Dkt. 5981/81503-A/GJG/ACK

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicants : James Hicks et al.

Serial No. : Not Yet Known

Filed : Herewith

For : VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

30 Rockefeller Plaza
20th Floor
New York, New York 10112
October 21, 2011

**BY EFS**
Commissioner for Patents
Alexandria, VA

**PETITION UNDER 37 CFR §1.84(a)(2) TO ACCEPT COLOR DRAWINGS FILED IN CONNECTION WITH THE ABOVE-IDENTIFIED APPLICATION**

This Petition is submitted under 37 CFR §1.84(a)(2) to accept color drawings filed with the above-identified application.

The subject application is being filed with one (1) set of color drawings consisting of eleven (11) sheets containing Figures 1-11. Such color drawings are necessary because certain features of the Figures are referenced by color in the specification and/or would not be identifiable or distinguishable in black and white.

For example, Figures 3, 4, 5, and 11 show data in a graphical form which cannot be adequately represented in black and white.

Another example is Figures 8-10 which use colors to differentiate strands of nucleic acid molecules and to identify their source.

In accordance with the guidance set forth in Section H of "LEGAL FRAMEWORK FOR EFS-WEB (06APRIL11)" published by the U.S. Patent and Trademark Office at www.uspto.gov/patents/process/file/efs/guidance/New_legal_framework.jsp, the requirement for three (3) sets of color drawings under 37 CFR §1.84(a)(2)(ii) is not applicable to color


IPR2025-01353; IPR2025-01355
CSHL EX2002

drawings submitted via EFS-Web. Therefore, applicants are submitting herewith via EFS-Web only one (1) set of such color drawings.

***Specification***

In accordance with 37 CFR §1.84(a)(2)(iii), the subject specification contains a paragraph immediately after the section heading titled "Brief Description Of The Figures" stating:

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

***Fee for Filing a Petition under 37 CFR §1.84(a)(2)***

Pursuant to 37 CFR §1.84(a)(2)(i), a fee of $130.00 is required for filing this Petition for accepting color drawings. Accordingly, authorization is hereby given to charge the amount of ONE HUNDRED THIRTY DOLLARS ($130.00) to Deposit Account No. 03-3125.

Applicants maintain that this Petition complies with the requirements of 37 CFR §1.84(a)(2), and respectfully request that this Petition be granted and the color drawings filed in connection with the subject application be accepted.


IPR2025-01353; IPR2025-01355
CSHL EX2002

No fee, other than the above-mentioned fee of $130.00, is deemed necessary in connection with the filing of this Petition. However, if any additional fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

Respectfully submitted,



Gary J. Gershik
Registration No. 39,992
Attorney for Applicants
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | |
| **Filing Date:** | |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER VARIATIONS |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | John P. White/Adam Krol |
| **Attorney Docket Number:** | 81503-A/GJG/ACK |

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 | 95 | 95 |
| Utility Search Fee | 2111 | 1 | 310 | 310 |
| Utility Examination Fee | 2311 | 1 | 125 | 125 |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| Petition fee- 37 CFR 1.17(h) (Group III) | 1464 | 1 | 130 | 130 |

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| **Extension-of-Time:** | | | | |
| **Miscellaneous:** | | | | |
| | | **Total in USD ($)** | | **660** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 11236578 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER VARIATIONS |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Adam Krol |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 81503-A/GJG/ACK |
| **Receipt Date:** | 21-OCT-2011 |
| **Filing Date:** | |
| **Time Stamp:** | 11:33:26 |
| **Application Type:** | Utility under 35 USC 111(a) |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | Deposit Account |
| Payment was successfully received in RAM | $660 |
| RAM confirmation Number | 7988 |
| Deposit Account | 033125 |
| Authorized User | |

## File Listing:

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



IPR2025-01353, IPR2025-01355
CSHL EX2002

| | | | | | |
|---|---|---|---|---|---|
| 1 | Transmittal of New Application | 111021_5981_81503_A_Transmittal_ACK.pdf | 195716<br>d9269c6d27a532408c6596235bbefc5a694666c2 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | | 111021_5981_81503_A_Application_ACK.pdf | 201755<br>55622c2510df08bf34b7dcd13c2de457a19b4225 | yes | 87 |

| | Multipart Description/PDF files in .zip description | | |
|---|---|---|---|
| | **Document Description** | **Start** | **End** |
| | Specification | 1 | 69 |
| | Claims | 70 | 86 |
| | Abstract | 87 | 87 |

| | | | | | |
|---|---|---|---|---|---|
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Drawings-other than black and white line drawings | 111021_5981_81503_A_Drawings_ACK.pdf | 1181727<br>c8287448249aedc8d408a11bdcec906fcab87767 | no | 11 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 4 | Oath or Declaration filed | 111021_5981_81503_A_DPOA_ACK.pdf | 269299<br>18f2f51a875555220aa8120d791bed6c71e7e6b4 | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 5 | | 111021_5981_81503_A_PrelimAmend_IDS_ACK.pdf | 10573042<br>89092248d6cea3779f922684997ee3bc7931c454 | yes | 92 |

| | Multipart Description/PