	7	2.95	0.17	5.33	0.03	D
APOA2	-1	2	-0.01	-0.05	-0.03	-0.01	ND
KISS	6825	130	154.79	4.52	279.65	0.82	D
ELK3	55	4	1.25	0.09	2.27	0.02	D
CBF2	72	3	1.64	0.07	2.96	0.01	D

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 248

(continued)

Gene	Avg counts*	error (counts)	concentration (fM)	error (conc.)	calculated copies/cell	error (copy/ cell)	Detected/ Not Detected (p<0.05)
IF130	625	47	14.16	1.10	25.59	0.20	D
RELB	35	5	0.78	0.11	1.42	0.02	D
CTCF	103	3	2.35	0.09	4.24	0.02	D
*Normalized and background subtracted.							

[0432] The hybridization methods described herein have been performed in single multiplexed reactions containing up to 120 different reporters with similar hybridization efficiencies and results.

12. EXAMPLE 7: CONSIDERATIONS REGARDING NANOREPORTER HYBRIDIZATION KINETICS

12.1 BACKGROUND

[0433] Solution hybridizations with a large excess of probe over target follow pseudo-first order kinetics. In this regime the speed of the reaction depends only on the probe concentration and not on the target concentration. For a two-probe, one-target strategy to provide accurate information on the concentration of a target in solution, the probes should both be present in excess of the target. The possible concentration range is preferably therefore bounded on the lower end by the concentration of the target. However, the useful concentration range for the nanoreporter technology described herein is practically bounded on the lower end by the amount of time needed to perform the hybridization.

12.2 HYBRIDIZATION KINETICS

[0434] In preferred embodiments, target detection and quantification assays are performed in which the target (T) must hybridize to both a reporter probe (R) and a ghost probe (G) to be detected (for example by affinity selection and detection of complexes comprising only (R) and (G), which in turn only form complexes in the presence of (T)). Assuming that these reactions are irreversible, there are four possible elementary reactions that occur.

$$R + T \xrightarrow{k_1} RT$$

$$T + G \xrightarrow{k_2} TG$$

$$RT + G \xrightarrow{k_3} RTG$$

$$R + TG \xrightarrow{k_4} RTG$$

[0435] Because RT and TG are intermediate complexes of two out of the three species, these four reactions can be simplified to

$$R + T + G \rightarrow RTG$$

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 249

[0436] However, to quantitatively calculate the rate of production of RTG (the reporter-target-ghost probe complex), all four reactions must be considered. The differential equations describing the system are:

$$\frac{dC_G}{dt} = -k_2 C_G C_T - k_3 C_G C_{RT} \quad \frac{dC_R}{dt} = -k_1 C_R C_T - k_4 C_R C_{TG} \quad \frac{dC_T}{dt} = -k_2 C_G C_T - k_2 C_R C_T$$

$$\frac{dC_{TG}}{dt} = k_2 C_G C_T - k_4 C_R C_{TG} \quad \frac{dC_{RT}}{dt} = k_1 C_R C_T - k_3 C_G C_{RT}$$

$$\frac{dC_{RTG}}{dt} = k_4 C_R C_{TG} + k_3 C_G C_{RT}$$

where C_R, C_T, C_G, C_{RT}, C_{TG}, and C_{RTG} are the concentrations of the various species, and kit-k_4 are the kinetic constants for the four elementary reactions. Values for these kinetic constants when the probes and targets are complementary single-stranded molecules (i.e. when there is no purification tag on the ghost probe and no reporter) can be calculated from data available in the literature (Wetmur, J. Annu. Rev. Biophys. Bioeng. 1976.5:337-361).

$$k = k_N \frac{\sqrt{L}}{N} \frac{\alpha_{salt}}{\alpha_{ref}}$$

[0437] In the above equation, k_N is the nucleation rate constant, L is the nucleic acid length (in base pairs), N is the nucleic acid complexity (equal to L for non-repetitive sequences) and a_{salt} and a_{ref} are corrections for salt concentration (Britten et al., 1974, Methods in Enzymology 29E:363-406). In the nanoreporter systems described herein, the kinetic constants will depend on the sizes of the attached ghost probe tags and reporter probe. Without being bound by any theory, it is the inventors' belief that the kinetic constants will have the same dependence on length that an elementary reaction has on the diffusion constants of the reactants.

$$k = k_N \frac{\sqrt{L}}{N} \frac{\alpha_{salt}}{\alpha_{ref}} \frac{D_1 + D_2}{2D_{50}}$$

[0438] In the above equation D_1 and D_2 are the diffusion constants of the two reacting species (see the reactions above) and D_{50} is the diffusion constant of a 50-mer single-stranded DNA molecule. Assuming a 100-base single-stranded target, 100-base single-stranded ghost probe, and 7200-base double stranded reporter, the relevant kinetic constants are

$$k_1 = 2.64 \times 10^5 \text{ L/mol/s}$$

$$k_2 = 6.55 \times 10^5 \text{ L/mol/s}$$

$$k_3 = 3.99 \times 10^5 \text{ L/mol/s}$$

$$k_4 = 1.91 \times 10^5 \text{ L/mol/s}$$

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 250

[0439] Numerically solving the system of differential equations with these kinetic constants (assuming at least a 10-fold excess of probes over target) yields the prediction that 5 pM reporter and 5 pM ghost probe will drive hybridization to 10% of completion in an overnight reaction (16-18 hours). At concentrations lower than 5 pM, the amount of completely hybridized molecules is likely impractical to measure. Thus, in a preferred embodiment, the lower concentration of a nanoreporter component (ghost probe and/or reporter probe) is 5pM.

12.3 ENTANGLEMENT OF REPORTERS

[0440] As probe concentrations increase, theory predicts that hybridization kinetics speed up without bound - the only limit being the solubility of the probes. However, the reporter probe can be very large compared to the target-specific sequence in the nanoreporter systems of the invention. Without being bound by any theory, the inventors believe that by its attachment to the reporter probe the kinetics of the target-specific sequence are altered from classical solution hybridization kinetics. Because the reporter probe is a large, polymeric molecule, it can have long-lived interactions (entanglements) with other nanoreporters when they come into contact. At low concentration the probability of two polymers becoming entangled is small, but as the concentration and/or size of a polymer in solution increases, these interactions become more and more common. In the extreme case of very long molecules at very high concentration the polymers form a permanent network, or gel, in solution. For solution hybridization to occur, a probe (*e.g.*, a nanoreporter probe)/target pair must diffuse through solution until they contact one another and a hybridization nucleus forms. Classically, hybridization reactions are not diffusion limited because the translational diffusion of the molecules is faster than the nucleation of the hybridization (*i.e.*, the probe and target diffuse together and interact many times before a nucleation occurs). In dilute solution its large size will slow the translational diffusion of the reporter probe, but may not significantly affect the kinetics. At some intermediate concentration, the reporter probes take up almost all of the space in the solution, effectively forming a permanently entangled gel, and can no longer diffuse in solution. However, the ghost probe and the targets are smaller molecules that are believed to still diffuse through the entangled reporter probes, allowing hybridization to take place (although possibly at a slower rate). The inventors also believe that at some higher concentration the reporter probe in solution will also hinder the movement of the ghost probe and the targets to the point that the reaction becomes diffusion limited. This concentration (which is not quantitatively known and depends upon the reporter probe structure, the ghost probe structure, and the target size) is the upper limit of the useful concentration range in the nanoreporter system, and can be emprically determined by one of skill in the art guided by the principles described herein.

12.4 LENGTH DEPENDENCE OF KINETICS

[0441] Since the limiting upper concentration for hybridization depends upon both the reporter structure and ghost probe structure (of which there are many possible variations), a theoretical framework to predict the permutations of useful concentration ranges is useful in the practice of the invention Classical theory predicts that hybridization kinetics depend only on the size of the smaller probe. Theory would therefore predict that the size of the reporter will not play a role in the hybridization kinetics as long as both the target molecule and the ghost probe are significantly smaller. Theory then predicts that the rate of hybridization (for a constant target length) depends on $1/L^{1/2}$, where L is the length of the ghost probe, due to steric inhibition of hybridization. Consequently, the kinetics of hybridization will be faster with smaller ghost probes. As the ghost probe length increases, the hybridization rate should decrease as $1/L^{1/2}$. If a constant ghost probe length is assumed, then the range of reporter lengths and concentrations that will result in a measurable mount of hybridization events can be defined. Once a reporter size has been defined, then the approximate range of ghost probe sizes can be determined. This is an iterative process, but may give good starting points from which to gather data to generate detailed empirical guidelines, given that the theories that the inventors' rationale is based upon were generated from hybridization data in systems that do not employ a reporter probe.

12.5 ENTANGLEMENT THRESHOLD

[0442] A reporter probe is essentially a polymer in free solution, which behaves as a random coil. The volume occupied by a single reporter, Vp, can be calculated from polymer physics theories according to the Freely-Jointed Chain model (FJC, for a flexible polymer, such as single-stranded DNA or RNA) or the Worm-Like Chain model (WLC, for a stiff polymer such as double-stranded DNA or a reporter). For either model

$$V_p = \frac{4}{3}\pi R_g^3$$

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 251

where Rg is the radius of gyration. For the FJC

$$R_g = b\left(\frac{N}{6}\right)^{0.6}$$

Where b is the segment length and N is the number of segments in the chain. For the WLC

$$R_g = \sqrt{\frac{1}{6}Nb^2 - \frac{b^2}{4} + \frac{b^2}{4N}\left(1 + \frac{1}{2N}\left(e^{-2N} - 1\right)\right)}$$

The entanglement threshold concentration is defined as the concentration where the entire volume of the solution is occupied by the reporters.

$$C^* = \frac{3}{4\pi R_g^3 N_A}$$

where N_A is Avogadro's number. Above this concentration it is assumed that the translational diffusion of the reporters is severely restricted. The entanglement threshold concentration varies with the reporter structure. As the reporter length increases, the entanglement threshold decreases (as $1/L^{1.5}$). From the equations above, the theoretical entanglement threshold for reporter probes with different spot sizes and different lengths can be calculated. The result of such calculations is shown in Figure 17, which shows that for a 7200 bp RNA/DNA hybrid reporter probe with 8 label attachment regions of about 900 bp each, the entanglement threshold is about 70 nM.

[0443] If both the target and the ghost probe are much smaller than the reporters, then they will most likely be free to diffuse through the solution even at these high concentrations of reporters. Initial data indicates that hybridization kinetics do not slow appreciably up to a concentration of 80 nM with a 7200-bp reporter probe, a 100-base target, and a 100-base ghost probe.

12.6 EFFECT OF ENTANGLEMENT THRESHOLD ON MULTIPLEXING

[0444] Assuming that the maximum concentration for reporters in a hybridization reaction is C*, then the concentration of each reporter (specific to a particular target) is equal to C*/M, where M is the multiplex of the reaction (number of different targets being addressed simultaneously). Conversely, the possible multiplex level for a particular reporter structure can be calculated from the lower limit of probe concentration (Cp from kinetics ~10 nM) and the entanglement threshold

$$M = \frac{C^*}{C_P}$$

[0445] If the number of nanoreporter codes available does not depend on reporter probe size, then the multiplexing of the nanoreporter depends primarily on the reporter probe size and concentration (since it is much larger than the ghost probe). Because the ghost probe makes an insignificant contribution to entanglement during hybridiziation, it is the inventors' belief that the concentration of the ghost probe can be increased far above the concentration of the reporter probe. In Table 4 below, the maximum total ghost probe concentration ([G]) is set to 1000 nM for all reporter concentrations. This difference in concentration of ghost probe and reporter probe is an adjustable parameter. Preliminary experiments show that in a multiplex hybridization reaction with a 7200 bp reporter and 100b ghost, 40pM of each reporter probe and 200 pM of each ghost probe results in near complete hybridization in an overnight reaction.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 252

12.7 OPTIMAL SIZE AND CONCENTRATION RANGES

[0446] Below in Table 4 is a summary of the optimal useful size and concentration ranges of the ghost probe and reporter probe at different multiplexing as approximated by the above theories. It is the inventors' belief that ghost probes up to about 200 bases will be practical for most applications.

Table 4: Optimal size and concentration ranges of reporter probe, ghost probe and target, as well as multiplicity of probes, in the nanoreporter systems of the invention.

Reporter Length (bp)	Ghost Length (b)	Minimum [R] (pM)	Minimum [G] (pM)	Maximum [R] (nM)	Maximum [G] (nM)	Max Multiplex
2000	100	5	5	603	1000	114417
2000	50	4	4	603	1000	161811
2000	200	7	7	603	1000	80905
3000	100	6	6	292	1000	45182
3000	50	5	5	292	1000	63897
3000	200	9	9	292	1000	31948
4000	100	7	7	178	1000	23912
4000	50	5	5	178	1000	33817
4000	200	11	11	178	1000	16908
5000	100	8	8	123	1000	14746
5000	50	6	6	123	1000	20854
5000	200	12	12	123	1000	10427
6000	100	9	9	91	1000	9988
6000	50	6	6	91	1000	14125
6000	200	13	13	91	1000	7062
7200	100	10	10	68	1000	6792
7200	50	7	10	68	1000	6792
7200	200	14	10	68	1000	6792
8000	100	11	11	57	1000	5444
8000	50	7	7	57	1000	7699
8000	200	15	15	57	1000	3850
10000	100	12	12	40	1000	3419
10000	50	8	8	40	1000	4835
10000	200	17	17	40	1000	2417

13. EXAMPLE 8: EXEMPLARY EMBODIMENTS FOR DUAL NANOREPORTER ASSEMBLY

[0447] This section describes an embodiment for assembly of a dual nanoreporter in which one probe is a ghost probe and the other probe is a reporter probe comprising color RNA segments assembled on an M 13 backbone. The ghost probe is attached to a biotinylated F-hook and the reporter probe is attached to a biotinylated G-hook. The dual nanoreporter is hybridized to a biomolecular sample to detect and quantify a target molecule. The steps below do not have to be performed in the order presented. Moreover, each particular step represents a specific embodiment that may be combined with embodiments other than those presented below.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 253

13.1 PREPARATION OF THE M13 SCAFFOLD

[0448] Single-stranded circular M13mp18 DNA (USB Corporation) is annealed to a 5-fold molar excess of an oligonucleotide complementary to the Bam HI recognition site (Bam Cutter oligo) and cut with Bam HI restriction enzyme to yield a linear single-stranded DNA backbone. An oligonucleotide complementary to the Bam Cutter oligonucleotide (anti-Bam oligonucleotide) is subsequently added in 50-fold excess to sequester free Bam Cutter oligonucleotide and thus prevent recircularization of the M13 during later steps.

[0449] The linear M13 molecule serves as a scaffold onto which RNA patches, or RNA segments, with incorporated fluorophores can be annealed.

13.2 ATTACHMENT OF A TARGET-SPECIFIC SEQUENCE TO THE SCAFFOLD

[0450] An oligonucleotide comprising a sequence (of, *e.g.*, 30-70 nucleotides) complementary to the target nucleic acid of interest, plus 9 bp of additional sequence used for ligation to the M13 scaffold, is generated and ligated to the 3' end of the linearized M13 scaffold.

13.3 ATTACHMENT OF G-TAGS TO THE SCAFFOLD

[0451] A G-tag (*e.g.*, an oligonucleotide having the sequence 5'-AACATCACACAGACC AACATCACACAGACC AACATCACACAGACC AACATCACACAGACC AGCCCTTTG-3', which includes 4 copies of the complement of the G-hook 5'-GGTCTGTGTGATGTT-3', followed by 9 bases of ligator sequence, and which is complementary to the G-hook) is attached to the 5' end of the linearized single-stranded M13 backbone to allow for (1) purification of the reporter following ligation and/or annealing of segments; and (2) immobilization of the reporter once it is "stretched" on a solid surface. The sequence of the ligator for attaching G-tag to the 5' end of single-stranded M 13 which has been linearized at the BamHI site can be 5'-CTCTAGAGGATCCAAAGGGCT-3'. The ligation reaction can be performed according to the following protocol to produce approximately 80 pmol of G-tag/M13 ligation product:

Materials:

[0452]

[100uM] anti-G4 tag oligo
[100uM] anti-G4 tag ligator oligo
[80 nM] Linear single-stranded M13
10X T4 DNA Ligase Buffer (Fermentas)
T4 DNA Ligase (Fermentas)
20X SSC (Ambion)
DEPC H2O (Ambion)

Method:

[0453]

1. Pre-anneal the G-tag and ligator:

25uM 2:1 G/Glig in 1X SSC
20 μl [100uM] G-tag Ligator
40 μl [100uM] G-tag
4 μl 20X SSC
16 μl DEPC H2O

* Anneal on the MJ Thermocycler
95°C, 3min; 72°C, 30sec, -1°C/cycle, x 68 cycles; hold at 4°C
2. Ligate the G-tag to the linear M13:

64nM M13-G4 in 1X Lig Buffer
1000 μl [80 nM] Linear M13
80 μl [25uM] 2:1 G/Glig in 1X SSC

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 254

124 μl 10X T4 DNA Ligase Buffer
40 μl T4 DNA Ligase

* Ligate in an aluminum heat block covered with foil at 37°C for 2hr then at 65°C for 15 minutes to inactivate the enzyme.

13.4 PREPARATION OF RNA SEGMENTS

[0454] Ten sets of oligonucleotide primer pairs are designed to create 10 different regions along the M13 scaffold. Each pair contains one primer which has a T7 RNA polymerase promoter at the 5' end. Regions 2-7 are designed to be 900 bases (approximately 300 nm) long, as this is the approximate size of a diffraction-limited spot (the smallest spot that can be achieved with standard optics). Regions 1 and 8 have both long and short versions: the long versions cover the whole 900-base region, while the short versions cover only a portion of the 900-base region to allow a target-specific sequence to be ligated. Thus a target-specific sequence can be attached to either end. The ends can also be used for attachment of anchors or tags.

[0455] PCR is performed using Taq polymerase and 0.5 ng of double-stranded M13mp18 (USB Corporation) as a template. Reactions are cleaned up using a Qiaquick purification kit from Qiagen. Each PCR reaction yields a double-stranded fragment corresponding to one specific segment as illustrated below. These fragments are used as templates for the in vitro transcription of the RNA segments.

[0456] Using the PCR products described above as double-stranded templates, RNA segments are generated using an in vitro transcription kit from Ambion (Megascript T7 kit). The products of the transcription reactions are purified (including treatment with DNAse I to remove template) using a RNeasy Kit from Qiagen.

13.5 LABELING OF THE RNA SEGMENTS

[0457] Using the PCR products described above as double-stranded templates, RNA segments for later dye-coupling are generated using an in vitro transcription kit from Ambion (MessageAmp ARENA kit). Aminoallyl-modified UTP nucleotides are incorporated into the RNA segments during transcription. The products of the transcription reactions are purified (including treatment with DNAse I to remove template) using a RNeasy Kit from Qiagen.

[0458] 20-100 μg of aminoallyl-modified RNA segment is coupled with NHS-ester dyes using Ambion Aminoallyl Labeling Kit. Dyes used include Alexa 488, Alexa 594 and Alexa 647 (Invitrogen/Molecular Probes) as well as Cy3 (Amersham).

[0459] Each segment is made separately in 4 colors so that each position on the scaffold can be filled with a segment in any of the four colors; thus different colors can be added at different positions to create many unique color combinations.

[0460] In this particular embodiment, adjacent segments are of different colors or there may be dark segments interspersed so that each segment is detected as an individual 'spot'. Dark segments may be used as part of the nanoreporter code.

13.6 ANNEALING OF THE RNA SEGMENTS TO THE SCAFFOLD

[0461] Segments for each position are annealed in a 2:1 ratio of segment to M 13 scaffold in 1X SSPE buffer at 70°C for 2 hours. An assembled nanoreporter with labeled RNA segments is depicted in Figure 3A-3B. Figure 3A depicts a nanoreporter in which only alternate "spots" (1, 3, 5 and 7) are labeled, and Figure 3B depicts a nanoreporter in which every spot is labeled.

13.7 PREPARATION OF THE GHOST PROBE

[0462] One or more oligonucleotides comprising sequences (of, *e.g.*, 30-70 nucleotides) complementary to different regions of the target nucleic acid(s) of interest than those to which the target-specific sequences of the reporter probe are complementary, are generated. Optionally, F-tags for F-hook attachment are ligated to the 5' end of the ghost probe using a a ligator oligonucleotide that is complementary to a short sequence on the 3' end of the F-hook as well as a short sequence on the 5' end of the ghost probe. The sequences that are complementary to the ligator oligonucleotide are not part of the F-hook sequence or the probe sequence, but are additional nucleotides added to those oligos in order to facilitate ligation.

13.8 ATTACHMENT OF F-TAGS TO THE GHOST PROBE

[0463] An F-tag (*e.g.*, an oligonucleotide having the sequence 5'-GATGGAGAC GTCTATCATCACAGC GTCTATCATCACAGC-biotin -3', which includes 2 copies of the complement of the F-hook 5'-GCTGTGATGATAGAC -3', followed

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 255

by 9 bases of ligator sequence and is complementary to the F-hook) is attached to the 3' end of the ghost probe to allow for (1) purification of the ghost-probe-target-reporter hybridization complex; and (2) attachment of the hybridization complex on the slide via the biotin moiety. The sequence of the ligator for attaching F-tag to the 3' end of the ghost probe can be 5'-GTCTCCATCTTCCGACAG-3'.

Materials:

[0464]

100uM F-biotin tag
100uM F ghost probe ligator
Fermentas 10X T4 DNA Ligase Buffer
1 uM ghost probes
Fermentas T4 DNA Ligase

Method:

[0465]

1. Pre-anneal the hook and ligator:

5uM F-biotin tag /ligator mix
5 µl [100uM] F-biotin tag
5 µl [100uM] F-ghost probe ligator
10 µl 10X T4 DNA Ligase Buffer
80 µl DEPC H_2O

Anneal on the MJ Thermocycler (95°C, 3min; 72°C, 30sec, -1°C/cycle x 68 cycles; hold at 4°C).

2. Set up the following ghost probe ligation:

300 nM anti-F2-biotin -GP
6.0 µl [1uM] Ghost Probe
4.8 µl [5uM] anti-F2-biotin tag /ligator mix
1.52 µl 10X T4 DNA Ligase Buffer
3.68 µl DEPC H_2O
4.0 µl T4 DNA Ligase

Ligate on the MJ Thermocycler (37°C, 18her; 65°C, 15 minutes; hold at 4°C)

3. QC the ligation on a 15% Novex TBE-Urea gel:

Prepare the following loading solutions:

Ligation	Neg Control-Ghost Probe
3.33 µl [300 nM] ligation	1 µl [1uM] ghost probe
1.67 µl DEPC H_2O	0.33 µl 10X T4 DNA Ligase Buffer
5 µl 2X Loading Buffer	3.67 µl DEPC H_2O
	5 µl 2X Loading Buffer

Neg Control- F-biotin tag /ligator mix
2 µl [0.5uM] F-biotin tag /ligator mix
0.33 µl 10X T4 DNA Ligase Buffer
2.67 µl DEPC H_2O
5 µl 2X Loading Buffer

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 256

(continued)

50 bp Oligo Ladder
4 μl Ladder
6 μl 2X loading buffer

Run of a 15% Novex TBE-Urea gel at 180V for 50 minutes.
Stain with SYBR Gold for 30 minutes.

13.8.1 ALTERNATIVE EMBODIMENTS

[0466] Rather than covalently coupling biotin to the single-stranded F-tag, the biotinylation of the ghost probe can also be accomplished by annealing a biotinylated oligonucleotide (DNA or RNA) with a sequence complementary to the common portion of the ghost probe. Such a sequence could be the F sequence itself, or another sequence which is added to the ghost probe in addition to the F sequence. If such an additional sequence is added, it could be from 10-100 bases long, from 1-10 copies, with the preferred configuration being a single copy from 50-100 bases long.

13.9 BIOTINYLATION OF TARGET MRNA

[0467] There are a number of commercially available kits available for the direct labeling of an mRNA sample including *Label* IT® $\mu Array_{TM}$ ^{Biotin} Mirus #MIR 8010) and Biotin-Chem-Link (Roche (1 812 149). Following manufacturer's procedures biotin labeled mRNA is added to the hybridization reaction as described in Section 3d (below) with the following modifications: Since most protocols suggest the use of poly A+ mRNA, the amount of RNA used could be reduced below the 100 ng total RNA in a typical hybridization to 10 ng and possibly lng. No ghost probe should be added to this reaction. F bead post-hybridization purification is no longer required. G-bead post-hybridization purification should be used to remove unhybridized biotinylated mRNA that might compete for binding to the slide. Depending on the amount of RNA used, this may or may not be required. Alternatively, total RNA could be biotinylated without the need for purification of the poly A+ fraction. In this case, the original amount of total RNA should be used (100 ng). The use of total RNA might require modifications of the manufacturer's protocol to increase labeling efficiency.

[0468] An alternative approach would be to enzymatically generate biotinylated 1st strand cDNA or biotinylated amplified RNA (aRNA) using commercially available kits and use these in place of total or mRNA. This approach would require a redesign of the reporter probes to be in the sense orientation. Both ghost probe and F-bead post-hybridization reactions would be omitted while G-bead purification would remain for removal of non-hybridized RNA.

13.10 HYBRIDIZATION OF DUAL NANOREPORTER TO TARGET

[0469] Many hybridization conditions are sufficient for achieving gene expression data. To shorten hybridization times while maintaining reasonable hybridization efficiency, several parameters can be altered: i) increasing ghost probe and reporter concentrations, ii) fragmenting of total RNA to average size range of 200-500 bp while lowering the pH of hybridization to 6.5, iii) using more total RNA in same hybridization volume, iv) lowering hybridization volume to approximately 10 μl. Blocking reagents such as Denhardt's and ssDNA can be removed without deleterious effects on hybridization efficiency or cross hybridization to mRNAs from different species.

[0470] The following protocol has been performed successfully with multiplexing from 1 to >500 nanoreporters with ghost probes (an example demonstrating a nanoreporter assay utilizing 25 nanoreporters is described in Example 6 above, and another example demonstrating a nanoreporter assay utilizing 509 nanoreporters is described in Example 9 below). The final concentration of all nanoreporters varies depending on 1) the concentration of each reporter and 2) the number of genes being multiplexed.

[0471] Typical total nanoreporter concentrations range from 40pM (1 gene @ 40pM) to 20nM (500 genes @40pM). Ghost probe concentrations also vary from 200pM (1 gene @ 200pM) to 100nM (500 genes @ 200pM). The example that follows describes a single multiplexed hybridization containing approximately 500 endogenous genes with positive and negative controls. Add, 11.1 μl of 2.7X hybridization mix [13.5X SSPE pH 7.5 (USB #75890), 0.27 μg/ μl sheared salmon sperm DNA (Sigma #D-7656), 0.27% tween 20 (Sigma #P-1379), and 13.5X Denhardt's reagent (Sigma D-2532)], 5 μl ofgene Nanoreporter mix (0.24nM each or 123nM total, includes 509 endogenous genes and 8 hybridization controls), 4.6 μl 513 gene ghost probe mix (1.3nM each or 667 nM total, includes 509 endogenous genes and 8 hybridization controls), 1 μl of purification control reporter mix (0.5pM), 1 μl of total cellular RNA (100ng/ul), 1 μl of 30X spike target mix (1.5nM-3fM) and 6.3 μl of DEPC treated water (Ambion #9922) to a 0.2ml thin wall tube (final volume 30 μl).

[0472] Final concentration of hybridization reagents should be 5X SSPE, 0.1% tween 20, 100ng/μl sheared salmon

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 257

sperm DNA, 5X Denhardt's reagent, 40 pM each Nanoreporter (~20nM total), 200pM each ghost probe (~100nM total) and 33ng/μl of total cellular RNA. Control spike targets typically vary in range from 50 M down to 0.1 fM in a single reaction. All reagents are most preferably free of all nuclease activity. For optimal results, all reagents should be free of nuclease activity.

[0473] Mix reagents well and incubate in temperature block with heated lid for 20 hours. After hybridization purify the nanoreporters with affinity reagents for both the ghost probe and the reporter probe.

13.10.1 ALTERNATIVE EMBODIMENT: HYBRIDIZATION PROTOCOL FOR WITHOUT SSDNA AND DENHARDT'S REAGENT

[0474] This protocol has been performed successfully with multiplexing from 1-500 nanoreporters and ghost probes. Removal of ssDNA and Denhardt's reagent from hybridizations performed with human reagents (Nanoreporters and ghost probes) had no effect on cross hybridization with mouse total RNA when compared to a hybridization containing ssDNA and Denhardt's. In addition, removal of ssDNA and Denhardt's does not result in an increased background signal (based on negative hybridization controls). Finally, there is no significant loss (or gain) of signal for endogenous genes hybridized in the presence or absence of ssDNA and Denhardt's (509 genes, R^2 value=0.998).

13.10.2 ALTERNATIVE EMBODIMENT: HYBRIDIZATION CONDITIONS FOR FRAGMENTED CELLULAR MRNA

[0475] Fragmentation of cellular RNA has been achieved by both thermal and cation catalyzed protocols. These protocols were designed to obtain fragment lengths between 100 and 700 bp (on average). Thermal fragmentation: Dilute total RNA sample to 200 ng/μl in RNAse free water. Heat sample to 95°C in temperature block with heated lid. Stop fragmentation by placing sample on ice. Use immediately or store at -80°C until use. Fragmentation via cation catalyzed reaction modified from manufacturer's protocol. (Ambion). Bring volume of RNA sample up to 9 μl with RNAse free water. Final concentration of total RNA should be between 0.2 and 2 μg/ml. Add 1 μl of 10X fragmentation buffer (Ambion10X fragmentation buffer). Incubate at 70°C for 5 minutes in temperature block. Longer times will result in smaller fragment size on average. Stop reaction by addition of 1 μl 200 mM EDTA. Use immediately or store at -80°C until use.

[0476] Fragmented RNA samples are hybridized as described herein except for the following modifications: i) pH of SSPE is reduced to 6.5 and ii) the time of reaction is reduced to 6hours (for hybridization reactions in which reporter probe and ghost probe concentrations are 200 pM).

13.11 PURIFICATION OF NANOREPORTER-TARGET COMPLEXES

[0477] Post-hybridization purification is preferred when the total reporter probe concentration is above 1 nM. Purification significantly decreases non-specific binding and increases specific binding efficiency to the slide at higher reporter and ghost probe concentrations. In the example provided above, a single F-bead purification is described (purifies hybridized complexes from the ghost-probe end). As described in Example 9 below, optimal results at high ghost probe concentrations (>5nM total) are obtained via a subsequent G-bead purification which purifies the hybridization complexes from the 5' end of the reporter effectively removing excess non-hybridized ghost probes. The preferred order of purification is F-bead, then G-bead but the order can be reversed and the protocols optimized accordingly. The exact sequences used in these affinity purifications can likely be changed and optimized in alternative embodimentsof the technology. These affinity purification steps and reagents are currently nucleic acid based but could theoretically be any sort of binding pairs that exhibit specific binding to one another and can be released by chemical treatment or alteration of binding conditions such that the interaction is disrupted and released. For example, an antibody/antigen pair, a protein/metal interaction, or ligand/receptor interaction, *etc.*

[0478] One example of purification is provided below.

[0479] After hybridization is complete, the salt of a hybridization sample (30 μl, starting at 5x SSPE = 825mM Na^+) is adjusted to a final concentration of approximately 1x SSPE. The diluted sample is added to 30 μl F-hook MyOne Dynabeads (F-MODB) and bound for 15 minutes at room temperature while rotating. The beads are sequestered with a magnet and the supernatant removed. The beads are washed twice with 150 μl 0.1xSSPE + 0.1% Tween at room temperature for 15 minutes w/ rotation and discarded. The purified reporters are eluted in 30 μl 0.1x SSPE at 45 °C for 15 minutes w/ rotation. At this point the hybridized reporters are purified from the contaminating un-hybridized reporters. The elution still contains contaminating un-hybridized ghost probes which will compete with the reporters for biotin-binding sites on the streptavidin coated slide. The 30 μl is added to 130 μl of 1xSSPE + 0.1% Tween to increase salt concentration. The sample (150 μl) is then loaded onto 30 μl of G-MODB and bound for 15 min at room temperature. The supernatant is discarded and the beads washed with 150 μl 0.1xSSPE + 0.1% Tween at room temperature for 15 minutes w/ rotation. The wash is discarded and the fully purified reporters eluted with 25 μl 0.1x SSPE at 45°C for 15 minutes w/ rotation. At this point only targets molecules that are hybridized to both a ghost probe (containing the anti-F

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 258

sequence) and a reporter (containing the anti-G sequence) will remain in solution.

13.12 IMMOBILIZATION AND STRETCHING AND IMAGING OF NANOREPORTER-TARGET COMPLEXES

[0480] Attachment to the slide and immobilization of the stretched complex may be achieved via a biotin-streptavidin interaction. In alternative embodiments, immobilization and stretching are achieved with other interaction pairs provided one of the two could be immobilized on the slide and the other attached to either the ghost probe or the reporter. Stretching does not have to be achieved via electrophoresis but can be done mechanically. The addition of bis-tris propane to the sample before binding is not required. The technology is not limited to the use of particular label monomers exemplified herein as long as the different label monomers can be separated by image processing.

[0481] One example of an immobilization and stretching protocol is provided below.

[0482] After purification, the hybridization products are loaded directly into an open well of a microfluidic device. The liquid is pulled into a microfluidic channel by capillary action where the hybridized molecules bind to the streptavidin-coated slide through the biotinylated ghost probe. The microfluidic device then intermittently tilts along the axis perpendicular to the length of the channels in alternating directions in order to force the reaction mixture to repeatedly pass through the channel and increase the binding efficiency.

[0483] After binding the hybridization reaction, the channel is washed with 1x TAE for 5 minutes by tilting the device at an angle. Fresh TAE is then added to each well to a level sufficient to contact platinum electrodes which are inserted in the wells (30 microliters in our current geometry). An electrical potential of 200V is then applied between the two wells connected by the microfluidic channel, stretching the reporters. After one minute of pre-electrophoresis to remove any remaining contaminating un-bound reporter molecules in the channel, a solution of 0.5 μM G-hooks in 1x TAE is added to the cathodic well (60 microliters of this solution). The electrical potential draws the G-hooks through the channel toward the anodic well. As they pass through the channel, the hooks hybridize with the free G-tag sequences on the free-end of the reporters which are bound to the surface and stretched. The streptavidin on the surface then binds the biotin on the G-hook and immobilizes the free end. When the potential is removed, the reporters remain stretched for imaging.

14. EXAMPLE 9: HYBRIDIZATION OF 509 CELLULAR GENES TO 100NG TOTAL RNA FROM A549 CELLS USING NANOSTRING REPORTER SYSTEM.

14.1 HYBRIDIZATION REACTION

[0484] Detection of 509 endogenous cellular genes was carried out in single multiplexed hybridization reaction. Eight non-human control sequences were spiked into each reaction that corresponded to approximately 0.1, 0.5, 1, 5, 10, 50, and 100 copies per cell as well as two reporters with no target (negative controls). There were also 4 reporters added that served as positive (3) and negative (1) controls for the post-hybridization purification process. A set of negative control hybridization was also performed containing the entire Nanostring reporter library but lacking cellular RNA.

[0485] Each sample was hybridized in triplicate. Final concentrations of the hybridization reagents were as follows: 20.8 nM total Nanoreporters (521 individual Nanoreporters at 40pM each), 103 nM total ghost probe (517 individual ghost probes @ 200pM each), 5X SSPE (pH 7.5), 5X Denhardt's reagent, 100ng/ul sheared salmon sperm DNA, 0.1 % tween 20, 50fM S11 spike target DNA, 10fM S10 spike target DNA, 5fM S9 spike target DNA, 1fM S8 spike target DNA, 0.5 fM S7 spike target DNA, 0.1fM S6 spike target DNA. S3 and S4 were added as negative controls. RNA was obtained from A549 lung epithelial cells under two different conditions. The final concentration of total RNA per hybridization was 33ng/ul. No total RNA was added to the negative control hybridizations. The final volume of the reaction was 30ul. Reagents were mixed and incubated at 65°C in thermocycler block with heated lid for 20 hours.

Master mix	(1 Reaction)	(9.3 reactions)
2.7 X hybridization mix*	11.1 μl	103.2 μl
513 endogenous gene reporters (0.24nM each)	5 μl	46.5 μl
513 endogenous gene ghost probes (1.3nM each)	4.6 μl	42.9 μl
Purification Control reporters (0.6nM each)	1 μl	9.3 μl
30X control target mix	1 μl	9.3 μl
Total	22.7 μl	211.2 μl
*Hybridization mix (13.5X SSPE, 13.5X Denhardt's reagent, 270 ng salmon sperm DNA, 0.27% tween 20)		

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 259

Reactions	1	2	3	4	5	6	7	8	9
Master mix	22.7	22.7	22.7	22.7	22.7	22.7	22.7	22.7	22.7
48.5ng/μl RNA #1	2.1	2.1	2.1	0	0	0	0	0	0
48.4ng/μl RNA #2	0	0	0	2.1	2.1	2.1	0	0	0
H_20	5.2	5.2	5.2	5.2	5.2	5.2	7.3	7.3	7.3
Total Rxn volume	30 μl	30 μl	30 μl	30 μl	30 μl	30 μl	30 μl	30 μl	30 μl
Incubate reactions in thermocycler with heated lid overnight (20hours).									

14.2 POST-HYBRIDIZATION PURIFICATION

[0486] Hybridization reactions were purified to remove unhybridized reporters using an oligonucleotide complimentary to ghost probe attached to magnetic beads (F-bead). Hybridization reactions were diluted 5 fold in 0.1 % Tween-20/TE to bring the final salt concentration to 1X SSPE. The diluted hybridization solution was then added to 100ul of F-beads (in 0.1% Tween-20) and allowed to bind to the beads at room temperature for 30min with continuous rotation. The beads were then washed three times in 150ul of 0.1X SSPE/0.1% Tween-20 and eluted in 100ul of 0.1X SSPE/0.1% Tween-20 for 15min at 45°C.

[0487] After F-bead elution, samples were purified from the opposite end of the hybridized complex using G-beads. Elutions were brought to a final concentration of 1X SSPE by the addition of 50ul of 3X SSPE/0.1% Tween-20 and bound to 30ul of G-beads (in 0.1% Tween-20) for 15min at room temperature with rotation. Beads were then washed as above and eluted in 30ul of 0.1X SSPE/Tween-20 and prepared for binding as described below.

14.3 BINDING, STRETCHING, AND IMMOBILIZATION

[0488] The samples was prepared for binding by addition of 1 ul of 1/5000 dilution of 0.1 uM Tetraspec™ fluorescent microspheres (product # T7279, Molecular Probes). Samples were loaded into a Nanostring fluidic device and attached to Accer18 Optichem® slide coated with strepavidin (product #TB0200) by tilting the device 45deg for 15mm and repeated a total of 4 times. After loading, slide surface was washed once with 90ul of 1X TAE. After wash buffer is removed the sample is prepared for electrostretching by addition of 40ul of TAE to each well. Attached complexes were stretched by applying 200V across the fluidic channel. After 1 minute the samples were immobilized in the stretched position by adding 60ul of 500nM of G-hook oligo solution to the well containing the negatively charged electrode while continuing to apply voltage for 5 minutes. After immobilization the TAE solution is removed and replaced with anti-photobleaching reagent for imaging.

14.4 IMAGING

[0489] Slides were imaged on Nikon Eclipse TE2000E equipped with a metal halide light source (X-cite 120, Exfo Corporation) and a 60X oil immersion lens (1.4 NA Plan Apo VC, Nikon). For each field of view, 4 images at different excitation wavelengths (480, 545, 580 and 622) were acquired with an Orca Ag CCD camera (Hamamatsu) under control of either Metamorph (Universal Imaging Corporation) or custom software. Images were processed with custom image processing software.

14.5 DATA ANALYSIS

[0490] Raw data was extracted from processed images using custom software. Data was normalized to the average counts for control spikes in each sample. To determine if a gene was "detected" by the system, the counts obtained for each gene from hybridizations containing RNA were compared to average counts of the two negative controls using a Student's ttest. The number of genes detected was 441 (87%) and 445 (88%) in sample #1 and #2, respectively.

[0491] A scatter plot (Figure shows normalized and average $\log_2$ signal values from each positive sample (n=3) for all 509 genes. The genes that were significantly different in the two samples were identified by a T-test of signal values in sample #2 against sample #1. In the graph below, the solid lines indicate the 2 fold upregulated threshold (black line) and 2-fold downregulated threshold (gray line) relative to sample #1. Genes with significant fold changes (p-value < 0.05) are shown in solid black diamonds. Genes whose fold change p-values were above this threshold are shown in open black squares.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 260

15. EXAMPLE 10: DETECTION OF SMALL SPOTS

[0492] As mentioned above, the label attachment regions of a nanoreporter scaffold region have a length anywhere from 10 nm to 10,000 rnn, but preferably corresponds closely to the smallest spot that can be detected with standard optics, which is about 300 nm. Spots of different color (spectrally distinguishable) are spatially resolvable at closer spacing than spots of the same color. It is possible to fit one, two, three or four spots of different colors between two spots of the same color, and yet spectrally and spatially resolve all the spots. It is also possible to significantly reduce the distance between two spots of the same color.

[0493] The limits of spatial resolution, i.e., differentiating closely spaced spots of the same color, are often thought of as hard limits, i.e., the Rayleigh Criteria (Inoue, S., Spring,. Video Microcopy (Plenum Press, 1997), p 30.) There are many techniques to drive beyond these limits that involve different imaging and/or image processing techniques. On the imaging side, structured illumination is one method to resolve spots of the same color that are spaced closer together. 50 nm has been demonstrated but, in theory, resolution with structured illumination is unlimited (Gustafsson, 2005, Proc. Nat'l. Acad. Sci. U.S.A. 102:13081-13086). On the image processing side, mixture modeling is an effective technique to push beyond commonly accepted limits (Thomann et ah, 2002, J. Microsc. 211:230-248). The combination of these techniques allows for drastically smaller nanoreporters with smaller spots, corresponding to label attachment regions of less than 50 nm.

[0494] These smaller spot spacing could allow for drastically shorter and more stable reporters, a larger number of codes, as well as a higher degree of multiplexing before the entanglement threshold is passed (for an explanation of entanglement thresholds, see Example 9 (described in Section 14) above. The tradeoff of making the spots much smaller and the reporters much shorter would be decreased signal and slower scan times. However, other technical advances, such as brighter light sources, and more efficient CCDs may offset the increased scan times making these approaches reasonable.

Claims

1. A probe pair comprising a first probe and a second probe,

(a) said first probe being a complex comprising:

(i) a first molecule comprising a first label attachment region to which are attached one or more label monomers that emit light constituting a first signal and a second label attachment region, which is non-overlapping with the first label attachment region, to which are attached one or more label monomers that emit light constituting a second signal;
(ii) a first target-specific sequence attached to the first molecule, and
(iii) an affinity tag attached to said first molecule,

(b) said second probe being a second molecule, comprising

(i) a second target-specific sequence;
(ii) optionally, a third label attachment region to which are attached one or more label monomers that emit light constituting a third signal; and
(iii) an affinity tag attached to said second molecule;

wherein the first target-specific sequence and the second target-specific sequence bind to different regions of the same target molecule,
wherein the target molecule is a naturally occurring molecule or a cDNA of a naturally occurring molecule or the complement of said cDNA,
wherein when said probe pair is bound to its target molecule, the identity of the first and second signals and their locations relative to each other constitute at least part of a code that identifies the target molecule,
wherein the first, second, and third label attachment regions are predetermined nucleotide sequences,
wherein the first, second, and third signals are spatially distinguishable, and
wherein a first DNA or RNA molecule is hybridized to the first label attachment region, to which said first DNA or RNA molecule is bound said one or more label monomers that emit light constituting said first signal; and wherein a second DNA or RNA molecule is hybridized to the second label attachment region, to which said second DNA or RNA molecule is bound said one or more label monomers that emit light constituting said second signal.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 261

2. A probe pair comprising a first probe and a second probe,

(a) said first probe being a complex comprising:

(i) a first molecule comprising a first label attachment region to which are attached one or more label monomers that emit light constituting a first signal and a second label attachment region, which is non-overlapping with the first label attachment region, to which are attached one or more label monomers that emit light constituting a second signal;
(ii) a first target-specific sequence attached to the first molecule, and
(iii) an affinity tag attached to said first molecule,

(b) said second probe being a second molecule, comprising

(i) a second target-specific sequence;
(ii) optionally, a third label attachment region to which are attached one or more label monomers that emit light constituting a third signal; and
(iii) an affinity tag attached to said second molecule;

wherein the first target-specific sequence and the second target-specific sequence bind to different regions of the same target molecule,
wherein the first signal, the second signal and the third signal are spatially distinguishable; wherein the target molecule is a naturally occurring molecule or a cDNA of a naturally occurring molecule or the complement of said cDNA,
wherein when said probe pair is bound to its target molecule, the identity of the first and second signals and their locations relative to each other constitute at least part of a code that identifies the target molecule,
wherein the first, second, and third label attachment regions are predetermined nucleotide sequences, and
wherein a first DNA or RNA molecule is hybridized to the first label attachment region, to which said first DNA or RNA molecule is bound said one or more label monomers that emit light constituting said first signal; and wherein a second DNA or RNA molecule is hybridized to the second label attachment region, to which said second DNA or RNA molecule is bound said one or more label monomers that emit light constituting said second signal.

3. The probe pair of claims 1 or 2, wherein a plurality of first DNA or RNA molecules are hybridized to the first label attachment region, to which DNA or RNA molecules are bound said one or more label monomers that emit light constituting said first signal; and wherein a plurality of second DNA or DNA molecules are hybridized to the second label attachment region, to which second DNA or RNA molecules are bound said one or more label monomers that emit light constituting said second signal.

4. The probe pair of any one of claims 1-3, wherein the second probe comprises a third label attachment region to which are attached one or more label monomers that emit light constituting a third signal.

5. The probe pair of claim 4, wherein said third label attachment region is a predetermined nucleotide sequence, and wherein a third RNA molecule is hybridized to the third label attachment region, to which third RNA molecule are attached said one or more label monomers that emit light constituting said third signal.

6. The probe pair of claim 5, wherein when the probe pair is bound to its target molecule, the code comprises the identity of the first signal, second signal and third signal and their locations relative to each other.

7. The probe pair of claims 1 or 2, wherein the code comprises the identity of the first and second signals, their locations relative to each other, and the size of the spot resulting from at least one of said signals.

8. The probe pair of any one of claims 1-3, wherein the first and second molecules are nucleic acid molecules.

9. The probe pair of claim 8, wherein the label attachment regions and target- specific sequences are predetermined nucleotide sequences.

10. The probe pair of claim 9 in which said one or more label monomers are covalently bound to nucleic acids hybridized to their respective attachment regions.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 262

11. The probe pair of claim 10, wherein said nucleic acids are hybridized to the respective attachment regions by way of one or more non-covalently bound bridging nucleic acids.

12. The probe pair of claims 1 or 2 in which the first and second target-specific sequences are unlabeled with any one or more of said label monomers.

13. The probe pair off any one of claims 1-3, wherein the second molecule further comprises a fourth label attachment region, which is non-overlapping with the third label attachment region, to which is attached one or more label monomers that emit light constituting a fourth signal.

14. The probe pair of claim 13, wherein when the probe pair is bound to its target molecule, the code comprises the identity of the first signal, second signal, third signal and fourth signal and their locations relative to each other.

15. The probe pair of any one of claims 1-3, wherein a first RNA molecule is hybridized to the first label attachment region, to which first RNA molecule are bound said one or more label monomers that emit light constituting said first signal; and wherein a second RNA molecule is hybridized to the second label attachment region, to which second RNA molecule are bound said one or more label monomers that emit light constituting said second signal.

16. The probe pair of any one of claims 1-3, wherein the affinity tags are covalently attached to the first molecule or to the second molecule.

17. The probe pair of claims 1 or 2 , wherein the label monomers attached to the first label attachment region emit light at the same wavelength, which light constitutes said first signal, wherein the label monomers attached to the second label attachment region emit light at the same wavelength, which light constitutes the second signal and wherein the label monomers attached to the third label attachment region, if present, emit light at the same wavelength, which light constitutes the third signal.

18. The probe pair of claims 1 or 2, wherein at least one of the first signal and the second signal comprises light at a plurality of different wavelengths.

19. The probe pair of claims 1 or 2, wherein the first, second and third signals are spectrally distinguishable.

20. The probe pair of any one of claims 1-3, wherein the first and second RNA molecules are each 100 to 3,000 nucleotides in length or wherein the first and second RNA molecules are each 500 to 1,500 nucleotides in length.

21. The probe pair of any one of claims 1-3, wherein the second probe is a nucleic acid complex comprising:

(a) the second molecule, wherein the second molecule comprises a third label attachment region to which is hybridized a third RNA molecule, to which third RNA molecule are bound one or more label monomers that emit light constituting a third signal; and
(b) the second target-specific sequence attached to the second molecule,

wherein the code comprises the identity of the first, second and third signals and their locations relative to each other.

22. The probe pair of claim 21, wherein the label monomers attached to the first label attachment regions emit light at the same wavelength, which light constitutes said first signal; wherein the label monomers attached to the second label attachment region emit light at the same wavelength, which light constitutes the second signal; and wherein the label monomers attached to the third label attachment region emit light at the same wavelength, which light constitutes the third signal.

23. The probe pair of claim 21, wherein at least one of the first signal, the second signal and the third signal comprises light at a plurality of different wavelengths.

24. The probe pair of claim 21, in which the first target-specific sequence and the second target-specific sequence are unlabeled with any one or more of said label monomers.

25. The probe pair of claim 21, wherein first, second and third signals are spectrally distinguishable.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 263

26. The probe pair of claim 21, wherein first and third signals emit at the same wavelength or wavelengths.

27. The probe pair of claim 21, wherein the affinity tag is covalently attached to the first molecule.

28. The probe pair of claim 21, wherein the affinity tag is covalently attached to the second molecule.

29. The probe pair of claim 21, wherein the first, second and third RNA molecules are each 100 to 3.000 nucleotides or wherein the first, second and third RNA molecules are each 500 to 1,500 nucleotides,

30. The probe pair of claims 21-30, wherein the RNA molecules comprise a single region comprising a regularly-repeated base.

31. The probe pair of claim 30, wherein the regularly repeated base is aminoallyl-modified uracil.

32. A method of detecting a target molecule in a biomolecular sample comprising:

(i) contacting said sample with a probe pair according to any one of claims 1-31 under conditions that allow binding of the first target-specific sequence and the second target-specific sequence to the target molecule and
(ii) detecting the code that identifies the target molecule.

33. The method of claim 32, further comprising quantitating the amount of said target molecule in said biomolecular sample.

34. The method of claims 32 or 33, wherein the concentration of each of said first and second target-specific sequences is in excess of the concentration of said target molecule.

35. The method of claims 32-34, further comprising the step of stretching the bound probe pair and target molecule prior to detecting the code that identifies the target molecule.

36. The method of claim 3 5, wherein the stretching is performed using a technique selected from the group consisting of flow-stretching, a receding meniscus technique, electrostretching and constriction in the flow of a liquid containing the nanoreporter in conjunction with an oscillating electric field.

37. A method of detecting a plurality of target molecules in a biomolecular sample comprising:

(i) contacting said samples with a population of probe pairs according to any one of claims 1-31, under conditions that allow binding of the first target-specific sequence and the second target-specific sequence of each probe pair to their respective target molecule, wherein each probe pair in said population when bound to its respective target molecule is associated with a distinguishable code, and
(ii) detecting the codes that identify the plurality of target molecules.

38. The method of claim 37, further comprising quantitating the amount of each of said plurality of target molecules in said biomolecular sample.

39. The method of claims 37-38, wherein the concentration of each of said first and second target-specific sequences of each probe pair of said population is in excess of the concentration of its respective target molecule.

40. The method of claims 37-39, further comprising the step of stretching the bound probe pairs and their respective target molecules prior to detecting the codes that identifies the plurality of target molecules.

41. The method of claim 40, wherein the stretching is performed using a technique selected from the group consisting of flow-stretching, a receding meniscus technique, electrostretching and constriction in the flow of a liquid containing the nanoreporter in conjunction with an oscillating electric field.

Patentansprüche

1. Sondenpaars, umfassend eine erste Sonde und eine zweite Sonde,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 264

(a) worin die erste Sonde einen Komplex darstellt, umfassend:

(i) ein erstes Molekül, umfassend eine erste Region zur Markierungsbefestigung, an welcher ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung eines ersten Signals, und eine zweite Region zur Markierungsbefestigung, welche nicht mit der ersten Region zur Markierungsbefestigung überlappt, an welcher ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung eines zweiten Signals;
(ii) eine erste, zielspezifische Sequenz, welche an dem ersten Molekül befestigt ist, und
(iii) einen Affinitäts-Tag, welcher an dem ersten Molekül befestigt ist,

(b) worin die zweite Sonde ein zweites Molekül ist, umfassend

(i) eine zweite, zielspezifische Sequenz;
(ii) optional eine dritte Region zur Markierungsbefestigung, an welcher ein oder mehrere Markierungsmonomere befestigt sind, weiche Licht emittieren unter Bildung eines dritten Signals; und
(iii) einen Affinitäts-Tag, welcher an dem zweiten Molekül befestigt ist;

worin die erste zielspezifische Sequenz und die zweite zielspezifische Sequenz an unterschiedliche Regionen des gleichen Zielmoleküls binden,
worin das Zielmolekül ein natürlich vorkommendes Molekül oder eine cDNA eines natürlich vorkommenden Moleküls oder das Komplement der cDNA ist,
worin, wenn das Sondenpaar an sein Zielmolekül gebunden ist, die Identität des ersten und zweiten Signals und ihre Positionen relativ zueinander zumindest einen Teil eines Codes bilden, welcher das Zielmolekül identifiziert,
worin die erste, zweite und dritte Region zur Markierungsbefestigung vorher festgelegte Nukleotidsequenzen sind,
worin das erste, zweite und dritte Signal räumlich unterscheidbar sind, und
worin ein erstes DNA- oder RNA-Molekül mit der ersten Region zur Markierungsbefestigung hybridisiert ist, worin an das erste DNA- oder RNA-Molekül ein oder mehrere Markierungsmonomere gebunden sind, welche Licht emittieren unter Bildung des ersten Signals;
und worin ein zweites DNA- oder RNA-Molekül mit der zweiten Region zur Markierungsbefestigung hybridisiert ist, worin an das zweite DNA- oder RNA-Molekül ein oder mehrere Markierungsmonomere gebunden sind, weiche Licht emittieren unter Bildung des zweiten Signals.

2. Sondenpaar, umfassend eine erste Sonde und eine zweite Sonde,

(a) worin die erste Sonde einen Komplex darstellt, umfassend:

(i) ein erstes Molekül, umfassend eine erste Region zur Markierungsbefestigung, an welcher ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung eines ersten Signals, und eine zweite Region zur Markierungsbefestigung, weiche nicht mit der ersten Region zur Markierungsbefestigung überlappt, an welcher ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung eines zweiten Signals;
(ii) eine erste, zielspezifische Sequenz, welche an dem ersten Molekül befestigt ist, und
(iii) einen Affinitäts-Tag, welcher an dem ersten Molekül befestigt ist,

(b) worin die zweite Sonde ein zweites Molekül ist, umfassend

(i) eine zweite, zielspezifische Sequenz;
(ii) optional eine dritte Region zur Markierungsbefestigung, an welcher ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung eines dritten Signals; und
(iii) einen Affinitäts-Tag, welcher an dem zweiten Molekül befestigt ist;

worin die erste zielspezifische Sequenz und die zweite zielspezifische Sequenz an unterschiedliche Regionen des gleichen Zielmoleküls binden,
worin das erste Signal, das zweite Signal und das dritte Signal räumlich unterscheidbar sind;
worin das Zielmolekül ein natürlich vorkommendes Molekül oder eine DNA eines natürlich vorkommenden Moleküls oder das Komplement der cDNA ist,
worin, wenn das Sondenpaar an sein Zielmolekül gebunden ist, die Identität des ersten und zweiten Signals und ihre Positionen relativ zueinander zumindest einen Teil eines Codes bilden, welcher das Zielmolekül identifiziert,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 265

worin die erste, zweite und dritte Region zur Markierungsbefestigung vorher festgelegte Nukleotidsequenzen sind, und
worin ein erstes DNA- oder RNA-Molekül mit der ersten Region zur Markierungsbefestigung hybridisiert ist, worin an das erste DNA- oder RNA-Molekül ein oder mehrere Markierungsmonomere gebunden sind, welche Licht emittieren unter Bildung des ersten Signals;
und worin ein zweites DNA- oder RNA-Molekül mit der zweiten Region zur Markierungsbefestigung hybridisiert ist, worin an das zweite DNA- oder RNA-Molekül ein oder mehrere Markierungsmonomere gebunden sind, welche Licht emittieren unter Bildung des zweiten Signals.

3. Sondenpaar nach Anspruch 1 oder 2, worin eine Vielzahl von ersten DNA- oder RNA-Molekülen mit der ersten Region zur Markierungsbefestigung hybridisiert sind, worin an die DNA- oder RNA-Moleküle ein oder mehrere Markierungsmonomere gebunden sind, welche Licht emittieren unter Bildung des ersten Signals;
und worin eine Vielzahl von zweiten DNA- oder RNA-Molekülen mit der zweiten Region zur Markierungsbefestigung hybridisiert sind, worin an die zweiten DNA- oder RNA-Moleküle ein oder mehrere Markierungsmonomere gebunden sind, weiche Licht emittieren unter Bildung des zweiten Signals.

4. Sondenpaar nach einem der Ansprüche 1-3, worin die zweite Sonde eine dritte Region zur Markierungsbefestigung umfasst, an welcher ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung eines dritten Signals.

5. Sondenpaar nach Anspruch 4, worin die dritte Region zur Markierungsbefestigung eine vorher festgelegte Nukleotidsequenz ist, und worin ein drittes RNA-Molekül mit der dritten Region zur Markierungsbefestigung hybridisiert ist, worin an dem dritten RNA-Molekül ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung des dritten Signals.

6. Sondenpaar nach Anspruch 5, worin, wenn das Sondenpaar an sein Zielmolekül gebunden ist, der Code die Identität des ersten Signals, des zweiten Signals und des dritten Signals und ihre Positionen relativ zueinander umfasst.

7. Sondenpaar nach Anspruch 1 oder 2, worin der Code die Identität des ersten und zweiten Signals, ihre Positionen relativ zueinander und die Größe des aus mindestens einem der Signale resultierenden Spots umfasst.

8. Sondenpaar nach einem der Ansprüche 1-3, worin das erste und zweite Molekül Nukleinsäuremoleküle sind.

9. Sondenpaars nach Anspruch 8, worin die Regionen zur Markierungsbefestigung und die zielspezifischen Sequenzen vorher festgelegte Nukleotidsequenzen sind.

10. Sondenpaar nach Anspruch 9, worin ein oder mehrere Markierungsmonomere kovalent an Nukleinsäuren gebunden sind, welche mit deren betreffenden Regionen zur Befestigung hybridisiert sind.

11. Sondenpaar nach Anspruch 10, worin die Nukleinsäuren mit den betreffenden Regionen zur Befestigung mittels einer oder mehreren nicht kovalent gebundenen brückenbildenden Nukleinsäuren hybridisiert sind.

12. Sondenpaar nach Anspruch 1 oder 2, worin die erste und zweite zielspezifische Sequenz nicht markiert sind mit einem oder mehreren Markierungsmonomeren.

13. Sondenpaar nach einem der Ansprüche 1-3, worin das zweite Molekül weiter eine vierte Region zur Markierungsbefestigung umfasst, welche nicht mit der dritten Region zur Markierungsbefestigung überlappt, an welcher ein oder mehrere Markierungsmonomere befestigt sind, welche Licht emittieren unter Bildung eines vierten Signals.

14. Sondenpaar nach Anspruch 13, worin, wenn das Sondenpaar an sein Zielmolekül gebunden ist, der Code die Identität des ersten Signals, des zweiten Signals, des dritten Signals und des vierten Signals und ihre Positionen relativ zueinander umfasst.

15. Sondenpaar nach einem der Ansprüche 1-3, worin ein erstes RNA-Molekül mit der ersten Region zur Markierungsbefestigung hybridisiert ist, worin an das erste RNA-Molekül ein oder mehrere Markierungsmonomere gebunden sind, welche Licht emittieren unter Bildung des ersten Signals, und worin ein zweites RNA-Molekül mit der zweiten Region zur Markierungsbefestigung hybridisiert ist, worin an das zweite RNA-Molekül ein oder mehrere Markierungsmonomere gebunden sind, welche Licht emittieren unter Bildung des zweiten Signals.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 266

16. Sondenpaar nach einem der Ansprüche 1-3, worin die Affinitäts-Tags kovalent an dem ersten Molekül oder an dem zweiten Molekül befestigt sind.

17. Sondenpaar nach Anspruch 1 oder 2, worin die Markierungsmonomere, welche an der ersten Region zur Markierungsbefestigung befestigt sind, Licht mit der gleichen Wellenlänge emittieren, worin das Licht das erste Signal bildet, worin die Markierungsmonomere, welche an der zweiten Region zur Markierungsbefestigung befestigt sind, Licht mit der gleichen Wellenlänge emittieren, worin das Licht das zweite Signal bildet, und worin die Markierungsmonomere, welche an der dritten Region zur Markierungsbefestigung befestigt sind, falls vorhanden, Licht mit der gleichen Wellenlänge emittieren, worin das Licht das dritte Signal bildet.

18. Sondenpaar nach Anspruch 1 oder 2, worin zumindest eines des ersten und des zweiten Signals Licht mit mehreren unterschiedlichen Wellenlängen umfasst.

19. Sondenpaar nach Anspruch 1 oder 2, worin das erste, zweite und dritte Signal spektral unterscheidbar sind.

20. Sondenpaar nach einem der Ansprüche 1-3, worin das erste und zweite RNA-Molekül jeweils 100 bis 3000 Nukleotide lang sind, oder worin das erste und zweite RNA-Molekül jeweils 500 bis 1500 Nukleotide lang sind.

21. Sondenpaar nach einem der Ansprüche 1-3, worin die zweite Sonde ein Nukleinsäurekomplex ist, umfassend:

(a) das zweite Molekül, worin das zweite Molekül eine dritte Region zur Markierungsbefestigung umfasst, mit welcher ein drittes RNA-Molekül hybridisiert ist, worin an das dritte RNA-Molekül ein oder mehrere Markierungsmonomere gebunden sind, welche Licht emittieren unter Bildung eines dritten Signals, und
(b) die zweite zielspezifische Sequenz, befestigt an dem zweiten Molekül, worin der Code die Identität des ersten, zweiten und dritten Signals und ihre Positionen relativ zueinander umfasst.

22. Sondenpaar nach Anspruch 21, worin die Markierungsmonomere, welche an den ersten Regionen zur Markierungsbefestigung befestigt sind, Licht mit der gleichen Wellenlänge emittieren, worin das Licht das erste Signal bildet, worin die Markierungsmonomere, welche an der zweiten Region zur Markierungsbefestigung befestigt sind, Licht mit der gleichen Wellenlänge emittieren, worin das Licht das zweite Signal bildet, und worin die Markierungsmonomere, welche an der dritten Region zur Markierungsbefestigung befestigt sind, Licht mit der gleichen Wellenlänge emittieren, worin das Licht das dritte Signal bildet.

23. Sondenpaar nach Anspruch 21, worin zumindest eines des ersten Signals, des zweiten Signals und des dritten Signals Licht mit mehreren unterschiedlichen Wellenlängen umfasst.

24. Sondenpaar nach Anspruch 21, worin die erste zielspezifische Sequenz und die zweite zielspezifische Sequenz nicht markiert sind mit einem oder mehreren Markierungsmonomeren.

25. Sondenpaar nach Anspruch 21, worin das erste, zweite und dritte Signal spektral unterscheidbar sind.

26. Sondenpaar nach Anspruch 21, worin das erste und dritte Signal mit der gleichen Wellenlänge oder den gleichen Wellenlängen emittieren.

27. Sondenpaar nach Anspruch 21, worin der Affinitäts-Tag kovalent an dem ersten Molekül befestigt ist.

28. Sondenpaar nach Anspruch 21, worin der Affinitäts-Tag kovalent an dem zweiten Molekül befestigt ist.

29. Sondenpaar nach Anspruch 21, worin das erste, zweite und dritte RNA-Molekül jeweils 100 bis 3000 Nukleotide aufweisen oder worin das erste, zweite und dritte RNA-Molekül jeweils 500 bis 1500 Nukleotide aufweisen.

30. Sondenpaar nach Anspruch 21-30, worin die RNA-Moleküle eine einzelne Region umfassen, welche eine sich regelmäßig wiederholende Base umfasst.

31. Sondenpaar nach Anspruch 30, worin die sich regelmäßig wiederholende Base Aminoallyl-modifiziertes Uracil ist.

32. Verfahren zum Nachweisen eines Zielmoleküls in einer biomolekularen Probe, umfassend:

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 267

(i) Inkontaktbringen der Probe mit einem Sondenpaar nach einem der Ansprüche 1 - 31 unter Bedingungen, welche eine Bindung der ersten zielspezifischen Sequenz und der zweiten zielspezifischen Sequenz an das Zielmolekül ermöglicht, und
(ii) Nachweisen des Codes, welcher das Zielmolekül identifiziert.

33. Verfahren nach Anspruch 32, weiter umfassend Quantifizieren der Menge des Zielmoleküls in der biomolekularen Probe.

34. Verfahren nach Anspruch 32 oder 33, worin die Konzentration sowohl der ersten als auch der zweiten zielspezifischen Sequenz größer ist als die Konzentration des Zielmoleküls.

35. Verfahren nach Anspruch 32-34, weiter umfassend den Schritt Strecken des gebundenen Sondenpaars und des Zielmoleküls vor dem Nachweisen des Codes, welcher das Zielmolekül identifiziert,

36. Verfahren nach Anspruch 35, worin das Strecken durchgeführt wird unter Verwendung eines Verfahrens, ausgewählt aus der Gruppe bestehend aus Strömungsstrecken ("Flow-Stretching"), einem Verfahren mittels zurückgehendem Meniskus, Elektrostrecken und Einengung in der Strömung einer Flüssigkeit, enthaltend den Nanoreporter, in Verbindung mit einem oszillierenden elektrischen Feld.

37. Verfahren zum Nachweisen mehrerer Zielmoleküle in einer biomolekularen Probe, umfassend:

(i) Inkontaktbringen der Probe mit einer Population von Sondenpaaren nach einem der Ansprüche 1-31, unter Bedingungen, welche eine Bindung der ersten zielspezifischen Sequenz und der zweiten zielspezifischen Sequenz jedes Sondenpaars an ihr entsprechendes Zielmolekül ermöglicht, worin jedes Sondenpaar in der Population bei Bindung mit seinem jeweiligen Zielmolekül mit einem unterscheidbaren Code verbunden wird, und
(ii) Nachweisen der Codes, welche die mehreren Zielmoleküle identifizieren.

38. Verfahren nach Anspruch 37, weiter umfassend Quantifizieren der Menge jedes der mehreren Zielmoleküle in der biomolekularen Probe.

39. Verfahren nach Anspruch 37-38, worin die Konzentration sowohl der ersten als auch der zweiten zielspezifischen Sequenzen jedes Sondenpaars der Population größer ist als die Konzentration ihres entsprechenden Zielmoleküls.

40. Verfahren nach Anspruch 37-39, weiter umfassend den Schritt Strecken der gebundenen Sondenpaare und ihrer entsprechenden Zielmoleküle vor dem Nachweisen der Codes, welche die mehreren Zielmoleküle identifizieren.

41. Verfahren nach Anspruch 40, worin das Strecken durchgeführt wird unter Verwendung eines Verfahrens, ausgewählt aus der Gruppe bestehend aus Strömungsstrecken ("Flow-Stretching"), einem Verfahren mittels zurückgehendem Meniskus, Elektrostrecken und Einengung in der Strömung einer Flüssigkeit, enthaltend den Nanoreporter, in Verbindung mit einem oszillierenden elektrischen Feld,

Revendications

1. Paire de sondes comprenant une première sonde et une deuxième sonde,

(a) ladite première sonde étant un complexe comprenant :

(i) une première molécule comprenant une première région d'attachement de marqueurs à laquelle sont attachés un ou plusieurs monomères marqueurs qui émettent une lumière constituant un premier signal et une deuxième région d'attachement de marqueurs, qui n'est pas en chevauchement avec la première région d'attachement de marqueurs, à laquelle sont attachés un ou plusieurs monomères marqueurs qui émettent une lumière constituant un deuxième signal ;
(ii) une première séquence à spécificité de cible attachée à la première molécule, et
(iii) une étiquette d'affinité attachée à ladite première molécule,

(b) ladite deuxième sonde étant une deuxième molécules comprenant :

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 268

(i) une deuxième séquence à spécificité de cible;
(ii) éventuellement une troisième région d'attachement de marqueurs à laquelle sont attachés un ou plusieurs monomères marqueurs qui émettent une lumière constituant un troisième signal ; et
(iii) une étiquette d'affinité attachée à ladite deuxième molécule ;

dans laquelle la première séquence à spécificité de cible et la deuxième séquence à spécificité de cible se lient à des régions différentes de la même molécule cible,
dans laquelle la molécule cible est une molécule naturelle ou un ADNc d'une molécule naturelle ou le complément dudit ADNc,
dans laquelle, quand ladite paire de sondes est liée à sa molécule cible, l'identité des premier et deuxième signaux et leurs positions l'un par rapport à l'autre constituent au moins une partie d'un code qui identifie la molécule cible, dans laquelle les première, deuxième et troisième régions d'attachement de marqueurs sont des séquences de nucléotides prédéterminées,
dans laquelle les premier, deuxième et troisième signaux sont spatialement distinguables, et
dans laquelle une première molécule d'ADN ou d'ARN est hybridée à la première région d'attachement de marqueurs, lesdits un ou plusieurs marqueurs monomères qui émettent une lumière constituant ledit premier signal étant liés à ladite première molécule d'ADN ou d'ARN ; et dans laquelle une deuxième molécule d'ADN ou d'ARN est hybridée à la deuxième région d'attachement de marqueurs, lesdits un ou plusieurs monomères marqueurs qui émettent une lumière constituant ledit deuxième signal étant liés à ladite deuxième molécule d'ADN ou d'ARN.

2. Paire de sondes comprenant une première sonde et une deuxième sonde,

(a) ladite première sonde étant un complexe comprenant :

(i) une première molécule comprenant une première région d'attachement de marqueurs à laquelle sont attachés un ou plusieurs monomères marqueurs qui émettent une lumière constituant un premier signal et une deuxième région d'attachement de marqueurs, qui n'est pas en chevauchement avec la première région d'attachement de marqueurs, à laquelle sont attachés un ou plusieurs monomères marqueurs qui émettent une lumière constituant un deuxième signal ;
(ii) une première séquence à spécificité de cible attachée à la première molécule, et
(iii)une étiquette d'affinité attachée à ladite première molécule,

(b) ladite deuxième sonde étant une deuxième molécule comprenant :

(i) une deuxième séquence à spécificité de cible ;
(ii) éventuellement une troisième région d'attachement de marqueurs à laquelle sont attachés un ou plusieurs monomères marqueurs qui émettent une lumière constituant un troisième signal; et
(iii) une étiquette d'affinité attachée à ladite deuxième molécule ;

dans laquelle la première séquence à spécificité de cible et la deuxième séquence à spécificité de cible se lient à des régions différentes de la même molécule cible,
dans laquelle le premier signal, le deuxième signal et le troisième signal sont spatialement distinguables ;
dans laquelle la molécule cible est une molécule naturelle ou un ADNc d'une molécule naturelle ou le complément dudit ADNc,
dans laquelle, quand ladite paire de sondes est liée à sa molécule cible, l'identité des premier et deuxième signaux et leurs positions l'un par rapport à l'autre constituent au moins une partie d'un code qui identifie la molécule cible, dans laquelle les première, deuxième et troisième régions d'attachement de marqueurs sont des séquences de nucléotides prédéterminées, et
dans laquelle une première molécule d'ADN ou d'ARN est hybridée à la première région d'attachement de marqueurs, lesdits un ou plusieurs marqueurs monomères qui émettent une lumière constituant ledit premier signal étant liés à ladite première molécule d'ADN ou d'ARN ; et dans laquelle une deuxième molécule d'ADN ou d'ARN est hybridée à la deuxième région d'attachement de marqueurs, lesdits un ou plusieurs monomères marqueurs qui émettent une lumière constituant ledit deuxième signal étant liés à ladite deuxième molécule d'ADN ou d'ARN.

3. Paire de sondes selon la revendication 1 ou 2, dans laquelle une pluralité de premières molécules d'ADN ou d'ARN sont hybridées à la première région d'attachement de marqueurs, lesdits un ou plusieurs monomères marqueurs qui émettent une lumière constituant ledit premier signal étant liés aux molécules d'ADN ou d'ARN ; et dans laquelle une pluralité de deuxièmes molécules d'ADN ou d'ARN sont hybridées à la deuxième région d'attachement de

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 269

marqueurs, lesdits un ou plusieurs monomères marqueurs qui émettent une lumière constituant ledit deuxième signal étant liés aux deuxièmes molécules d'ADN ou d'ARN.

4. Paire de sondes selon l'une quelconque des revendications 1 à 3, dans laquelle la deuxième sonde comprend une troisième région d'attachement de marqueurs à laquelle sont attachées un ou plusieurs monomères marqueurs qui émettent une lumière constituant un troisième signal.

5. Paire de sondes selon la revendication 4, dans laquelle ladite troisième région d'attachement de marqueurs est une séquence de nucléotides prédéterminée, et dans laquelle une troisième molécule d'ARN est hybridée à la troisième région d'attachement de marqueurs, lesdits un ou plusieurs monomères marqueurs qui émettent une lumière constituant ledit troisième signal étant attachés à la troisième molécule d'ARN.

6. Paire de sondes selon la revendication 5, dans laquelle, quand la paire de sondes est liée à sa molécule cible, le code comprend l'identité du premier signal, du deuxième signal et du troisième signal, et leurs positions les uns par rapport aux autres.

7. Paire de sondes selon la revendication 1 ou 2, dans laquelle le code comprend l'identité des premier et deuxième signaux, leurs positions l'un par rapport à l'autre, et la taille du point résultant d'au moins l'un desdits signaux.

8. Paire de sondes selon l'une quelconque des revendications 1 à 3, dans laquelle les première et deuxième molécules sont des molécules d'acide nucléique.

9. Paire de sondes selon la revendication 8, dans laquelle les régions d'attachement de marqueurs et les séquences à spécificité de cible sont des séquences de nucléotides prédéterminées.

10. Paire de sondes selon la revendication 9, dans laquelle lesdits un ou plusieurs monomères marqueurs sont liés de manière covalente à des acides nucléiques hybridés à leurs régions d'attachement respectives.

11. Paire de sondes selon la revendication 10, dans laquelle lesdits acides nucléiques sont hybridés aux régions d'attachement respectives au moyen d'un ou plusieurs acides nucléiques de pontage liés de manière non covalente.

12. Paire de sondes selon la revendication 1 ou 2, dans laquelle les première et deuxième séquences à spécificité de cible ne sont pas marquées par un ou plusieurs quelconques desdits monomères marqueurs.

13. Paire de sondes selon l'une quelconque des revendications 1 à 3, dans laquelle la deuxième molécule comprend en outre une quatrième région d'attachement de marqueurs, qui n'est pas en chevauchement avec la troisième région d'attachement de marqueurs, à laquelle sont attachés un ou plusieurs monomères marqueurs qui émettent une lumière constituant un quatrième signal.

14. Paire de sondes selon la revendication 13, dans laquelle, quand la paire de sondes est liée à sa molécule cible, le code comprend l'identité du premier signal, du deuxième signal, du troisième signal et du quatrième signal, et leurs positions les uns par rapport aux autres.

15. Paire de sondes selon l'une quelconque des revendications 1 à 3, dans laquelle une première molécule d'ARN est hybridée à la première région d'attachement de marqueurs, lesdits un ou plusieurs monomères marqueurs qui émettent une lumière constituant ledit premier signal étant liés à la première molécule d'ARN ; et dans laquelle une deuxième molécule d'ARN est hybridée à la deuxième région d'attachement de marqueurs, lesdits un ou plusieurs monomères marqueurs qui émettent une lumière constituant ledit deuxième signal étant liés à la deuxième molécule d'ARN.

16. Paire de sondes selon l'une quelconque des revendications 1 à 3, dans laquelle les étiquettes d'affinité sont attachées de manière covalente à la première molécule ou à la deuxième molécule.

17. Paire de sondes selon la revendication 1 ou 2, dans laquelle les monomères marqueurs attachés à la première région d'attachement de marqueurs émettent une lumière à la même longueur d'onde, laquelle lumière constitue ledit premier signal, dans laquelle les monomères marqueurs attachés à la deuxième région d'attachement de marqueurs émettent une lumière à la même longueur d'onde, laquelle lumière constitue le deuxième signal, et dans laquelle les monomères marqueurs attachés à la troisième région d'attachement de marqueurs, s'ils sont présents,

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 270

émettent une lumière à la même longueur d'onde, laquelle lumière constitue le troisième signal.

18. Paire de sondes selon la revendication 1 ou 2, dans laquelle au moins l'un parmi le premier signal et le deuxième signal comprend une lumière à une pluralité de longueurs d'onde différentes.

19. Paire de sondes selon la revendication 1 ou 2, dans laquelle les premier, deuxième et troisième signaux sont spectralement distinguables,

20. Paire de sondes selon l'une quelconque des revendications 1 à 3, dans laquelle les première et deuxième molécules d'ARN ont chacune une longueur de 100 à 3000 nucléotides, ou dans laquelle les première et deuxième molécules d'ARN ont chacune une longueur de 500 à 1500 nucléotides.

21. Paire de sondes selon l'une quelconque des revendications 1 à 3, dans laquelle à deuxième sonde est un complexe d'acide nucléique comprenant :

(a) la deuxième molécule, laquelle deuxième molécule comprend une troisième région d'attachement de marqueurs à laquelle est hybridée une troisième molécule d'ARN, un ou plusieurs monomères marqueurs qui émettent une lumière constituant un troisième signal étant liés à la troisième molécule d'ARN ; et
(b) la deuxième séquence à spécificité de cible attachée à la deuxième molécule,

dans laquelle le code comprend l'identité des premier, deuxième et troisième signaux et leurs positions les uns par rapport aux autres.

22. Paire de sondes selon la revendication 21, dans laquelle les monomères marqueurs attachés aux premières régions d'attachement de marqueurs émettent une lumière à la même longueur d'onde, laquelle lumière constitue ledit premier signal ; dans laquelle les monomères marqueurs attachés à la deuxième région d'attachement de marqueurs émettent une lumière à la même longueur d'onde, laquelle lumière constitue le deuxième signal ; et dans laquelle les monomères marqueurs attachés à la troisième région d'attachement de marqueurs émettent une lumière à la même longueur d'onde, laquelle lumière constitue le troisième signal.

23. Paire de sondes selon la revendication 21, dans laquelle au moins l'un parmi le premier signal, le deuxième signal et le troisième signal comprend une lumière à une pluralité de longueurs d'ondes différentes.

24. Paire de sondes selon la revendication 21, dans laquelle la première séquence à spécificité de cible et la deuxième séquence à spécificité de cible ne sont pas marquées par un ou plusieurs quelconques desdits monomères marqueurs.

25. Paire de sondes selon la revendication 21, dans laquelle les premier, deuxième et troisième signaux sont spectralement distinguables.

26. Paire de sondes selon la revendication 21, dans laquelle les premier et troisième signaux émettent à la ou aux mêmes longueurs d'onde.

27. Paire de sondes selon la revendication 21, dans laquelle l'étiquette d'affinité est attachée de manière covalente à la première molécule.

28. Paire de sondes selon la revendication 21, dans laquelle l'étiquette d'affinité est attachée de manière covalente à la deuxième molécule.

29. Paire de sondes selon la revendication 21, dans laquelle les première, deuxième et troisième molécules d'ARN ont chacune 100 à 3000 nucléotides ou dans laquelle les première, deuxième et troisième molécules d'ARN ont chacune 500 à 1500 nucléotides.

30. Paire de sondes selon les revendications 21 à 30, dans laquelle les molécules d'ARN comprennent une seule région comprenant une base régulièrement répétée.

31. Paire de sondes selon la revendication 30, dans laquelle la base régulièrement répétée est l'uracile à modification aminoallyle.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 271

32. Procédé pour détecter une molécule cible dans un échantillon biomoléculaire, comprenant :

(i) la mise en contact dudit échantillon avec une paire de sondes selon l'une quelconque des revendications 1 à 31 dans des conditions qui permettent la liaison, à la molécule cible, de la première séquence à spécificité de cible et de la deuxième séquence à spécificité de cible, et
(ii) la détection du code qui identifie la molécule cible.

33. Procédé selon la revendication 32, comprenant en outre la quantification de la quantité de ladite molécule cible dans ledit échantillon biomoléculaire,

34. Procédé selon la revendication 32 ou 33, dans lequel la concentration de chacune desdites première et deuxième séquences à spécificité de cible est supérieure à la concentration de ladite molécule cible.

35. Procédé selon les revendications 32 à 34, comprenant en outre l'étape consistant à étirer la paire de sondes liée et la molécule cible avant détection du code qui identifie la molécule cible.

36. Procédé selon la revendication 35, dans lequel l'étirement est effectué au moyen d'une technique choisie dans l'ensemble constitué par un étirement par écoulement, une technique de ménisque de recul, un électro-étirement et une constriction dans l'écoulement d'un liquide contenant le nanorapporteur conjointement avec un champ électrique oscillant.

37. Procédé pour détecter une pluralité de molécules cibles dans un échantillon biomoléculaire, comprenant :

(i) la mise en contact dudit échantillon avec une population de paires de sondes selon l'une quelconque des revendications 1 à 31, dans des conditions qui permettent la liaison de la première séquence à spécificité de cible et de la deuxième séquence à spécificité de cible de chaque paire de sondes à leur molécule cible respective, chaque paire de sondes dans ladite population, lorsqu'elle est liée à sa molécule cible respective, étant associée à un code distinguable, et
(ii) la détection du code qui identifie la pluralité de molécules cibles.

38. Procédé selon la revendication 37, comprenant en outre la quantification de la quantité de chacune parmi ladite pluralité de molécules cibles dans ledit échantillon biomoléculaire.

39. Procédé selon les revendications 37 et 38, dans laquelle la concentration de chacune parmi lesdites première et deuxième séquences à spécificité de cible de chaque paire de sondes de ladite population est supérieure à la concentration de sa molécule cible respective.

40. Procédé selon les revendications 37 à 39, comprenant en outre l'étape consistant à étirer les paires de sondes liées et leurs molécules cibles respectives avant détection des codes qui identifient la pluralité de molécules cibles.

41. Procédé selon la revendication 40, dans lequel l'étirement est effectué au moyen d'une technique choisie dans l'ensemble constitué par un étirement par écoulement, une technique de ménisque de recul, un électro-étirement et une constriction dans l'écoulement d'un liquide contenant le nanorapporteur conjointement avec un champ électrique oscillant.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 272

FIGURE 1A-1C

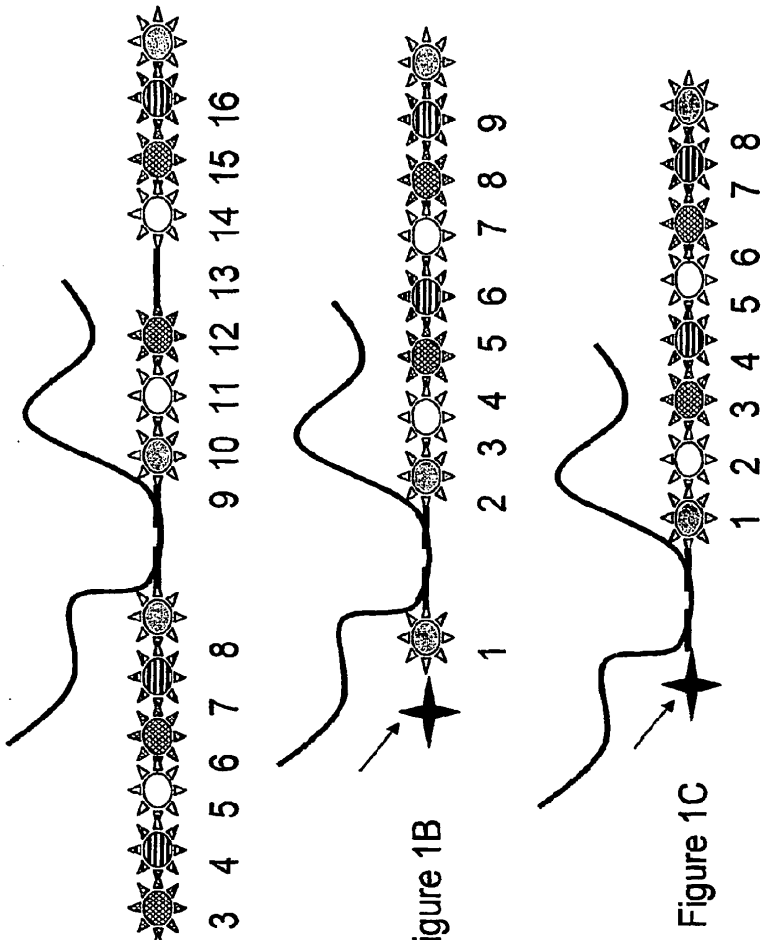


Figure 1A

Figure 1B

Figure 1C

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 273

FIGURE 1D-1F

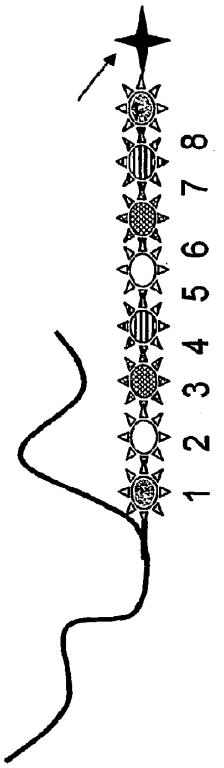


Figure 1D

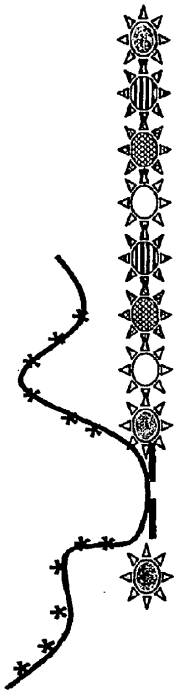


Figure 1E

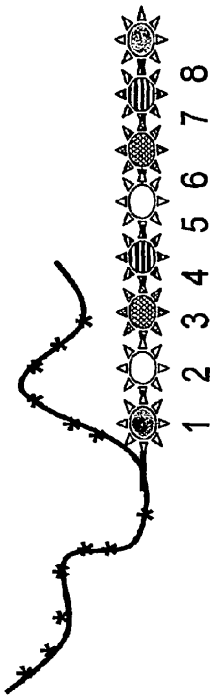
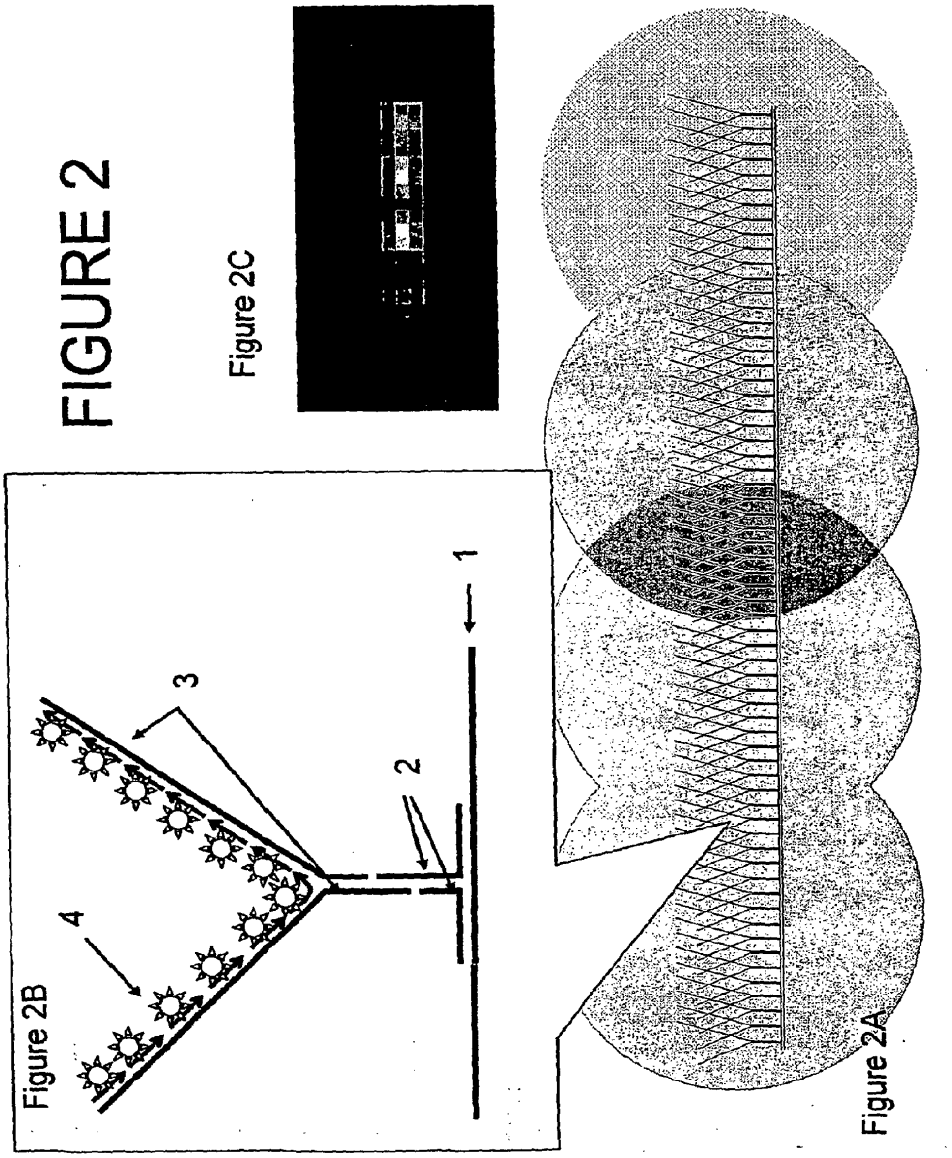


Figure 1F

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 274



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 275

FIGURE 3

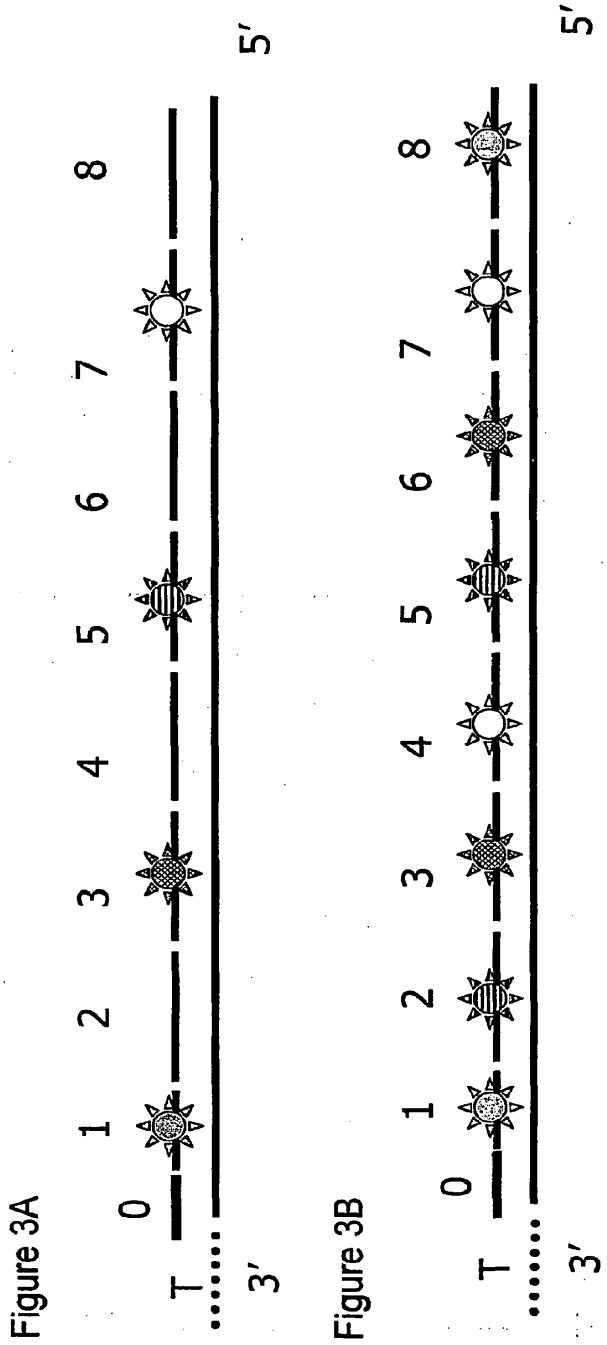
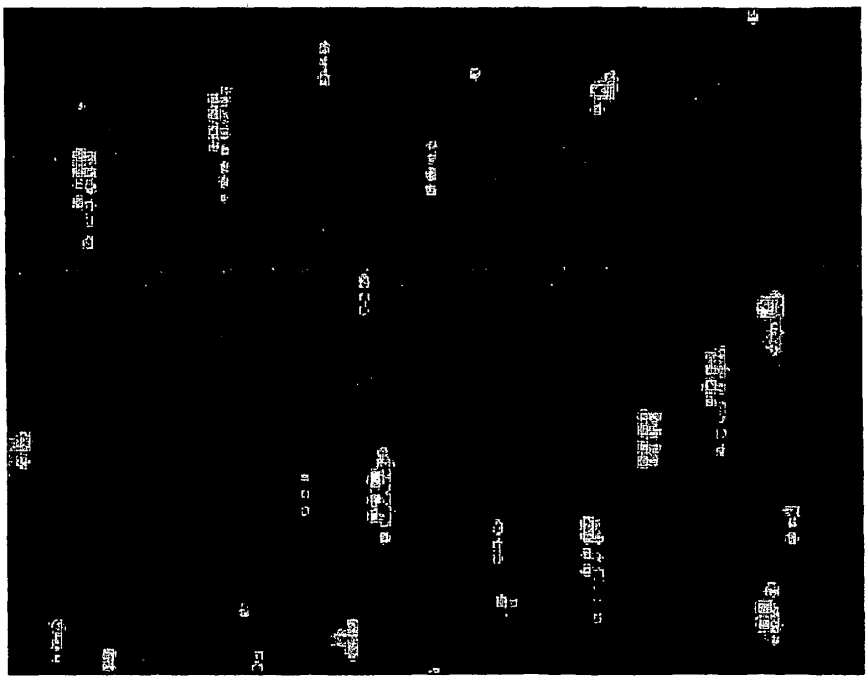


Figure 3A

Figure 3B

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 276

FIGURE 4



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 277

FIGURE 5

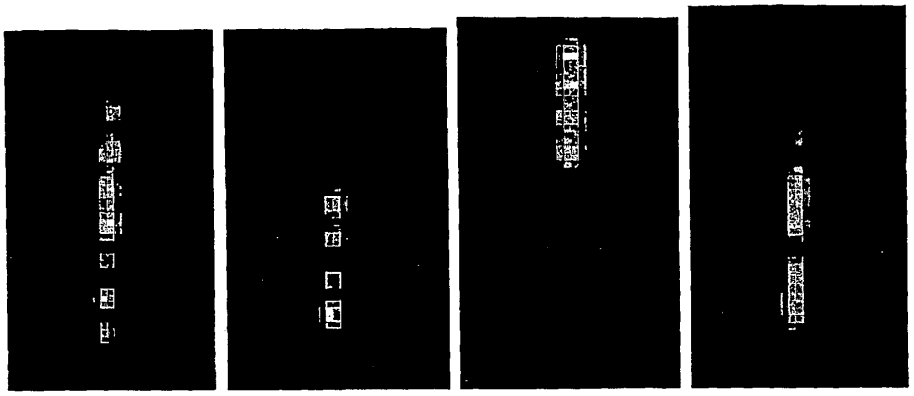


Figure 5A

Figure 5B

Figure 5C

Figure 5D

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 278

FIGURE 6

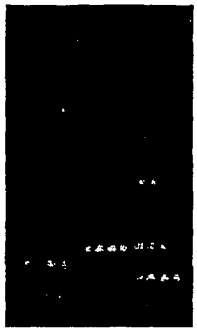


Figure 6E

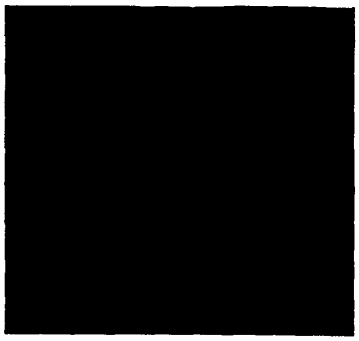


Figure 6D

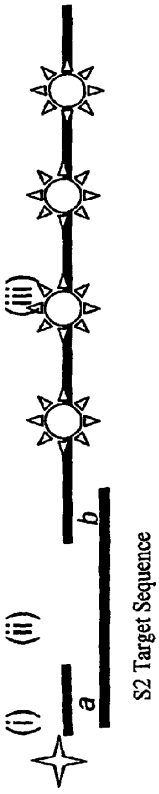


Figure 6A

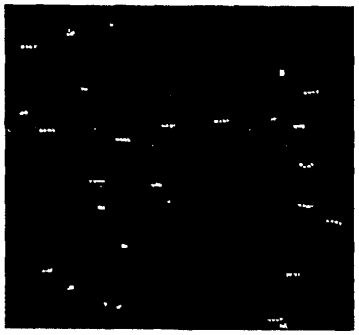


Figure 6C

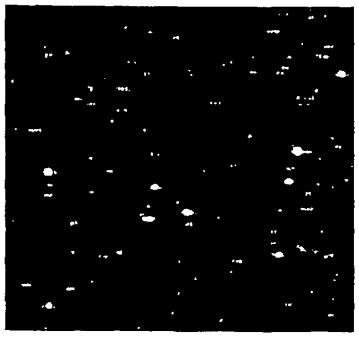
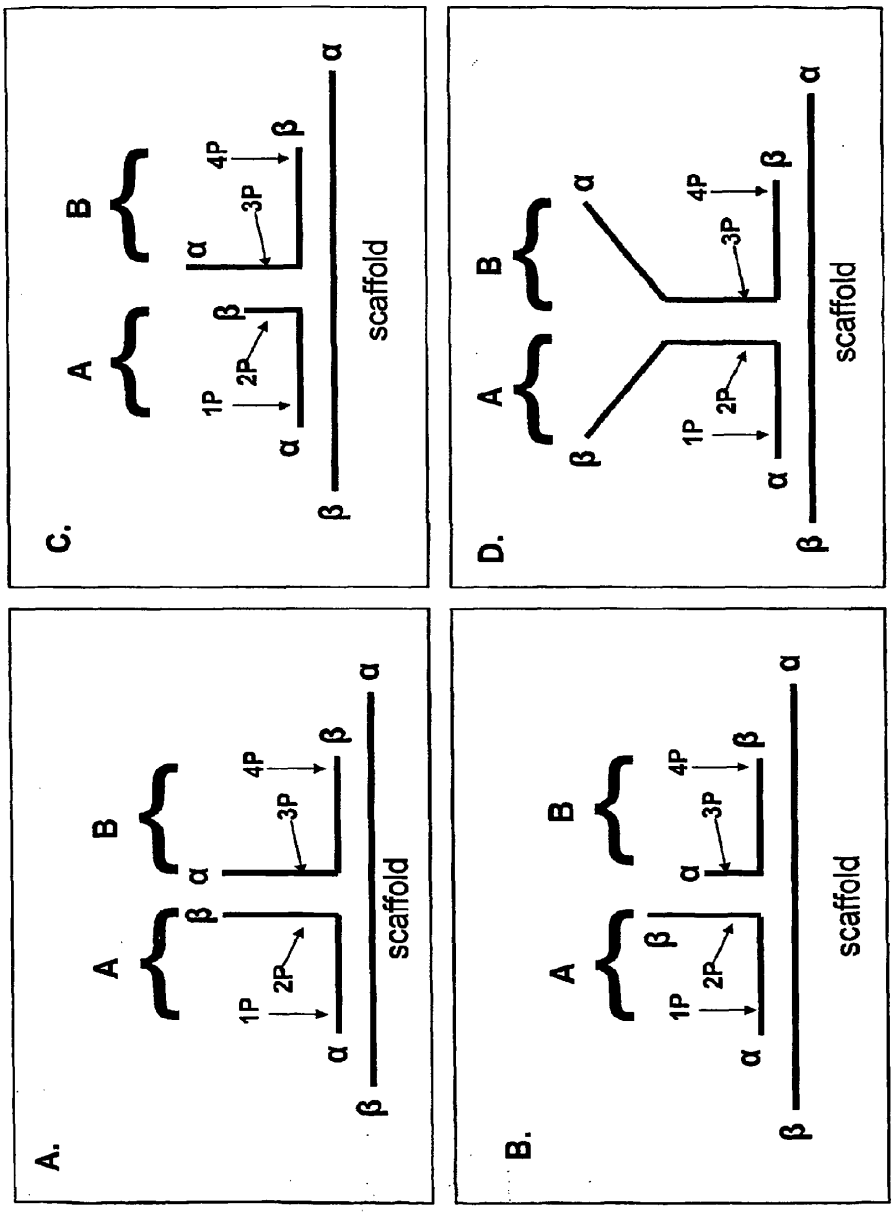


Figure 6B

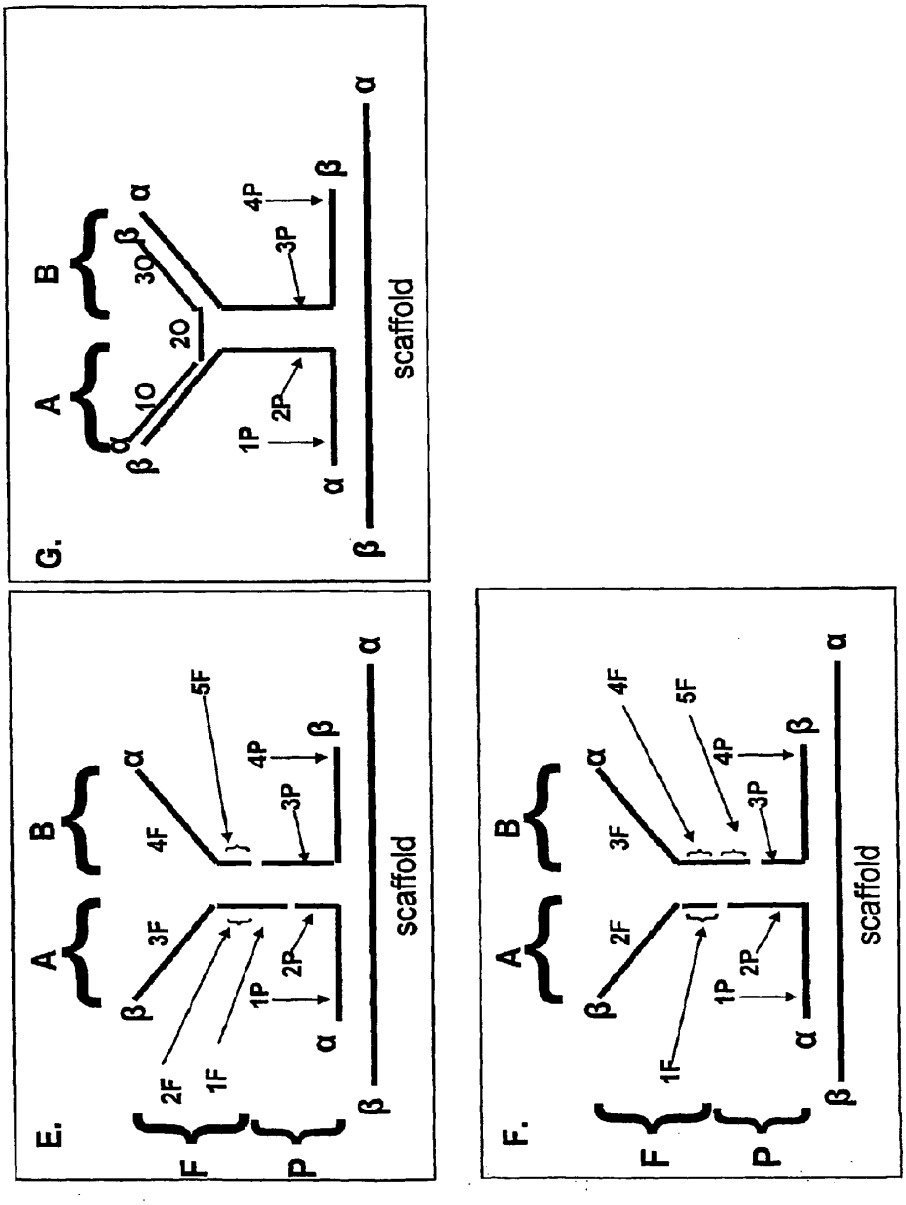
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 279

FIGURE 7A-D



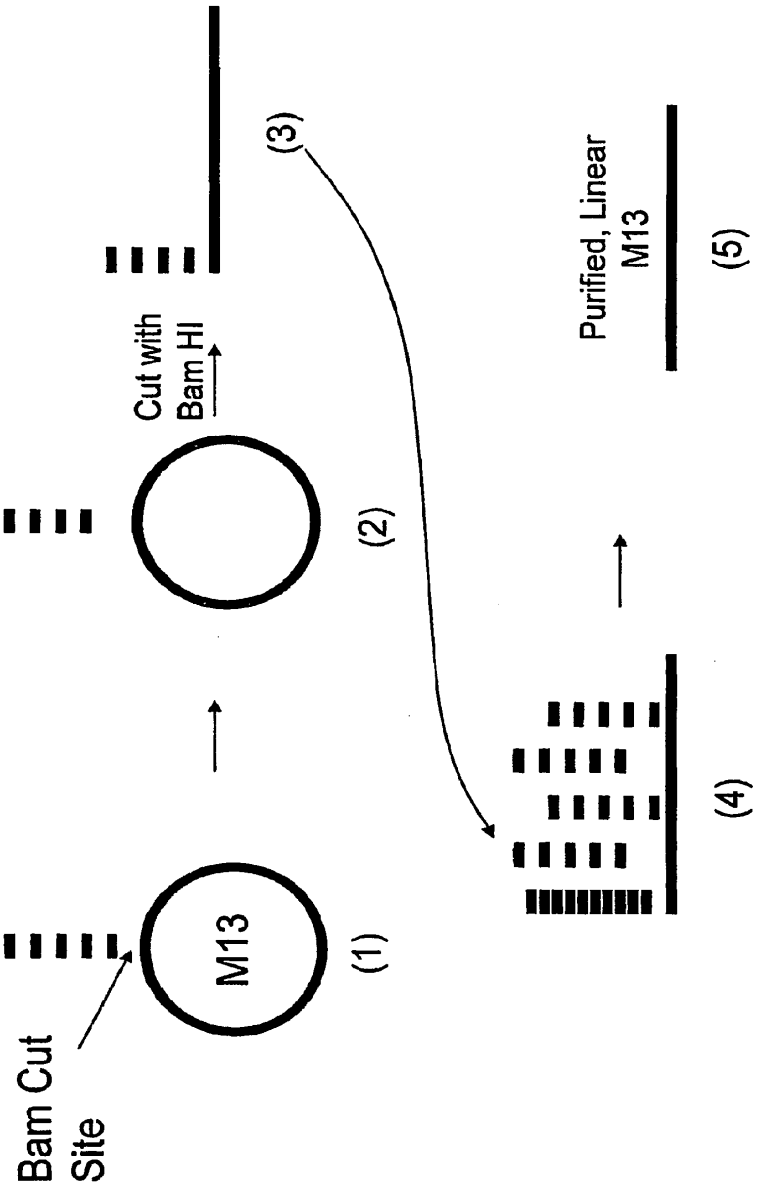
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 280

FIGURE 7E-G



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 281

FIGURE 8



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 282

FIGURE 9

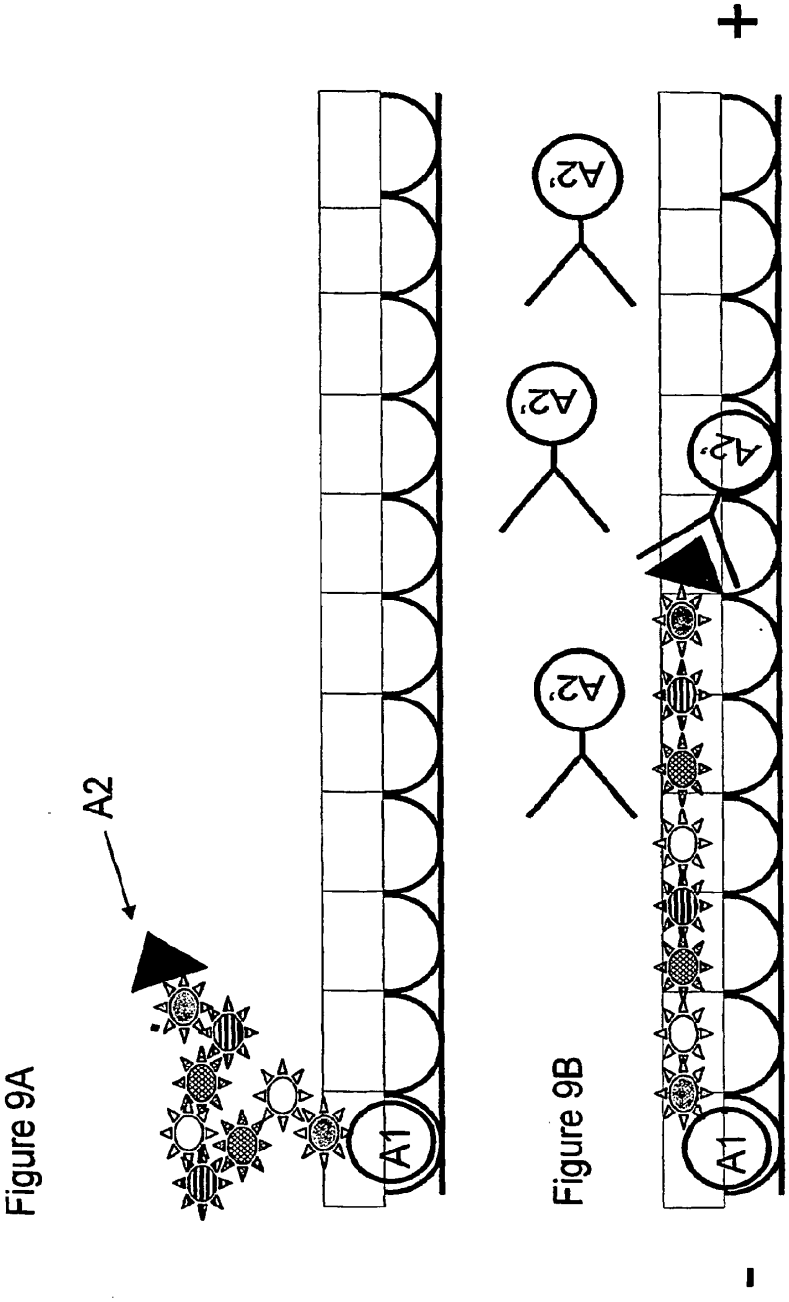


Figure 9A

Figure 9B

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 283

FIGURE 10

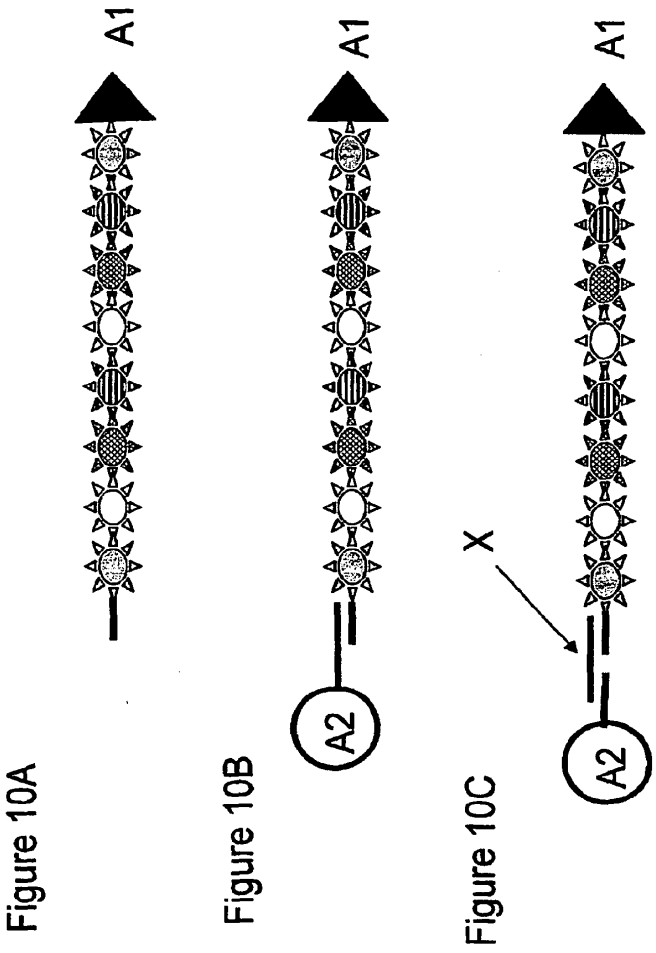


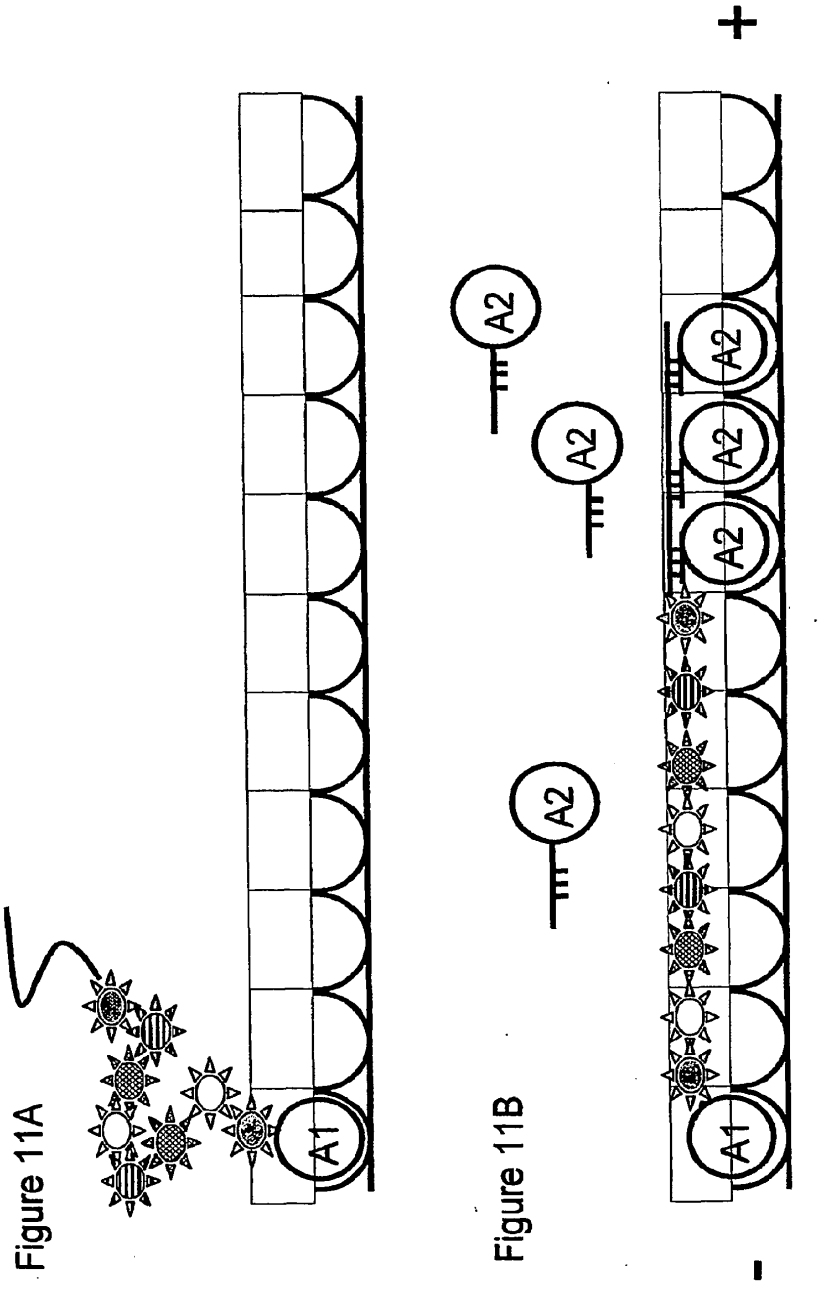
Figure 10A

Figure 10B

Figure 10C

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 284

FIGURE 11



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 285

FIGURE 12A-12B

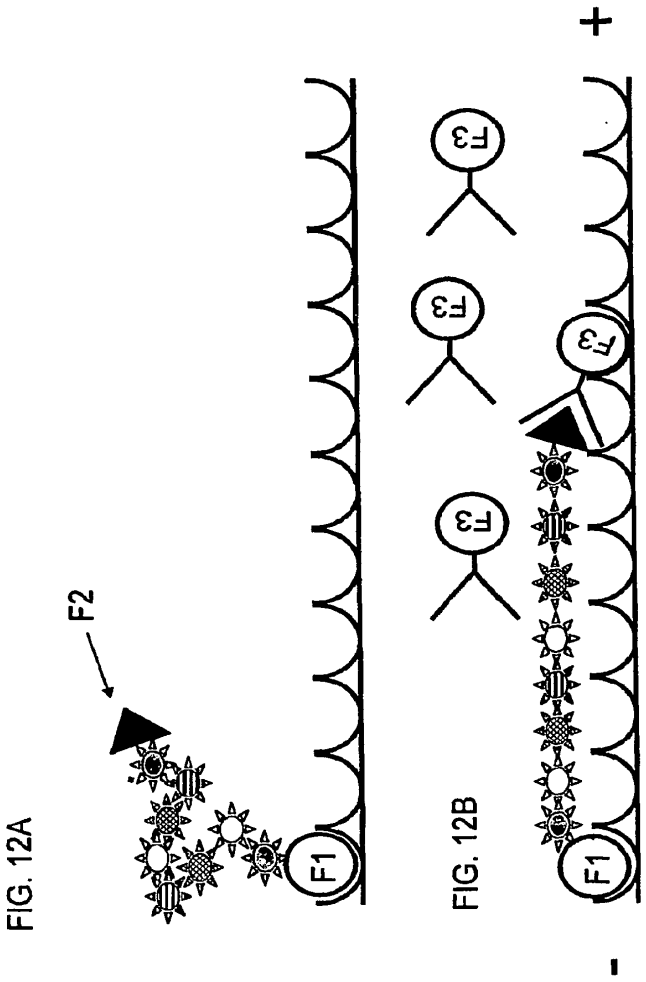


FIG. 12A

FIG. 12B

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 286

FIGURE 13A-13C

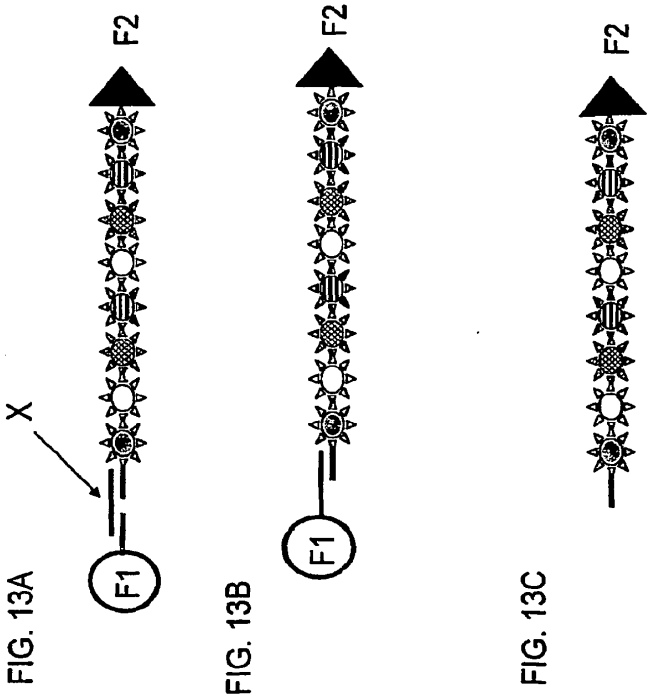
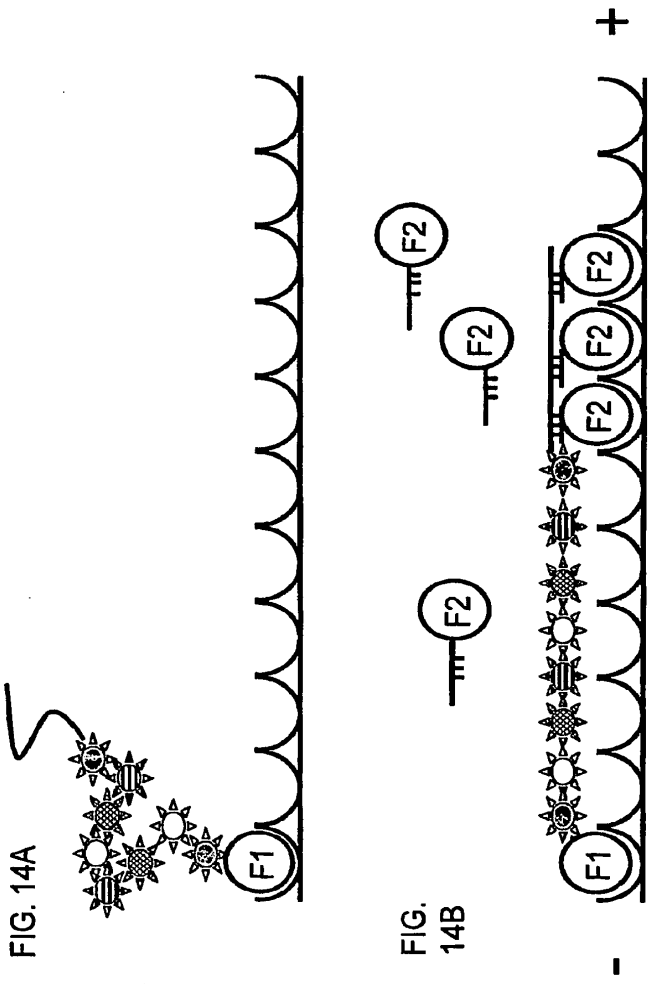


FIG. 13A FIG. 13B FIG. 13C

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 287

FIGURE 14A-14B



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 288

FIGURE 14C-14D

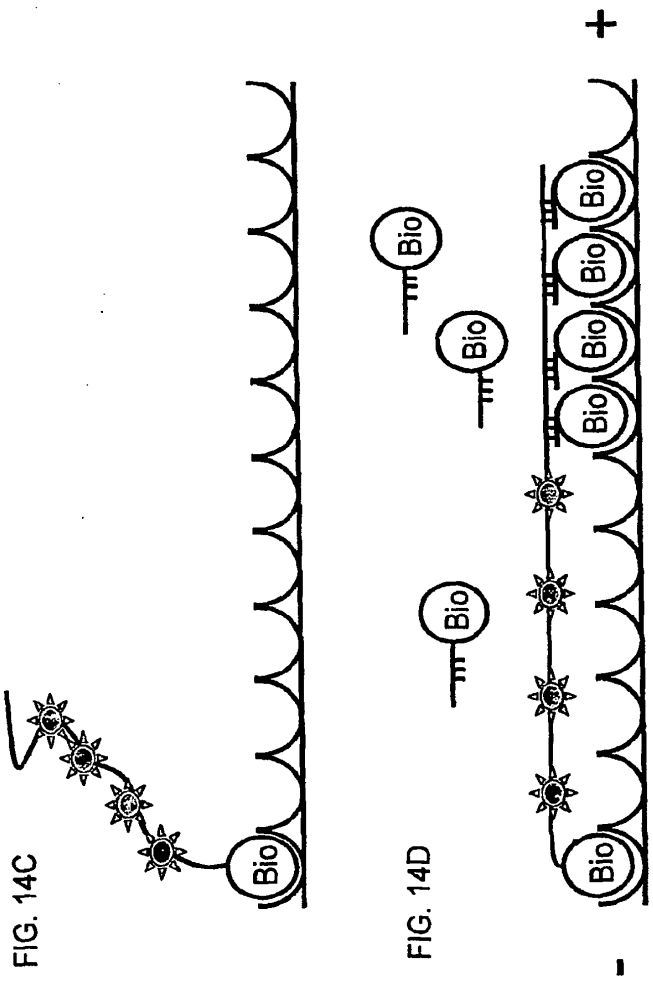


FIG. 14C

FIG. 14D

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 289

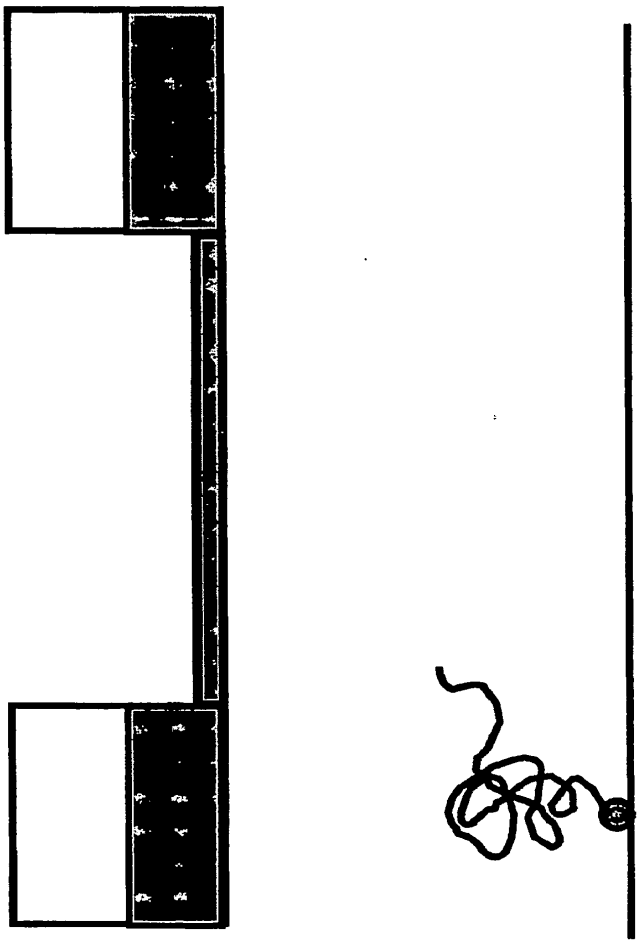


FIGURE 15A

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 290

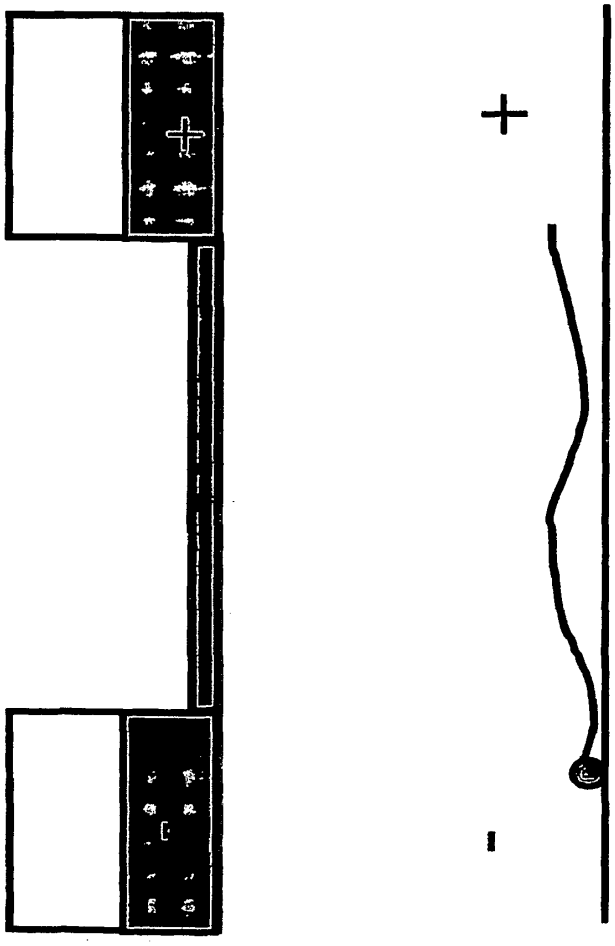


FIGURE 15B

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 291

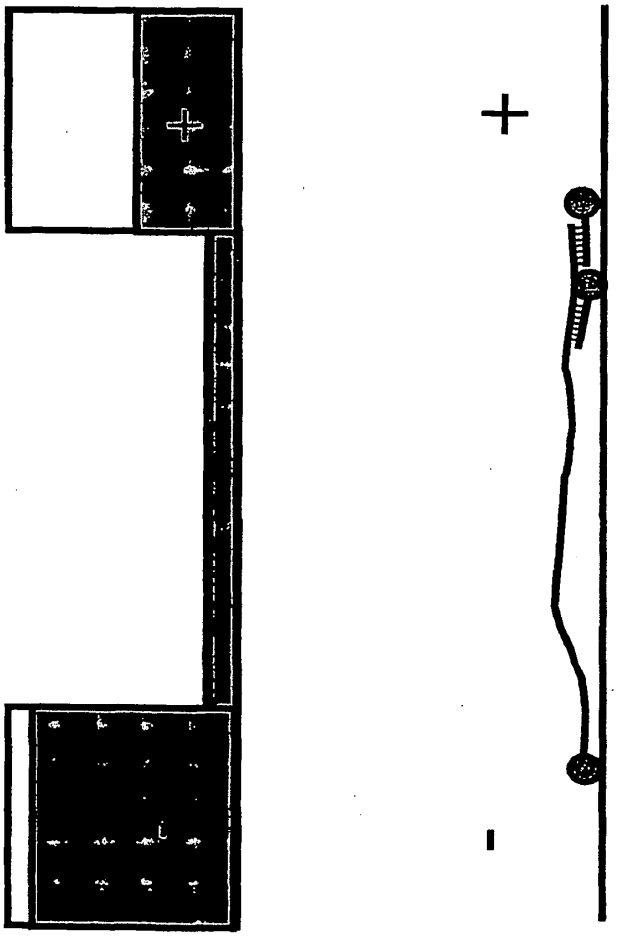


FIGURE 15C

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 292

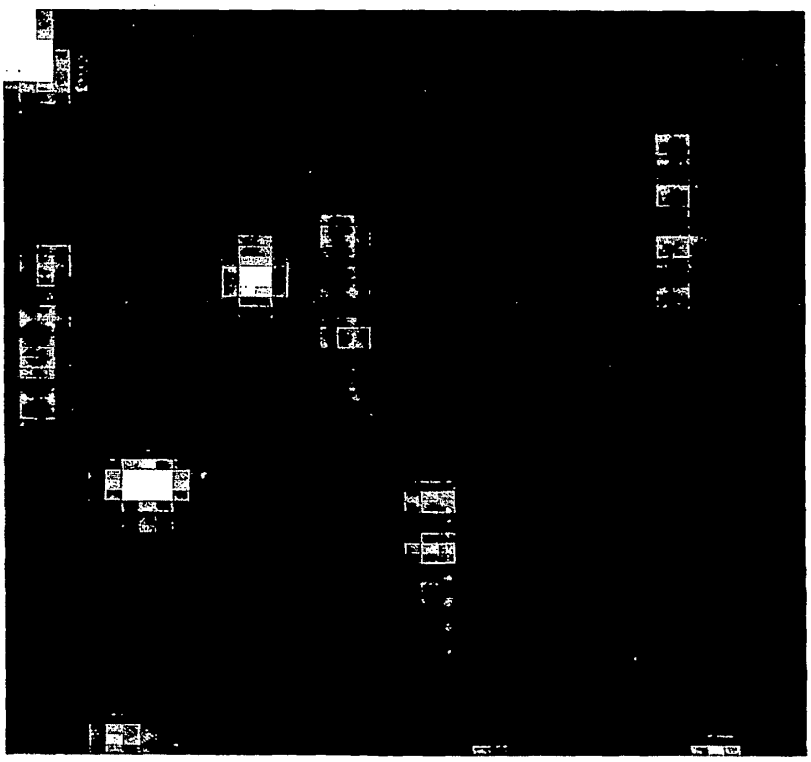
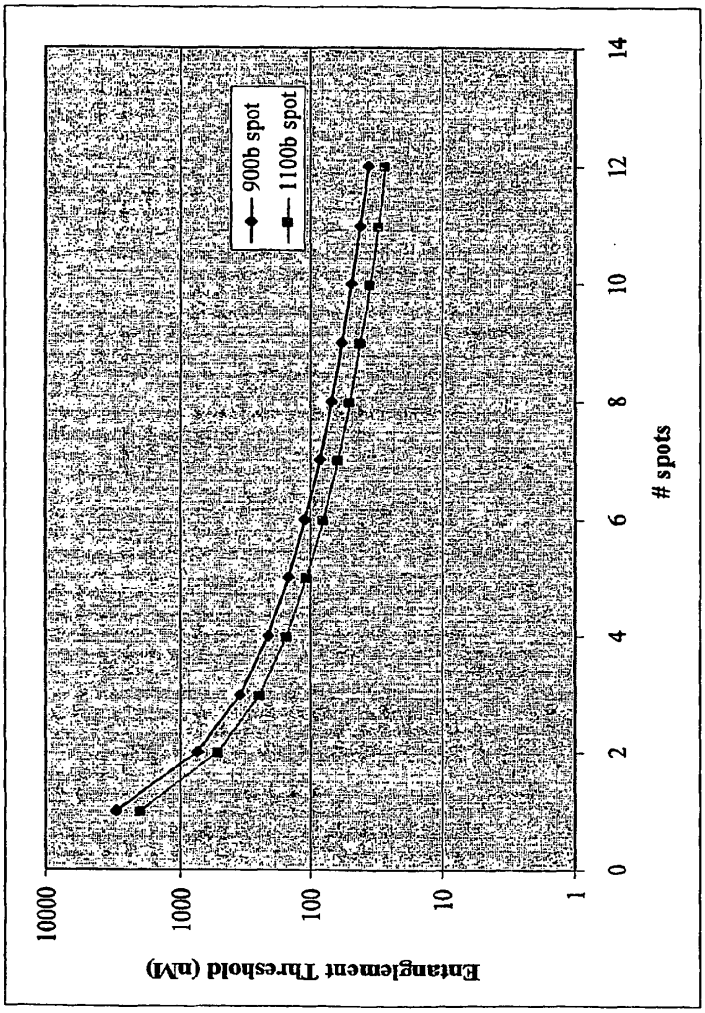


FIGURE 16

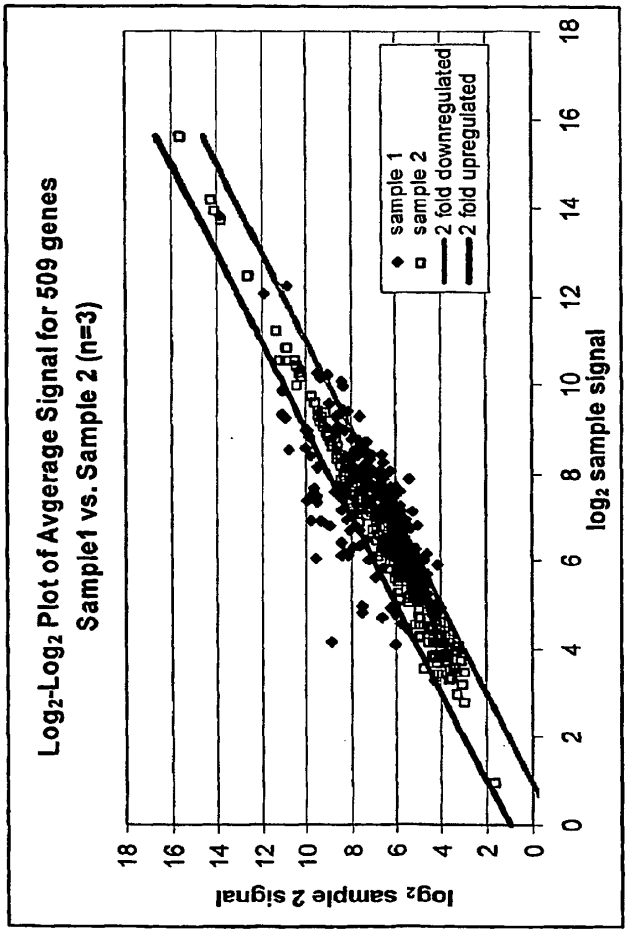
IPR2024-01177
California Institute of Technology - Ex. 2003 Page 293

FIGURE 17



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 294

FIGURE 18



IPR2024-01177
California Institute of Technology - Ex. 2003 Page 295

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 2005071401 A **[0006]**
- US 6268147 B **[0007]**
- WO 03003810 A **[0008]**
- US 20100261026 A1 **[0078]**
- US 5293050 A **[0175]**
- US 5354707 A, Chapple Sokol, **[0175]**
- US 5496934 A **[0196]**
- US 5202247 A **[0196]**
- US 5137819 A **[0196]**
- US 6001983 A **[0253]**
- US 6548255 B **[0272] [0274]**
- US 6344319 B **[0272] [0274]**
- US 6303296 B **[0272] [0274]**
- US 6265153 B **[0272] [0274]**
- US 6225055 B **[0272] [0274]**
- US 6054327 B **[0272] [0274]**
- US 5840862 B **[0272] [0274]**

Non-patent literature cited in the description

- **Schena et al.** *Science,* 1995, vol. 270, 467-470 **[0009]**
- **Lockhart et al.** *Nature Biotechnology,* 1996, vol. 14, 1675-1680 **[0009]**
- **Blanchard et al.** *Nature Biotechnology,* 1996, vol. 14, 1649 **[0009]**
- **Zuker.** *Nucleic Acids Res.,* 2003, vol. 31 (13), 3406-15 **[0135]**
- **Mathews et al.** *J. Mol. Biol.,* 1999, vol. 288, 911-940 **[0135]**
- **Hyrup et al.** *Bioorganic & Medicinal Chemistry,* vol. 4 (1), 5-23 **[0140]**
- **Hyrup et al.** *Bioorganic & Medicinal Chemistry,* 1996, vol. 4 (1), 5-23 **[0140]**
- **Perry-O'Keefe et al.** *Proc. Natl. Acad. Sci. USA,* 1996, vol. 93, 14670-675 **[0140]**
- Metal-chelate affinity chromatography. **Petty et al.** Current Protocols in Molecular Biology. Greene Publish. Assoc. & Wiley Interscience, 1996, vol. 2 **[0196]**
- **Smith.** *Methods Mol. Cell Bio.,* 1993, vol. 4, 220-229 **[0196]**
- **Guan et al.** *Gene,* 1987, vol. 67, 21-30 **[0196]**
- **Tomme et al.** *Protein Eng.,* 1994, vol. 7, 117-123 **[0196]**
- **Henegariu et al.** *Biotechniques,* 2001, vol. 31, 246-50 **[0224]**
- **Yokota et al.** *Nuc. Acids Res.,* 1997, vol. 25, 1064-1070 **[0224]**
- **Matsuura et al.** *Nuc. Acids Res.,* 2001, vol. 29, E79 **[0224]**
- **Uhlman ; Peyman.** *Chemical Review,* 1990, vol. 90 (4), 544 584 **[0253]**
- **Goodchild.** *Bioconjugate Chem.,* 1990, vol. 1 (3), 165 186 **[0253]**
- **Egholm et al.** *J. Am. Chem. Soc.,* 1992, vol. 114, 1895 **[0253]**
- **Gryaznov et al.** *J. Am. Chem. Soc.,* vol. 116, 3143 3144 **[0253]**
- **Fasman.** CRC Practical Handbook of Biochemistry and Molecular Biology. 1985, 385-392 **[0253]**
- Beilstein's Handbuch der Organischen Chemie. Springer Verlag **[0253]**
- **Matsuura et al.** *JBiomol Struct Dyn.,* 2002, vol. 20 (3), 429-36 **[0271]**
- **Ferree ; Blanch.** *Biophys J.,* 2003, vol. 85 (4), 2539-46 **[0271]**
- **Stigter ; Bustamante.** *Biophys J.,* 1998, vol. 75 (3), 1197-210 **[0271]**
- **Matsuura et al.** *Nucleic Acids Res.,* 2001, vol. 29 (16 **[0271]**
- **Ferree ; Blanch.** *Biophys J.,* 2004, vol. 87 (1), 468-75 **[0271]**
- **Bensimon et al.** *Science,* 1994, vol. 265, 2096-2098 **[0272]**
- **Henegariu et al.** *BioTechniques,* 2001, vol. 31, 246-250 **[0272]**
- **Kraus et al.** *Human Genetics,* 1997, vol. 99, 374-380 **[0272] [0273]**
- **Michalet et al.** *Science,* 1997, vol. 277, 1518-1523 **[0272] [0273]**
- **Yokota et al.** *Nucleic Acids Res.,* 1997, vol. 25 (5), 1064-70 **[0272] [0273]**
- **Otobe et al.** *Nucleic Acids Research,* 2001, vol. 29, 109 **[0272]**
- **Zimmerman ; Cox.** *Nucleic Acids Res.,* 1994, vol. 22 (3), 492-7 **[0272]**
- **Ashkin.** *Optics Letters,* 1986, vol. 11, 288-290 **[0275]**
- **Ashkin et al.** *Science,* 1987, vol. 235, 1517-1520 **[0275]**
- **Ashkin et al.** *Nature,* vol. 330, 769-771 **[0275]**

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 296

- **Perkins et al.** *Science,* 1994, vol. 264, 822-826 **[0275]**
- **Simmons et al.** *Biophysical Journal,* 1996, vol. 70 (1813-1), 22 **[0275]**
- **Block et al.** *Nature,* 1990, vol. 348, 348-352 **[0275]**
- **Grier.** *Nature,* 2003, vol. 424, 810-816 **[0275]**
- **Asbury et al.** *Electrophoresis,* 2002, vol. 23 (16), 2658-66 **[0305]**
- **Kabata et al.** *Science,* 1993, vol. 262 (5139), 1561-3 **[0305]**
- **Asbury ; van den Engh.** *Biophys J.,* 1998, vol. 74, 1024-30 **[0305]**
- Surface Treatment. Wiley Encyclopedia of Packaging Technology. John Wiley & Sons, 1997, 867 874 **[0311]**
- **Atkinson ; Smith.** Oligonucleotide Synthesis: A Practical Approach. IRL Press, 1984, vol. Solid Ph, 45 49 **[0311]**
- **Pease.** *Proc. Natl. Acad. Sci. USA,* 1994, vol. 91, 5022 5026 **[0311]**
- **Lloyd Williams et al.** Chemical Approaches to the Synthesis of Peptides and Proteins. CRC Press, 1997 **[0311]**
- **Wetmur.** *J. Annu. Rev. Biophys. Bioeng.,* 1976, vol. 5, 337-361 **[0436]**
- **Britten et al.** *Methods in Enzymology,* 1974, vol. 29E, 363-406 **[0437]**
- **Inoue, S.** Video Microcopy. Plenum Press, 1997, 30 **[0493]**
- **Gustafsson.** *Proc. Nat'l. Acad. Sci. U.S.A.,* 2005, vol. 102, 13081-13086 **[0493]**
- **Thomann.** *J. Microsc.,* 2002, vol. 21 (1), 230-248 **[0493]**

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 297

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 July 2002 (18.07.2002)

PCT

(10) International Publication Number
WO 02/056014 A2

(51) International Patent Classification[7]: G01N 33/53

(21) International Application Number: PCT/US01/49132

(22) International Filing Date:
18 December 2001 (18.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/747,003 22 December 2000 (22.12.2000) US

(71) Applicant *(for all designated States except US)*: **GLAXO GROUP LIMITED** [GB/GB]; Glaxo Wellcome House, Berkeley Avenue, Greenford, Middlesex UB6 ONN (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **NELSEN, Anita, J.** [US/US]; GlaxoSmithKline, Five Moore Drive, P.O. Box 13398, Research Triangle Park, NC 27709 (US). **PEPPERS, Lottie, L.** [US/US]; GlaxoSmithKline, Five Moore Drive, P.O. Box 13398, Research Triangle Park, NC 27709 (US). **WEINER, Michael, Phillip** [US/US]; c/o 454 Corporation, 20 Commercial Street, Branford, CT 06405 (US).

(74) Agents: **LEVY, David, J.** et al.; GlaxoSmithKline, Five Moore Drive, P.O. Box 1398, Research Triangle Park, NC 27709 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, 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, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, 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, TR), OAPI patent (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*

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 02/056014 A2

(54) Title: METHODS FOR ENCODING AND DECODING COMPLEX MIXTURES IN ARRAYED ASSAYS

(57) Abstract: The present invention is a method of encoding a complex mixture of assay constituents comprising using combinations of detectable tags and a total number of detectable tags less than the total number of constituents to be encoded. The method comprises determining the total number of constituents to be encoded; determining the number of detectable tags in each combination, wherein the number of detectable tags in each combination is more than one and less than or equal to the number of prime numbers in the number of constituents to be encoded; and determining the total number of detectable tags, wherein the total number of detectable tags equals a sum of a set of factors of the total number of constituents, wherein the number of factors equals the number of detectable tags in each combination. The encoding methods are useful in a multiplexed assay using complex mixtures of assay constituents.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 298

METHODS FOR ENCODING AND DECODING COMPLEX MIXTURES IN ARRAYED ASSAYS

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

This invention relates generally to the fields of molecular biology and chemical analysis. More specifically, the invention relates to methods of encoding and decoding complex mixtures in multiplexed assays in order to minimize time and expense necessary for assaying numerous constituents in a single assay.

BACKGROUND ART

In the past few years the genomes of several organisms have been completely (or nearly completely) sequenced, including those of *Saccharomyces cerevisiae, Drosophila melanogaster, Escherichia coli, Caenorhabditis elegans* and, most recently, the human genome. To make use of this wealth of available genomic data, rapid, high-throughput methods for analyzing all of the predicted gene products and their roles in the structural and functional organization of the cell were needed. Specifically needed were encoding and decoding means for analyzing the functional information of the thousands of genes in a complete genome.

One technology previously introduced used unique "bar-coding" tags for each of thousands of yeast genes and a silicon chip for decoding (Shoemaker et al., 1996). This method was useful for examining differential gene expression of a population of yeast strains. However, it required several thousand tags and a relatively expensive readout platform.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 299

Prior to the present invention, no multiplexed method had been provided that could be scaled to accommodate assays of varying complexity using multiplexed, inexpensive, high-throughput methods.

SUMMARY OF THE INVENTION

In accordance with the purpose(s) of this invention, as embodied and broadly described herein, this invention, in one aspect, relates to a method of encoding a complex mixture of assay constituents comprising using combinations of detectable tags and a total number of detectable tags that is less than the total number of constituents to be encoded. More specifically, the invention relates to a method further comprising determining the total number of constituents to be encoded; determining the number of detectable tags in each combination, wherein the number of detectable tags in each combination is more than one and less than or equal to the number of prime numbers in the number of constituents to be encoded; and determining the total number of detectable tags, wherein the total number of detectable tags equals a sum of a set of factors of the total number of constituents, wherein the number of factors equals the number of detectable tags in each combination.

In yet another aspect, the invention relates to a method of performing a multiplexed assay using complex mixtures of assay constituents encoded according to the encoding method of the invention. Specifically, the invention relates to a method comprising performing an assay to produce assay constituents using an array of the complex mixtures, wherein each constituent in a single complex mixture is detectably tagged with a unique combination of detectable tags; detecting which complex mixtures of assay constituents in the array have a positive response; and decoding the constituents in the complex mixtures having the positive response to determine which specific constituent or constituents are positive.

In another embodiment, the invention relates to a kit for performing a multiplexed assay using complex mixtures of encoded assay constituents,

comprising a means of detectably tagging assay constituents encoded according to the encoding method of the invention, and an arraying means for a plurality of complex mixtures, and a container therefor.

The advantages of this invention include scalability, throughput and low-cost as compared to the currently available methods. Additional advantages of the invention will be set forth in part in the description that follows, and in part will be obvious from the description, or may be learned by practice of the invention. The advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments of the invention and, together with the description, serve to explain the principles of the invention.

Figure 1 shows a schematic of the bait vector, pMW101, and the prey vector, pMAR101, used in the Yeast Two Hybrid (Y2H) study.

Figure 2 shows a schematic of the synthesis of the 96 pMAR101 prey vectors used in the Y2H study. Each 5' ZipCode was bracketed by a similar DNA sequence (5'- TGGGCGACTTCTCCAAAC -3', (SEQ ID NO:2) which was labeled the "Watson" sequence). And each 3' sequence was bracketed by a second DNA sequence (5'- CTTGCAGATTCGGCAGTT -3' (SEQ ID NO:3), which was labeled the "NCrick" sequence). PCR amplification was used to generate 96 different fragments of the Cm^r gene; each fragment with the following order: 5'-Watson-$ZipCode_{1-12}$-Cm^r – $ZipCode_{A-H}$-NCrick-3'. Fragments were cloned into the pMAR101 vector at a unique *Swa*I site.

Figure 3 shows the method of bead-based genotyping by hybridization to

Luminex beads, which was used to decode the Y2H positive wells. Following the Y2H assay, clones in positive wells were PCR amplified using biotinylated Watson and nCrick primers. For a given fragment, querying which pair of 3' and 5' ZipCodes were contained therein involved hybridizing the fragment to the cZipCodes on the microsphere. Flow cytometry was used to detect the label captured on a particular pair of microspheres.

Figure 4 shows an example of a decode of 20 PCR products hybridized to a set of 20 ZipCode beads. A set of 96 vectors, each encoding a unique region containing two ZipCodes bracketed by a Watson and nCrick was used. DNA sequence served as a PCR template in a reaction containing Watson and nCrick primers. The PCR product was then used in a microsphere-based genotyping method and both of the ZipCodes on either side of the Cmr gene were decoded by hybridization to a set of 20 different beads. Shown are the MFI values obtained from the first twenty PCR products of the 96-member set.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention may be understood more readily by reference to the following detailed description of preferred embodiments of the invention and the Examples included therein and to the Figures and their previous and following description.

Before the present compounds, compositions, articles, devices, kits, and/or methods are disclosed and described, it is to be understood that this invention is not limited to specific assay methods, specific means or methods of detection, or to particular encoding or decoding means, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

As used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a microsphere" includes mixtures of various

microspheres, reference to "an assay constituent" includes mixtures of two or more constituents, and the like.

Ranges may be expressed herein as from "about" one particular value and/or to "about" another particular value. When such a range is expressed, another embodiment includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by use of the antecedent "about," it will be understood that the particular value forms another embodiment. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint and independently of the other endpoint.

In this specification and in the claims that follow, reference will be made to a number of terms that shall be defined to have the following meanings:

"Optional" or "optionally" means that the subsequently described event or circumstance may or may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not. For example, the phrase "detectable tags optionally are contained in or coupled to microspheres" means that the detectable tags may or may not be contained in or coupled to microspheres and that the description includes both detectable tags contained in or coupled to microspheres and detectable tags otherwise used to label the desired assay constituents.

The present invention provides a method of encoding a complex mixture of assay constituents comprising using combinations of detectable tags and a total number of detectable tags that is less than the total number of constituents to be encoded. This method offers an advantage over the prior art because it reduces the number of labels necessary to detect a given number of constituents and lends itself to highly complex, multiplexed formats that are useful in high-throughput assays with pooled samples.

As used throughout, "encoding" refers to tagging an assay constituent with one or more detectable tags so that the tag(s) can be detected and the constituent

identified by decoding (i.e., attributing the detectable tag(s) to a specific assay constituent).

An assay constituent can be either a reactant or a product of the assay. The constituents are selected from the group consisting of proteins, peptides, amino acids, small molecules, nucleotides, fatty acids, sugars, cofactors, receptors, receptor ligands, protein domains, oligonucleotides, transcription factors, nucleic acids, and small compounds.

As used throughout, "an assay" can be a chemical assay, protein assay, pharmacologic assay, hybrid assay (e.g., yeast two hybrid, prokaryotic two-hybrid, reverse-two hybrid, or three-hybrid assay), display assay (e.g., phage display-, F pilli- and lacI-fusion), protein readout assay, binding assay (ligand, nucleic acid, antibody, small molecule, or small compound binding assay), cell-based assay, genomic assay, read-out assay (transcriptional or protein read-out assay), or the like.

A "detectable tag" refers to any label that can be detected with a detection means and can include the absence of a label. Thus, if one hundred thousand assay constituents are to be encoded, the methods of the present invention provide that less than one hundred thousand detectable tags are used, even if one of those tags is the absence of a label. Preferably, the detectable tags are directly or indirectly coupled to the constituents. The detectable tags as used in the methods of the present invention optionally are contained in or coupled to a solid support that binds the constituents either directly or through an intermediary. Thus, the detectable tags can be coupled to a non-mobile solid support, like a plate or a chip, or a mobile solid support, like microspheres. Optionally, each detectably tagged microsphere used in the methods of the present invention is coupled to a means of specifically binding a constituent. For example, in one embodiment, the coupled means is a nucleic acid (called a "ZipCode"), which is complementary to a nucleic acid in or bound to the constituent to be encoded.

Preferably the detectable tags are selected from the group consisting of radiolabels, dyes, fluorescent labels, Quantum Dot® (Quantum Dot Corp.), and

combinations thereof. "Dyes" include, but are not limited to, chemiluminescent, magnetic, and radiofrequency labels.

Optionally, the method of the present invention further comprises determining the total number of constituents to be encoded; determining the number of detectable tags in each combination, wherein the number of detectable tags in each combination is more than one and less than or equal to the number of prime numbers in the number of constituents to be encoded; and determining the total number of detectable tags, wherein the total number of detectable tags equals a sum of a set of factors of the total number of constituents, wherein the number of factors equals the number of detectable tags in each combination. For example, if 100,000 genes are to be screened using a traditional method of encoding for each assay constituent to be screened, then 100,000 different detectable tags would have to be used in a traditional one dimensional assay (i.e., one detectable tag for each constituent). In the present method, however, the number of different detectable tags can be reduced using a multi-step process. Where the total number of constituents to be screened is 100,000, the number of detectable tags in each combination can be any number between two and ten, because ten is the number of prime numbers that are multiples of 100,000 (i.e., 2 X 2 X 2 X 2 X 2 X 5 X 5 X 5 X 5 X 5 = 10 prime numbers). Thus, if three detectable tags will be used in each combination to encode a single constituent, the total number of detectable tags needed is calculated by determining three factors of the total number of constituents (e.g., 10 X 100 X 100) and adding those three factors together (i.e., 210) to determine the total number of detectable tags. Using this paradigm, the entire 100,000 genes could be screened using a total of 210 detectable tags.

Table 1: Examples of assay arrays and detectable tags for an assay of 100,000 constituents

Number of detectable tags in each combination	Example number of constituents to be screened (factors)	Example total number of detectable tags needed
1	100,000 (100,000)	100,000
2	100,000 (250 X 400)	650
	(500 X 200)	700
	(160 X 625)	785
	(125 X 800)	925
	(100 X 1,000)	1,100
	(80 X 1250)	1,330
	(40 X 2,500)	2,540
	(50 X 2000)	2050
	(32 X 3125)	3,157
	(25 X 4000)	4,025
	(20 X 5,000)	5,020
	(16 X 6,250)	6,266
	(10 X 10,000)	10,010
	(8 X 12,500)	12,508
	(5 X 20,000)	20,005
	(4 X 25,000)	25,004
	(2 X 50,000)	50,002
3	100,000 (50 X 25 X 80)	155
	(100 X 25 X 40)	165
	(32 X 25 X 125)	182
	(20 X 125 X 40)	185
	(16 X 125 X 50)	191
	(10 X 100 X 100)	210
	(160 X 25 X 25)	210
	(10 X 125 X 80)	215
	(8 X 125 X 100)	233
	(20 X 25 X 200)	245
	(160 X 5 X 125)	290
	(16 X 25 X 250)	291
	(100 X 5 X 200)	305
	(50 X 8 X 250)	308
	(4 X 125 X 200)	329
	(250 X 5 X 80)	335
	(250 X 4 X 100)	354
	(10 X 25 X 400)	435
	(250 X 2 X 200)	452
	(50 X 5 X 400)	455
	(400 X 2 X 125)	527
	(8 X 25 X 500)	533
	(500 X 5 X 40)	545
	(50 X 4 X 500)	554
	(100 X 2 X 500)	602
	(8 X 625 X 20)	653
	(16 X 625 X 10)	651

	(32 X 5 X 625)	662
	(4 X 625 X 40)	669
	(2 X 625 X 80)	707
	(800 X 5 X 25)	830
	(10 X 10 X 1000)	1,020
	(20 X 5 X 1000)	1,025
	(4 X 25 X 1000)	1,029
	(50 X 2 X 1000)	1,052
	(10 X 8 X 1250)	1,268
	(16 X 5 X 1250)	1,271
	(1250 X 4 X 20)	1,274
	(1250 X 2 X 40)	1,292
	(10 X 5 X 2000)	2,015
	(2 X 25 X 2000)	2,027
	(8 X 5 X 2500)	2,513
	(10 X 4 X 2500)	2,514
	(20 X 2 X 2500)	2,522
	(4 X 3125 X 8)	3,137
	(2 X 3125 X 16)	3,143
	(4,000 X 5 X 5)	4,010
	(2 X 125 X 4000)	4,127
	(10 X 2 X 5000)	5,012
	(4 X 5 X 5000)	5,009
	(6,250 X 4 X 4)	6,258
	(6,250 X 2 X 8)	6,260
	(2 X 5 X 10,000)	10,007
	(12,500 X 2 X 4)	12,506
	(25,000 X 2 X 2)	25,004
4	100,000 (e.g., 10 X 10 X 10 X 100)	130
5	100,000 (e.g., 10 X 10 X 10 X 10 X 10)	50
6	100,000 (e.g., 2 X 5 X10 X 10 X 10 X 10)	47
7	100,000 (e.g., 2 X 2 X 5 X 5 X 10 X 10 X 10)	44
8	100,000 (e.g., 2 X 2 X 2 X 5 X 5 X 5 X 10 X 10)	41
9	100,000 (e.g., 2 X 2 X 2 X 2 X 5 X 5 X 5 X 5 X 10)	38
10	100,000 (2 X 2 X 2 X 2 X 2 X 5 X 5 X 5 X 5 X 5)	35

As demonstrated by the above table, the total number of detectable tags is minimized by selecting factors of the total number of constituents that are equal or approximate. Thus, in one embodiment of the present method, the total number of detectable tags is minimized by selecting factors of the total number of constituents that are equal or approximate.

Using the encoding method of the present invention, an assay can be designed based on the total number of detectable tags available, based on the total number of tags in each combination, based on the arraying means (e.g., the number of wells on a plate that can be read using automated readers), or a combination. Thus, if 100,000 constituents are to be assayed and only 35 total detectable tags are available, then there must be 10 detectable tags in each combination. Alternatively, if there are practical limitations to the number of tags that can be detected in combination, then the assay could limit the number of detectable tags in combination to, for example, three and the total number of detectable tags could be, for example, 210.

The invention further provides a method of performing a multiplexed assay using complex mixtures of assay constituents encoded according to the encoding method of the invention. Specifically, the method comprises performing an assay to produce assay constituents using an array of the complex mixtures, wherein each constituent in a single complex mixture is detectably tagged with a unique combination of detectable tags; detecting which complex mixtures of assay constituents in the array have a positive response; and decoding the constituents in the complex mixtures having the positive response to determine which specific constituent or constituents are positive. As used herein, "an array" includes a multiwell plate or any other arraying means. Thus, an array using a multiwell plate can be eight wells in one dimension and twelve wells in another dimension as in a 96 well plate. An array could also be sixteen wells in one dimension and twelve wells in another dimension, using two 96 well plates.

The detection means is selected as specific for the detectable tags. For example,

if the detectable tag is fluorescent and is contained in or coupled to microspheres, then flow cytometry with a fluorescence detection device or a FAC sorter can be used to detect and distinguish a tag or combination of tags. When a specific mixture (e.g., a complex mixture in a specific well in an assay plate) has a positive response, then that particular mixture is decoded to identify the positive constituent in that mixture. For example, if the well contained ten genes to be screened, then the decoding method would identify which of the ten genes had a positive response. Thus, the decoding is performed by detecting and distinguishing with a detection device the detectable tags in each complex mixture of constituents.

In one embodiment the method of encoding and decoding is used with a complex mixture of arrayed cDNA clones in a yeast two-hybrid analysis. The steps comprise using an array of complex mixtures of yeast host cells comprising an encoded set of cDNAs made by cloning each individual cDNA into a member of a set of vectors, wherein each member of the set of vectors comprises a yeast two-hybrid activation domain and a selected pair of identifying nucleic acid sequences ("ZipCodes"), wherein the selected pair of identifying nucleic acids is specific for each individual cDNA, and wherein the yeast host cells containing the set of vector are combined to create complex mixtures of cDNA clones; mating the arrayed host cells with a yeast expressing a bait protein and one or more reporter genes; detecting an interaction or absence of an interaction between the bait protein and the activation domains in each complex mixture of the array by determining the expression of the reporter gene or genes; performing PCR amplification of each complex mixture that shows an interaction, wherein the PCR amplification is performed using labeled primers; and decoding the PCR products using a genotyping assay. In one embodiment, the first member of the selected pair of identifying nucleic acids is at the 5' end of an antibiotic resistance gene and the second member is at the 3' end of the antibiotic resistance gene. Preferably, each vector in the set has two primer nucleic acid sequences present in each vector, wherein the first primer nucleic acid is at the 5' end of the first member of the identifying pair and the second primer nucleic acid is at the 3' end of the second member of the identifying pair. In one

embodiment, the antibiotic resistance gene is a chloramphenicol gene.

In one embodiment of the yeast two hybrid method, each identifying nucleic acid is 25 bases. In another embodiment, twenty different identifying nucleic acids are used in combinations to form 96 different pairs of identifying nucleic acids. Thus, each complex mixture can contain up to about 96 different cDNAs and the array of complex mixtures of cDNA clones in host cells can contain up to about 9,220 different cDNAs.

In one embodiment of the yeast two hybrid method, the reporter gene is a β-galactosidase gene. In another embodiment the reporter gene is a Leu2 gene. In yet another embodiment, the reporter genes are both the β-galactosidase gene and the Leu2 gene.

The genotyping assay as used in the yeast two hybrid method comprises contacting, under conditions that allow formation of hybridization products, the labeled PCR products of each mixture with a set of microspheres, wherein each member of the set of microspheres is distinguishably labeled and is coupled with a capture nucleic acid complementary to one of the identifying nucleic acid sequences; detecting the label of the PCR product and the label of the microsphere in two or more hybridization products. The presence of a labeled PCR product in two different hybridization products indicates the cDNA specific to the pair of identifying nucleic acid sequences. Preferably, the distinguishable label of the microsphere is a fluorescent label, wherein the PCR product is fluorescently labeled, and wherein the fluorescent label of the microsphere and the PCR product can be detected in the same reaction product or products.

In one embodiment of the yeast two hybrid method, the microspheres are carboxylated and amino groups at the 5' end of the capture nucleic acids are coupled to the carboxyl groups.

Preferably, the capture nucleic acid further comprises a luciferase cDNA. Optionally, the luciferase cDNA has the sequence CAGGCCAAGTAACTTCTTCG (SEQ ID NO:1). The capture oligonucleotide can be directly coupled to the microsphere or can be indirectly coupled to the microsphere by a carbon spacer.

In the yeast two hybrid method, the label of the PCR product and the label of the microsphere in two or more hybridization products is preferably detected using flow cytometry.

The present invention further provides a kit for performing a multiplexed assay using complex mixtures of encoded assay constituents, comprising a means of detectably tagging assay constituents encoded according to the encoding method of the invention, and an arraying means for a plurality of complex mixtures, and a container therefor. The means of detectably tagging assay constituents can include, for example, a set of microspheres, wherein each member of the set of microspheres is detectably tagged and binds selectively to an assay constituent. For example, the kit can comprise a set of detectably tagged microspheres that bind selectively to cDNA clones in a yeast two-hybrid analysis. The kit for performing a yeast two hybrid can further comprise one or more of the following: a set of yeast vectors comprising a reporter gene, a yeast two-hybrid activation domain, and a selected pair of identifying nucleic acid sequences, wherein the selected pair of identifying nucleic acids is specific for each vector; a means for homologously recombining the cDNAs to be encoded into the vectors of the set into yeast host cells; a means for combining the yeast host cells containing the set of vectors to create the complex mixture of cDNA clones; a set of yeast bait cells expressing a bait protein; a means for mating the yeast cells containing the set of vectors and the yeast bait cells; a set of labeled PCR primers; or a set of microspheres, wherein each member of the set of microspheres is detectably labeled and is coupled with a capture nucleic acid complementary to one of the identifying nucleic acid sequences.

Experimental

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how the compounds, compositions, articles, devices and/or methods claimed herein are made and evaluated, and are intended to be purely exemplary of the invention and are not

intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers (e.g., amounts, temperature, etc.), but some deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, temperature is in °C or is at ambient temperature, and pressure is at or near atmospheric.

Example 1: Yeast 2-Hybrid Analysis

Reagents. Restriction and DNA modification enzymes were purchased from various manufacturers and used according to their recommendations. AmpliTaq Gold DNA polymerase and Big Dye Terminator Cycle Sequencing reagent were purchased from Applied Biosystems (Foster City, CA, USA). Unmodified oligonucleotides were purchased from either Keystone Biosource (Camarillo, CA, USA), MWG Research (High Point, NC, USA) or IDT (Coralville, IA, USA). 2-[N-Morpholino]ethanesulfonic acid (MES) and 1-Ethyl-3-(3-Dimethylaminopropyl) carbodiimide Hydrochloride (EDC) were purchased from Sigma (St. Louis, IL, USA) and Pierce (Rockford, IL, USA), respectively. Streptavidin phycoerythrin was purchased from Becton Dickinson (San Jose, CA, USA). Yeast cell preparation and transformation reagents were purchased from Zymo Research (Orange, CA, USA).

Preparation of microspheres. Carboxylated fluorescent polystyrene microspheres (5.5 μm in diameter) were purchased from the Luminex Corp. (Austin, TX, USA). Oligonucleotides, synthesized to contain a 5' amino group, $C_{(15\text{-}18)}$ spacer, 20 base luciferase sequence and 25 base Zipcode-complimentary sequence, were ordered from Oligos, Etc. (Wilsonville, OR, USA) or from Applied Biosystems. Oligonucleotides were covalently coupled to the microspheres as described by Ianonne and co-workers (Chen et al. (2000); Iannone et al (2000)).

Yeast strains, plasmids, and media. Yeast strains EGY48 and L40 have been described (Finley & Brent (1994)). Plasmids pHybLex/Zeo, pYesTrp2 and

pMW101 have also been described (Finley & Brent (1994); Gyuris et al (1993); Watson et al (1996)). Selective yeast media were prepared as described in Gyuris et al. (1993). The plasmid pBC SK+ was purchased from Stratagene (La Jolla, CA, USA).

Construction of pMAR101 and pMAR101 derivatives. To construct pMAR101, a PCR product derived from the amplification of pMW101 with forward primer (5'- GCCGAAGCTTGCGGTTGGGGTATTCGCAACGGCGACTGG -3') (SEQ ID NO:28) and reverse primer (5'-ATACGCATGCAATTCGCCCGGAATTAGCTT GGCTGCAGGT –3') (SEQ ID NO:29) was digested with restriction endonucleases *Hin*dIII and *Sph*I and ligated overnight at 16 °C with *Hin*dIII, *Sph*I-digested, agarose gel purified plasmid pYesTrp2 (Invitrogen, Carlsbad, CA). Addition of this PCR product incorporates regions of approximately 23 bases adjacent to and on either side of the multiple cloning site. The resulting plasmid enables simultaneous homologous recombination of amplified genes into both bait and prey vectors (Figure 1). To construct pMAR101.1-pMAR101.96 twenty primers (Table 2) containing a common sequence for amplification, a 25 base zipcode and a terminal end of the chloramphenicol (Cm^r) gene were used to amplify the Cm^r marker from the plasmid pBC SK+. Following amplification under standard conditions in Optiprime buffer #5 (Stratagene), 2 units of *Pfu* polymerase were added to each of the 50 µl reactions and incubated at 72 °C for 20 minutes. The resulting 96 unique blunt-ended fragments were ligated with *Swa*I digested pMAR101 for 16 hours at 16 °C. Ligation products were transformed into electrocompetent DH10B cells (LTI, Gaitherburg, MD, USA) and clones were selected on LB agar plates containing carbenicillin (50 µg/ml) and chloramphenicol (12.5 µg/ml) (Figure 2). Colonies were screened by PCR to confirm incorporation of the cassette as well as to select for uniform orientation of the cassette. Common sequence primers, "Watson" (5'-TGGGCGACTTCTCCAAAC-3') (SEQ ID NO:2) and "nCrick" (5'-CTTGCAGATTCGGCAGTT-3') (SEQ ID NO:3), were used to confirm that the

1241 bp fragment was incorporated into the plasmid. Primers Watson and a plasmid-specific oligo, "pYesTrp Forward" (Invitrogen) were used to screen for the orientation of the Cm^r gene.

Table 2. DNA Primers and Associated ZipCode Sequences

	DNA Sequence[a]		
Primer	**Watson/NCrick**	**ZipCode**	**Cam gene**
1 (SEQ ID NO:4)	TGGGCGACTTCTCCAAAC	**GATGATCGACGAGACACTCTCGCCA**	CTGTGACGGAAGATCACTTCGC
2 (SEQ ID NO:5)	TGGGCGACTTCTCCAAAC	**CGGTCGACGAGCTGCCGCGCAAGAT**	CTGTGACGGAAGATCACTTCGC
3 (SEQ ID NO:6)	TGGGCGACTTCTCCAAAC	**GACATTCGCGATCGCCGCCCGCTTT**	CTGTGACGGAAGATCACTTCGC
4 (SEQ ID NO:7)	TGGGCGACTTCTCCAAAC	**CGGTATCGCGACCGCATCCCAATCT**	CTGTGACGGAAGATCACTTCGC
5 (SEQ ID NO:8)	TGGGCGACTTCTCCAAAC	**GCTCGAAGAGGCGCTACAGATCCTC**	CTGTGACGGAAGATCACTTCGC
6 (SEQ ID NO:9)	TGGGCGACTTCTCCAAAC	**CACCGCCAGCTCGGCTTCGAGTTCG**	CTGTGACGGAAGATCACTTCGC
7 (SEQ ID NO:10)	TGGGCGACTTCTCCAAAC	**GTAAATCTCCAGCGGAAGGGTACGG**	CTGTGACGGAAGATCACTTCGC
8 (SEQ ID NO:11)	TGGGCGACTTCTCCAAAC	**CTTTTCCCGTCCGTCATCGCTCAAG**	CTGTGACGGAAGATCACTTCGC
9 (SEQ ID NO:12)	TGGGCGACTTCTCCAAAC	**GGCTGGGTCTACAGATCCCCAACTT**	CTGTGACGGAAGATCACTTCGC
10 (SEQ ID NO:13)	TGGGCGACTTCTCCAAAC	**GAACCTTTCGCTTCACCGGCCGATC**	CTGTGACGGAAGATCACTTCGC
11 (SEQ ID NO:14)	TGGGCGACTTCTCCAAAC	**TTTCGGCACGCGCGGGATCACCATC**	CTGTGACGGAAGATCACTTCGC
12 (SEQ ID NO:15)	TGGGCGACTTCTCCAAAC	**CTCGGTGGTGCTGACGGTGCAATCC**	CTGTGACGGAAGATCACTTCGC
A (SEQ ID NO:16)	CTTGCAGATTCGGCAGTT	**TCAACGTGCCAGCGCCGTCCTGGGA**	CTCCACGGGGAGAGCCTGAGCA
B (SEQ ID NO:17)	CTTGCAGATTCGGCAGTT	**GCGAAGGAACTCGACGTGGACGCCG**	CTCCACGGGGAGAGCCTGAGCA
C (SEQ ID NO:18)	CTTGCAGATTCGGCAGTT	**CGGGGATACCGATCTCGGGCGCAC**	ACTCCACGGGGAGAGCCTGAGCA
D (SEQ ID NO:19)	CTTGCAGATTCGGCAGTT	**GGAGCTTACGCCATCACGATGCGAT**	CTCCACGGGGAGAGCCTGAGCA

E CTTGCAGATTCGGCAGTT**CGTGGCGGTGCGGAGTTTCCCCGAA**CTCCACGGGGAGAGCCTGAGCA (SEQ ID NO:20)

F CTTGCAGATTCGGCAGTT**CGATCCAACGCACTGGCCAAACCTA**CTCCACGGGGAGAGCCTGAGCA (SEQ ID NO:21)

G CTTGCAGATTCGGCAGTT**CTGAATCCTCCAACCGGGTTGTCGA**CTCCACGGGGAGAGCCTGAGCA (SEQ ID NO:22)

H CTTGCAGATTCGGCAGTT**TTCGGCGCTGGCGTAAAGCTTTTGG**CTCCACGGGGAGAGCCTGAGCA (SEQ ID NO:23)

---

[a]DNA Sequences are as follows: 3' cam gene, TCCACGGGGAGAGCCTGAGCA (SEQ ID NO:24); 5' cam gene, CTGTGACGGAAGATCACTTCGC (SEQ ID NO:25); Watson, TGGGCGACTTCTCCAAAC (SEQ ID NO:2); NCrick, CTTGCAGATTCGGCAGTT (SEQ ID NO:3). The ZipCode sequences (in bold) are shown between the Watson/NCrick and Cam sequences.

Preparation of pMAR101.1-pMAR101.96. pMAR101.x plasmid DNAs were purified using Qiatip-500 columns (Qiagen, Valencia, CA, USA). The resulting DNAs were digested with *Eco*RI and *Xho*I restriction enzymes and purified on 1% preparative agarose gels. Digested plasmid was transformed into competent EGY48 cells and plated on agar plates containing YNB –Trp + glucose to determine background.

Cloning by Homologous Recombination. Plasmids were constructed *in vivo* in yeast as described by Oldenburg et al. (1997). Briefly, genes of interest were amplified from plasmid DNAs isolated from commercially available cDNA libraries. Primers for amplification were designed to include portions of both the pMAR101 plasmid as well as portions of the gene of interest. The forward primer contained 23 bases of vector sequence immediately adjacent to and 5' of the *Eco*RI restriction site (GCAACGGCGACTGGCTGGAATTC) (SEQ ID NO:26) fused to approximately 25 bases of the 5' end of the gene to be amplified. This primer does not require a start codon, but does require the gene to be in-frame with the *Eco*RI site. The reverse primer contained 23 bases of vector sequence adjacent to and 3' of the *Xho*I site (GCTTGGCTGCAGGTCGACTCGAG) (SEQ ID NO:27) fused to

approximately 25 bases of the 3' end of the gene to be amplified. The 3' primer does require a termination codon. Amplification was carried out in 25 µl reactions, each containing 100 ng cDNA template, 200 nm primers, 1X RedTaq buffer (Sigma), 200 µM dNTP mix (Sigma), 0.75 units Red*Taq* polymerase (Sigma) and 0.125 units *Pfu* polymerase. The $MgCl_2$ concentration was adjusted to a final concentration of 3.0 mM. Reactions were amplified for 30 cycles (94°C for 30 seconds, 56°C for 45 seconds and 72 °C for 2 minutes) followed by a final extension of 72°C for 7 minutes. Products were verified by electrophoresis of 5 µl on analytical agarose gels. PCR products were then cloned into pMAR101.x by homologous recombination and co-transformed into yeast. Competent EGY48 yeast cells (10 µ) were combined with 50 ng of *Eco*RI, *Xho*I digested vector, 100 ng PCR products and 100 µl EZ3 solution, vortex mixed, and incubated at 30°C. After at least 30 minutes, the entire reaction was plated on agar plates containing YNB – Trp+glucose and incubated at 30°C for 72 hours. Colonies were screened for inserts by PCR with the vector-specific primers pYesTrp Forward and Reverse (Invitrogen). Ninety-six different genes were cloned into the 96 unique pMAR101.x vectors and pooled for analysis against the bait clones. Using a similar method, baits for Y2H were cloned and co-transformed into either pMW101/RFY206/pSH1834T or pHybLex/L40. Bait clones were selected on YNB+Ura +His+glucose or YPD+Zeocin(300 µg/ml), respectively.

Preparation of yeast bait and prey cultures. For validation experiments, several colonies from the bait transformation plates were inoculated into 15 mL of selection media and grown 48 hours at 30°C. For high-throughput experiments, 96 different baits were arrayed in 96-well V-bottom microplates containing 200 µl of selection media and grown 48 hours at 30 °C. For both applications, yeast library plates (containing genes cloned into the 96 pMAR101.x vectors) were thawed and 200 µl of selection media (YNB–Trp +glucose with antibiotics) was added to each well. These plates were also incubated at 30°C for 48 hours.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 316

Yeast two-hybrid assay. Liquid mating of yeast was performed essentially as described in Buckholz et al. (1999). To validate this method, 5 µl from each well of the prey yeast library was transferred into a fresh 96-well V-bottom plate. Bait cultures were spun down and the cells resuspended in 45 mL YEP galactose + raffinose broth with antibiotics. A 25 µl aliquot of bait culture was added to the 5 µl of prey culture in each well. Following a 48 hour incubation at 30°C, 200 µl of minimal selective dropout media minus uracil, histidine, tryptophan and leucine, plus 2% galactose and 1% raffinose (SGR-UHWL) was added diluting the rich YPD media 1:10. After incubating an additional 48 hours, 5 µl of samples were transferred to a new microtiter plate and diluted 1:40 using SGR-UHWL to a final volume of 205 µl. The diluted matings were incubated for an additional 3-5 days at 30 °C. The mating mixture (25 µl) was transferred to 96-well assay plates and β-galactosidase assay was performed as described.

For high-throughput analysis 175 µl of each prey culture was transferred to a 50 ml conical centrifuge tube and spun down, the supernatant removed, and the cells resuspended in 45 mL YNB –Trp + glucose with antibiotics. A 5 µl aliquot of each bait culture was transferred into a fresh 96-well V-bottom plate and 25 µl of the pooled prey culture was added to each well.

Decoding assay. Positive-interactors detected via the β-gal assay served as template for PCR in which the chloramphenicol cassettes of the interacting prey (pMAR101.x) were amplified using biotinylated primers, Watson and nCrick, and standard PCR conditions. Products were hybridized against a pool of microspheres containing 250 beads per µl in 1.5 M NaCl. The pool was populated with 20 different types of microspheres; where each type was coupled to the complement of one of the ZipCode sequences used in pMAR101.x construction. During the hybridization, samples were denatured at 96 °C for 2 minutes and incubated at 45 °C for greater than one hour. Following hybridization, samples were washed in 1X SSC containing 0.2% Tween-20, resuspended in a 1.5M NaCl solution containing streptavidin-phycoerythrin and incubated for 20 minutes at room temperature. The

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 317

reactions were diluted with 60 µlof 1X SSC containing 0.2% Tween-20 prior to analysis on the LX100 (Luminex, Austin, TX).

Validating the vectors. We synthesized 96 "library" vectors, each vector identifiable by a unique pair of 2 of 20 possible 25-base ZipCodes. The ZipCodes we used were DNA sequences derived from a sequence of the *M. tuberculosis* genome. Chen et al. (2000); Iannone (2000). The particular ZipCode sequences were chosen to: 1) be absent from the known human genome sequence (by BLAST analysis), 2) have no discernable secondary structure, and 3) not hybridize to any of the other ZipCodes under the conditions (determined empirically) used for the genotyping and decoding analysis.

Twelve of the 20 ZipCodes ($ZipCodes_{1-12}$) have been placed 5' upstream of a Cm^r gene while the other 8 ($ZipCodes_{A-H}$) have been located at the 3' end of the Cm^r gene. Each 5' ZipCode was bracketed by a similar DNA sequence (5'-TGGGCGACTTCTCCAAAC-3' (SEQ ID NO:2), a translation of the amino acid sequence WATSON which we have labeled the "Watson" sequence); each 3' sequence was bracketed by a second DNA sequence (5'-CTTGCAGATTCGGCAGTT-3' (SEQ ID NO:3), a translation of the amino acid sequence NCRICK which we have labeled the "nCrick" sequence). PCR amplification was used to generate 96 different cassettes containing the Cm^r gene; each fragment with the following order: 5'-Watson-$ZipCode_{(1-12)}$-Cm^r-$ZipCode_{(A-H)}$-nCrick-3'. The 96 different cassettes were cloned into a unique *Swa* I site of the pMAR101 vector to synthesize the vectors pMAR101.1 ... pMAR101.96.

The set of vectors was purified and biotin-labeled Watson and nCrick primers were used to generate a PCR fragment. This PCR fragment was then hybridized against the set of 20 microspheres (Figure 3). Figure 4 illustrates the results from a subset of the analysis and demonstrates clear discrimination of the appropriate pair of the beads to the labeled fragment.

Identifying interactors using a non-random array. Prey clones were constructed

in groups of (for this example) 96 where each novel cDNA fragment of the group was cloned into a unique vector of the 96 library vectors. A Y2H analysis was performed using a bait protein against the multiplexed prey. Positive clones were isolated and the cassettes containing the Cm^r genes were amplified using biotin-dye-labeled 'Watson' and 'nCrick' primers. The PCR product was then used in a bead-based decoding assay and both of the ZipCodes on either side of the Cm^r gene were identified by hybridization to a set of 20 different beads. The data points were used to decode which member of the 96 vector set contained the interactor cDNA fragment which interacted with our bait protein. To further validate this result, we used vector specific primers to amplify the region containing the cloned cDNA fragment and submitted this fragment for DNA sequencing. Results were analyzed by BLAST and compared with the results of the bead-based assay, shown in Table 3.

The data show that this method can be used for high-throughput yeast two-hybrid analysis where the proteins of interest are encoded by known or predicted cDNA sequences. The method is easily automated and can be scaled to accommodate projects of varying complexity. In this analysis of 96 clones per well, 20 ZipCodes in two "dimensions," the dimensions containing 12 and 8 elements (i.e., a 96 well plate), respectively, were used. In practice, a reduced set of 13 elements in 6 dimensions (2x2x2x2x2x3) can be used to encode the 96 different beads. In fact, the complexity of any number of analyses per well can be reduced to the sum of the prime factors defining the maximum number of clones one desires to analyze. For example, a 10-fold increase in well complexity, i.e., 960 clones per well, (12x8x10) could be encoded by either 30 beads or by the sum of the prime factors making up 960 (2x2x2x2x2x3x2x5), i.e., with just 20 beads.

To access the complete set of human genes with just a set of the currently existing 100 bead set, it is reasonable to expect that a multiplex of 1000 clones per well could be decoded using a 10x12x8 matrix or array. A single master library plate could be used to represent close to the predicted number of genes.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 319

Table 3. Example of results from an assay of approximately 1500 bait clones against 350 prey clones.[1]

Sample[2]	Avg. β-gal Score	Bead Set Values [3]				Hit Identity	Sequence Verification[4]
		1	2	3	4	Plasmid	
B707.P677.F05	8603	79 C 2964.922	58 10 2155.084	71 11 67.3227	8 E 52.8565	C10 pMAR101.75	(+)
B707.P675.E05	2470	58 10 1811.079	77 B 703.7803	53 H 16.65197	8 E 14.02155	B10 pMAR101.74	(+)
B691.P679.G06	2224	56 9 1436.165	8 E 761.8906	34 2 1.283019	81 D -2.06527	E09 pMAR101.69	(+)
B674.P675.G06	6480	52 6 2102.083	53 H 1564.84	50 5 18.56156	33 G 16.50951	H06 pMAR101.48	(+)
B674P675.A11	4556	38 4 1008.487	75 A 351.1794	79 C 36.33744	53 H 34.78622	A04 pMAR101.25	(-)[5]
B678P679.D01	2784	50 5 2203.989	53 H 1553.246	73 12 36.82699	52 6 26.41147	D01 pMAR101.04	(+)
B680P679.H04	5444	36 3 1957.287	53 H 1955.586	33 G 333.8939	50 5 236.088	H03 pMAR101.24	(+)
B674P675.G05	12924	50 5 2120.277	33 G 1602.547	73 12 20.97098	79 C 16.38077	G05 pMAR101.39	(+)
B674P675.G06	6480	52 6 2019.826	53 H 1536.088	33 G 42.05269	50 5 26.03052	H06 pMAR101.47	(+)

[1] All samples have been assayed in duplicate. Samples in which both assays were successful have been included in the data.
[2] Sample names reflect ("bait" plate. "prey" pool."bait" well)
[3]Luminex bead values for 4 highest scoring beads (out of 20-bead set) after background subtraction.
[4]Sequence results of inserts cloned by homologous recombination into pMAR101.x were analyzed by BLAST.
[5]Top BLAST hit reveals a protein family member closely related to cloned gene.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 320

Example 2: Phage display-, F pilli- and lacI-fusion display systems.

Using the protocol set forth in Example 1, 96 or more differentially-labeled M13 clones are made. These are used in either gpIII or gpVIII vectors in an arrayed format as in the Y2H example. Alternatively, the ZipCodes are generated in a 'wild-card' (i.e., random) synthesis and then cloned into the vectors as pools of up to several thousands to millions. These libraries are used in a typical phage display experiment and the ZipCodes identified after panning.

LacI and F pilli are other display-systems and can be used in a manner identical to the phage display methods.

Example 3: Baculovirus (in insect and eukaryotic) adenovirus, adeno-associated virus (AAV), and retrovirus.

Similar to that described for phage display and yeast two-hybrid, these eukaryotic viruses are used to express foreign protein in eukaryotic cells. Baculovirus can be used to infect both insect and mammalian cells and also as a fusion vector (in a manner similar to the M13 gp-fusion vectors).

Example 4: Cell-based assay, cell-surface marker hapten or cell-surface protein.

The transcriptional readout used in the Y2H system is engineered such that the reporters are either small haptens or peptides that are transcribed and that eventually appear on the surface of the cell as *in vivo* fusions (for example, in yeast the Mat alpha gene product, in *E.coli* the *F pilli* gene product, and in mammalian cell lines the CD40 gene product). These cell lines are used in panning experiments against a set of antibodies or other specific-interactors. The interactors are fused to beads or labeled with a reporter molecule. In a 2-dimensional analysis of 20 different fusions, 96 (using 8 x 12) different types of cells are analyzed at once. In a manner similar to that described for yeast two-hybrid, the number of dimensions or factors within a dimension are increased to increase the number of cell lines that could be

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 321

simultaneously examined.

Example 5: Cell-based assay, other markers

In place of haptens as in the previous example, the reporters are external protein fusions or nucleic acid molecules (for example, DNA and/or RNA). The reporter molecules can be one of the many forms of green flourescent proteins (GFPs) available.

Example 6: Protein tags with antibodies

Proteins are synthesized with genetic fusion tags. In a specific example, a set of twenty tags is designed that do not cross-react. Sets of tags are used in a manner similar to the dimensions used in the yeast two-hybrid experiment described.

Throughout this application, various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention without departing from the scope or spirit of the invention. Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.

References

1. Buckholz, R.G., et al. (1999) Automation of yeast two-hybrid screening. J. Molec. Microbiol. Biotechnol. 1: 34-38.
2. Chen, J., et al. (2000) A Microsphere-Based Assay for Multiplexed Single

Nucleotide Plymorphism Analysis Using Single Base Chain Extension. Genome Research 10:549-557.

3. Finley, R.L., Brent, R. (1994) Interaction mating reveals binary and ternary connections between Drosophila cell cycle regulators. Proc. Natl. Acad. Sci. USA 91: 12980-12984.
4. Gyuris, J., et al. (1993) Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. Cell 75: 791-803.
5. Iannone, M.A., et al. (2000) Multiplexed Single Nucleotide Polymorphism Genotyping by Oligonucleotide Ligation and Flow Cytometry. Cytometry 39:131-140.
6. Oldenburg, K., et al. (1997) Recombination-mediated PCR-directed plasmid construction in vivo in yeast. Nucleic Acids Research 25:451-452.
7. Shoemaker, D.D., et al. (1996) Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. Nat. Genet. 14:450-456.
8. Watson, M.A., et al. (1996) Vectors encoding Alternative Antibiotic Resistance for Use in the Yeast Two-Hybrid System. BioTechniques 21:255-259 .

What is claimed is:

1. A method of encoding a complex mixture of assay constituents comprising using combinations of detectable tags and a total number of detectable tags that is less than the total number of constituents to be encoded, to encode the complex mixture of assay constituents.

2. The method of claim 1, further comprising:

   (a) determining the total number of constituents to be encoded;

   (b) determining the number of detectable tags in each combination, wherein the number of detectable tags in each combination is more than one and less than or equal to the number of prime numbers in the number of constituents to be encoded; and

   (c) determining the total number of detectable tags, wherein the total number of detectable tags equals a sum of a set of factors of the total number of constituents, wherein the number of factors equals the number of detectable tags in each combination.

3. The method of claim 2, wherein the total number of detectable tags is minimized by selecting factors of the total number of constituents that are equal or approximate.

4. The method of claim 1, wherein the constituents are selected from the group consisting of proteins, peptides, amino acids, small molecules, nucleotides, fatty acids, sugars, cofactors, receptors, receptor ligands, protein domains, oligonucleotides, transcription factors, nucleic acids, and small compounds.

5. The method of claim 1, wherein the detectable tags are directly or indirectly coupled to the constituents.

6. The method of claim 1, wherein the detectable tags are selected from the group consisting of radiolabels, dyes, fluorescent labels, and combinations thereof.

7. The method of claim 5, wherein the detectable tags are contained in or coupled to microspheres.

8. The method of claim 7, wherein the microspheres are coupled to a means of specifically binding the constituents.

9. The method of claim 8, wherein the coupled means is a nucleic acid complementary to a nucleic acid in or bound to the constituent to be bound.

10. A method of performing a multiplexed assay using complex mixtures of assay constituents encoded according to the method of claim 2, comprising:

    (a) Performing an assay to produce assay products using an array of the complex mixtures, wherein each constituent in a single complex mixture is detectably tagged with a unique combination of detectable tags;

    (b) Detecting which complex mixtures of assay products in the array have a positive response; and

    (c) Decoding the constituents in the complex mixtures having the positive response to determine which specific constituent or constituents are positive.

11. The method of claim 10, wherein the assay is selected from the group consisting of a chemical assay, protein assay, pharmacologic assay, antibody binding assay, hybrid assay, display assay, and genomic assay readout assay.

12. The method of claim 10, wherein the constituents are selected from the group

consisting of nucleic acids, proteins, peptides, amino acids, small molecules, nucleotides, fatty acids, sugars, cofactors, receptors, receptor ligands, protein domains, oligonucleotides, transcription factors, and small compounds.

13. The method of claim 10, wherein the detectable tags are directly or indirectly coupled to the constituents.

14. The method of claim 10, wherein the detectable tags are selected from the group consisting of radiolabels, dyes, fluorescent labels, and combinations thereof.

15. The method of claim 10, wherein the detectable tags are contained in or coupled to microspheres.

16. The method of claim 15, wherein the microspheres are coupled to a means of specifically binding the constituents.

17. The method of claim 16, wherein the coupled means is a nucleic acid complementary to a nucleic acid in or bound to the constituent to be bound to the microspheres.

18. The method of claim 10, wherein the decoding is performed by detecting and distinguishing with a detection device the detectable tags in each complex mixture of constituents.

19. A kit for performing a multiplexed assay using complex mixtures of encoded assay constituents, comprising a means of detectably tagging assay constituents encoded according to method of claim 2 and an arraying means for a plurality of complex mixtures.

BAIT AND PREY VECTORS

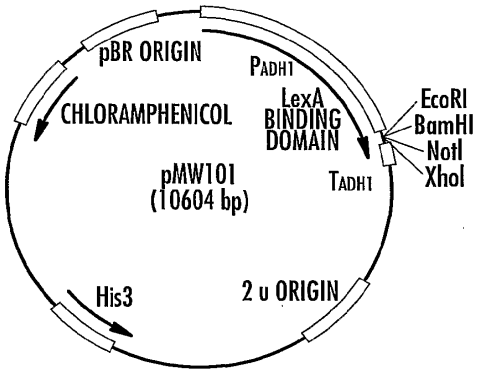
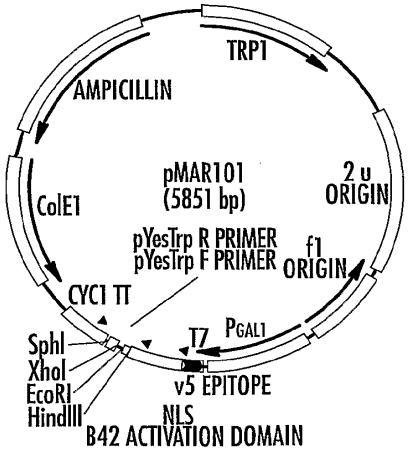


FIG. 1.



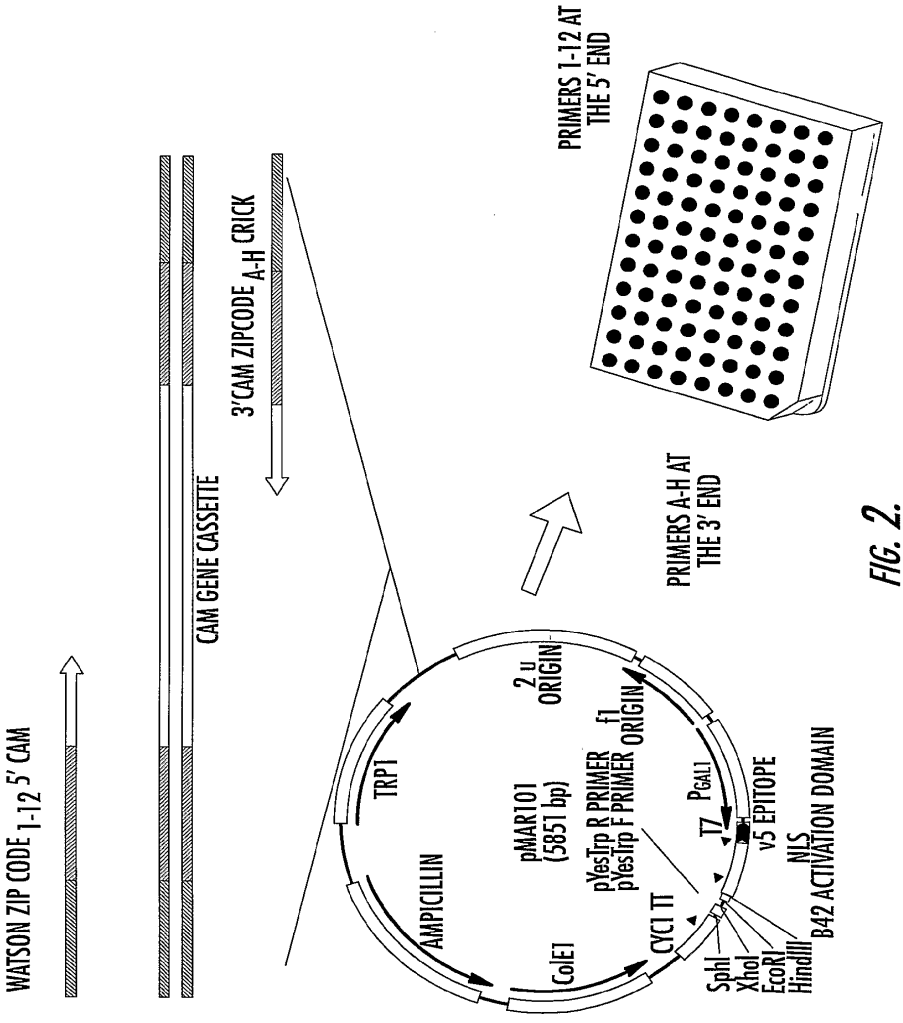


FIG. 2.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 328

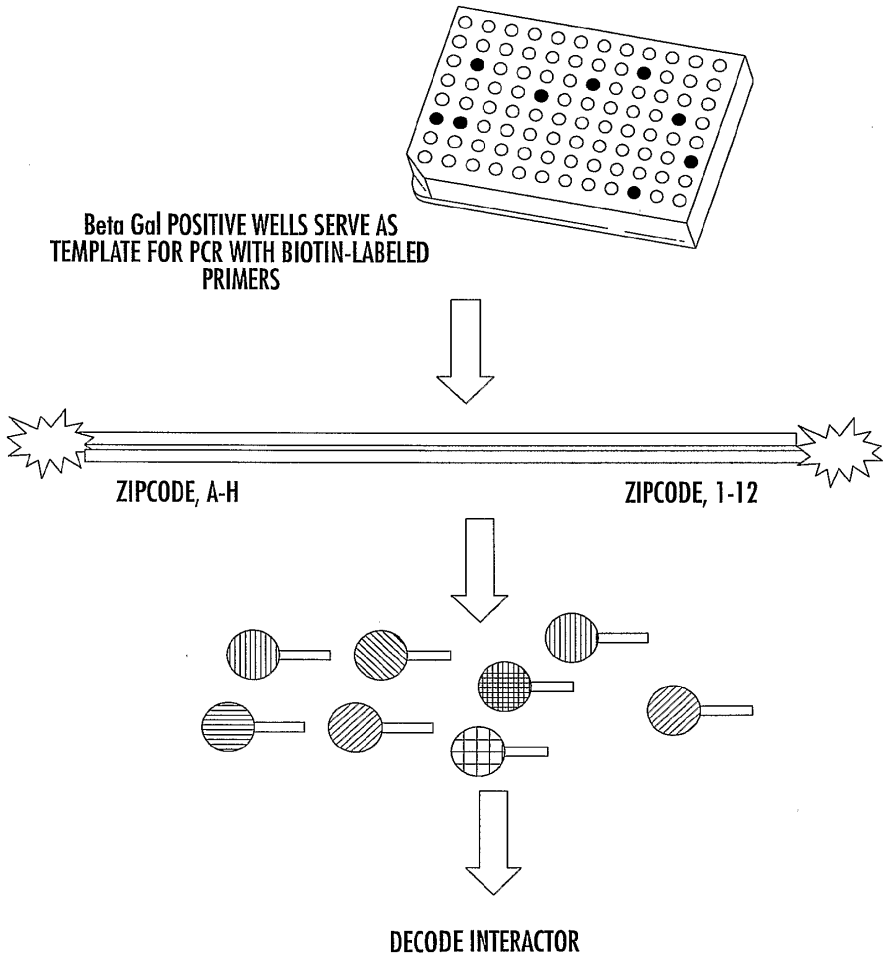

Beta Gal POSITIVE WELLS SERVE AS TEMPLATE FOR PCR WITH BIOTIN-LABELED PRIMERS
ZIPCODE, A-H
ZIPCODE, 1-12
DECODE INTERACTOR

FIG. 3.

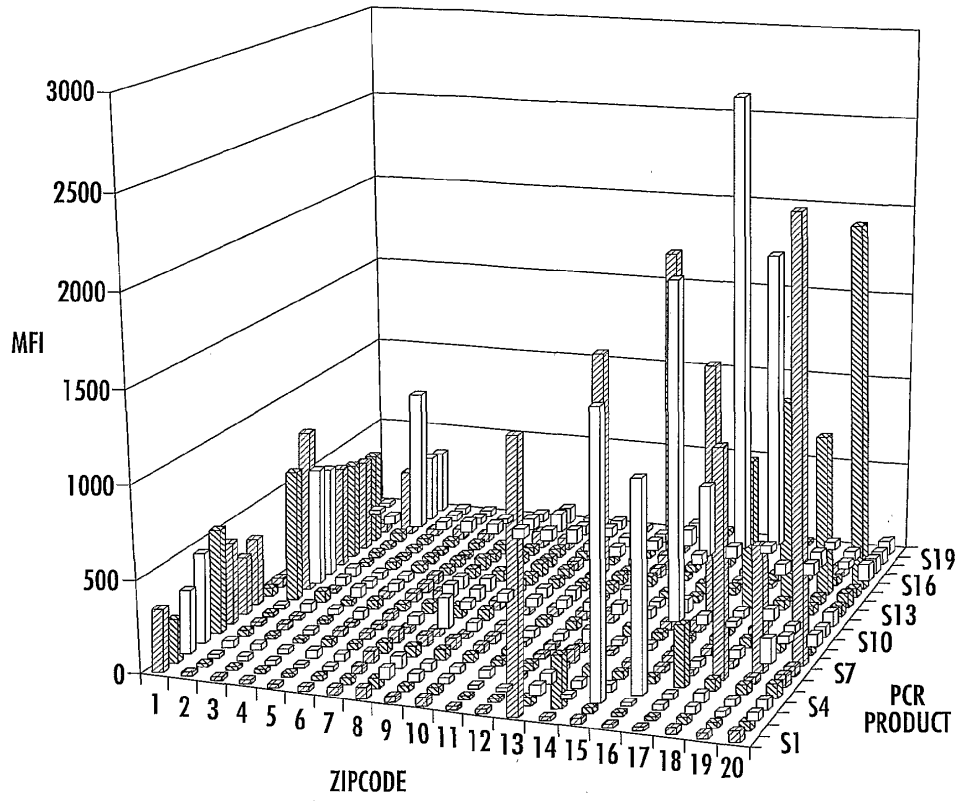

3000
2500
2000
MFI
1500
1000
500
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
ZIPCODE
S19
S16
S13
S10
S7
S4
S1
PCR
PRODUCT

FIG. 4.

SEQUENCE LISTING

<110> Glaxo Group Limited
Michael P. Weiner
Lottie L. Peppers
Anita J. Nelsen

<120> METHODS FOR ENCODING AND DECODING
COMPLEX MIXTURES IN ARRAYED ASSAYS

<130> 07083.0010P1

<160> 29

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 1
caggccaagt aacttcttcg 20

<210> 2
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 2
tgggcgactt ctccaaac 18

<210> 3
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 3
cttgcagatt cggcagtt 18

<210> 4
<211> 65
<212> DNA
<213> Artificial Sequence

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 331

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 4
tgggcgactt ctccaaacga tgatcgacga gacactctcg ccactgtgac ggaagatcac 60
ttcgc 65

<210> 5
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 5
tgggcgactt ctccaaaccg gtcgacgagc tgccgcgcaa gatctgtgac ggaagatcac 60
ttcgc 65

<210> 6
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 6
tgggcgactt ctccaaacga cattcgcgat cgccgcccgc tttctgtgac ggaagatcac 60
ttcgc 65

<210> 7
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 7
tgggcgactt ctccaaaccg gtatcgcgac cgcatcccaa tctctgtgac ggaagatcac 60
ttcgc 65

<210> 8
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 8
tgggcgactt ctccaaacgc tcgaagaggc gctacagatc ctcctgtgac ggaagatcac 60
ttcgc 65

<210> 9
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 9
tgggcgactt ctccaaacca ccgccagctc ggcttcgagt tcgctgtgac ggaagatcac 60
ttcgc 65

<210> 10
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 10
tgggcgactt ctccaaacgt aaatctccag cggaagggta cggctgtgac ggaagatcac 60
ttcgc 65

<210> 11
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 11
tgggcgactt ctccaaacct tttcccgtcc gtcatcgctc aagctgtgac ggaagatcac 60
ttcgc 65

<210> 12
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 12
tgggcgactt ctccaaacgg ctgggtctac agatccccaa cttctgtgac ggaagatcac 60
ttcgc 65

<210> 13
<211> 65
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 13
tgggcgactt ctccaaacga acctttcgct tcaccggccg atcctgtgac ggaagatcac 60
ttcgc 65

<210> 14
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 14
tgggcgactt ctccaaactt tcggcacgcg cgggatcacc atcctgtgac ggaagatcac 60
ttcgc 65

<210> 15
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 15
tgggcgactt ctccaaacct cggtggtgct gacggtgcaa tccctgtgac ggaagatcac 60
ttcgc 65

<210> 16
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 16
cttgcagatt cggcagtttc aacgtgccag cgccgtcctg ggactccacg gggagagcct 60
gagca 65

<210> 17
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 17
cttgcagatt cggcagttgc gaaggaactc gacgtggacg ccgctccacg gggagagcct 60
gagca 65

<210> 18
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 18
cttgcagatt cggcagttcg gggataccga tctcgggcgc acactccacg gggagagcct 60
gagca 65

<210> 19
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 19
cttgcagatt cggcagttgg agcttacgcc atcacgatgc gatctccacg gggagagcct 60
gagca 65

<210> 20
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 20
cttgcagatt cggcagttcg tggcggtgcg gagtttcccc gaactccacg gggagagcct 60
gagca 65

<210> 21
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 21
cttgcagatt cggcagttcg atccaacgca ctggccaaac ctactccacg gggagagcct 60
gagca 65

<210> 22
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =

synthetic construct

<400> 22
cttgcagatt cggcagttct gaatcctcca accgggttgt cgactccacg gggagagcct 60
gagca 65

<210> 23
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 23
cttgcagatt cggcagtttt cggcgctggc gtaaagcttt tggctccacg gggagagcct 60
gagca 65

<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 24
tccacgggga gagcctgagc a 21

<210> 25
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 25
ctgtgacgga agatcacttc gc 22

<210> 26
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 26
gcaacggcga ctggctggaa ttc 23

<210> 27
<211> 23
<212> DNA
<213> Artificial Sequence

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 336

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 27
gcttggctgc aggtcgactc gag 23

<210> 28
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 28
gccgaagctt gcggttgggg tattcgcaac ggcgactgg 39

<210> 29
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 29
atacgcatgc aattcgcccg gaattagctt ggctgcaggt 40

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 July 2002 (18.07.2002)

PCT

(10) International Publication Number
WO 02/056014 A3

(51) International Patent Classification[7]: G01N 33/53, C12Q 1/00, G01N 33/543

(21) International Application Number: PCT/US01/49132

(22) International Filing Date:
18 December 2001 (18.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/747,003 22 December 2000 (22.12.2000) US

(71) Applicant *(for all designated States except US)*: **GLAXO GROUP LIMITED** [GB/GB]; Glaxo Wellcome House, Berkeley Avenue, Greenford, Middlesex UB6 ONN (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **NELSEN, Anita, J.** [US/US]; GlaxoSmithKline, Five Moore Drive, P.O. Box 13398, Research Triangle Park, NC 27709 (US). **PEPPERS, Lottie, L.** [US/US]; GlaxoSmithKline, Five Moore Drive, P.O. Box 13398, Research Triangle Park, NC 27709 (US). **WEINER, Michael, Phillip** [US/US]; c/o 454 Corporation, 20 Commercial Street, Branford, CT 06405 (US).

(74) Agent: **GRASSLER, Frank, P.**; GlaxoSmithKline, Five Moore Drive, P.O. Box 1398, Research Triangle Park, NC 27709 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, 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, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, 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, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, 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*

(88) Date of publication of the international search report:
9 October 2003

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 02/056014 A3

(54) Title: METHODS FOR ENCODING AND DECODING COMPLEX MIXTURES IN ARRAYED ASSAYS

(57) Abstract: The present invention is a method of encoding a complex mixture of assay constituents comprising using combinations of detectable tags and a total number of detectable tags less than the total number of constituents to be encoded. The method comprises determining the total number of constituents to be encoded; determining the number of detectable tags in each combination, wherein the number of detectable tags in each combination is more than one and less than or equal to the number of prime numbers in the number of constituents to be encoded; and determining the total number of detectable tags, wherein the total number of detectable tags equals a sum of a set of factors of the total number of constituents, wherein the number of factors equals the number of detectable tags in each combination. The encoding methods are useful in a multiplexed assay using complex mixtures of assay constituents.

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 338

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/49132

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N33/53 C12Q1/00 G01N33/543

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 7 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)
EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FULTON R J ET AL: "Advanced multiplexd analysis with the FlowMetrix(TM) system" CLINICAL CHEMISTRY, AMERICAN ASSOCIATION FOR CLINICAL CHEMISTRY. WINSTON, US, vol. 43, no. 9, September 1997 (1997-09), pages 1749-1756, XP002142645 ISSN: 0009-9147	1,4-6
Y	the whole document	1-19
X	WO 99 52708 A (LUMINEX CORPORATION) 21 October 1999 (1999-10-21) page 17, line 27 -page 19; claims 1-17	1-19
A	WO 99 64867 A (AMERSHAM PHARMACIA BIOTECH UK LIMITED) 16 December 1999 (1999-12-16) the whole document	1-19

-/--

[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
22 July 2003	06/08/2003

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	Moreno de Vega, C

Form PCT/ISA/210 (second sheet) (July 1992)

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 339

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	CHEN J ET AL: "A MICROSPHERE-BASED ASSAY FOR MULTIPLEXED SINGLE NUCLEOTIDE POLYMORPHISM ANALYSIS USING SINGLE BASE CHAIN EXTENSION" GENOME RESEARCH, COLD SPRING HARBOR LABORATORY PRESS, US, vol. 10, no. 4, April 2000 (2000-04), pages 549-557, XP000927257 ISSN: 1088-9051 the whole document ---	1-19
Y	WO 00 68434 A (YALE UNIVERSITY) 16 November 2000 (2000-11-16) the whole document -----	1-19

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 340

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int ional Application No
PCT/US 01/49132

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9952708 A	21-10-1999	AU 3555599 A	01-11-1999
		EP 1079967 A1	07-03-2001
		WO 9952708 A1	21-10-1999
WO 9964867 A	16-12-1999	CA 2313047 A1	16-12-1999
		EP 1036332 A1	20-09-2000
		WO 9964867 A1	16-12-1999
		JP 2002517759 T	18-06-2002
WO 0068434 A	16-11-2000	AU 4824800 A	21-11-2000
		CA 2371843 A1	16-11-2000
		EP 1196632 A2	17-04-2002
		JP 2002543813 T	24-12-2002
		WO 0068434 A2	16-11-2000

Form PCT/ISA/210 (patent family annex) (July 1992)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 July 2007 (05.07.2007)

PCT

(10) International Publication Number
WO 2007/076128 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2006/049274

(22) International Filing Date:
22 December 2006 (22.12.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/753,758 23 December 2005 (23.12.2005) US
60/843,528 8 September 2006 (08.09.2006) US

(71) Applicants *(for all designated States except US)*: **NANOSTRING TECHNOLOGIES, INC.** [US/US]; 201 Elliot Avenue West, Suite 300, Seattle, WA 98119 (US). **INSTITUTE FOR SYSTEMS BIOLOGY** [US/US]; 1441 North 34th Street, Seattle, WA 98103-8904 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **GEISS, Gary, K.** [US/US]; 647 Nw 78th Street, Seattle, WA 98117 (US). **FERREE, Sean, M.** [US/US]; 7018 3rd Avenue Nw, Seattle, WA 98117 (US). **WEBSTER, Philippa, J.** [US/US]; 4317 Densmore Avenue N, Seattle, WA 98103 (US). **DIMITROV, Krassen, M.** [BG/US]; 8804 Ravenna Avenue NE, Seattle, WA 98115 (US).

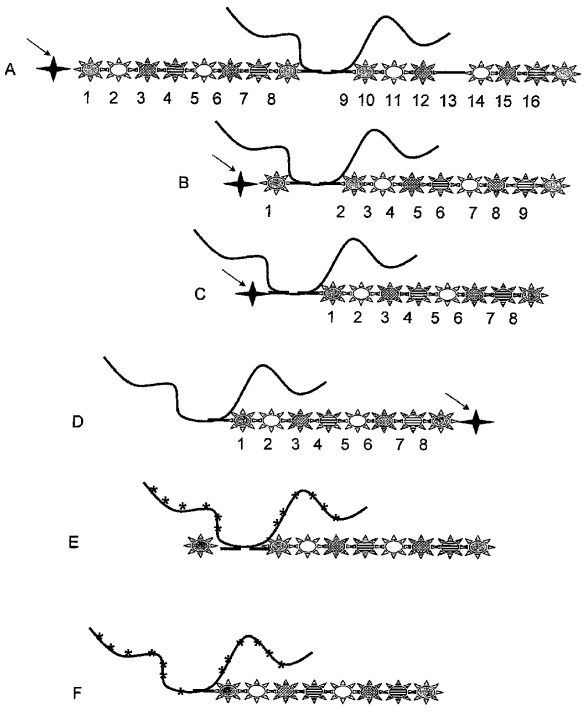
(74) Agents: **ANTLER, Adriane, M.** et al.; Jones Day, 222 East 41st Street, New York, NY 10017-6702 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, 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, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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,

[Continued on next page]

(54) Title: NANOREPORTERS AND METHODS OF MANUFACTURING AND USE THEREOF



(57) Abstract: The present invention relates to compositions and methods for detection and quantification of individual target molecules in biomolecular samples. In particular, the invention relates to coded, labeled probes that are capable of binding to and identifying target molecules based on the probes' label codes. Methods of making and using such probes are also provided. The probes can be used in diagnostic, prognostic, quality control and screening applications.

WO 2007/076128 A2

IPR2024-01177
California Institute of Technology - Ex. 2003 Page 342

GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, 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*

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.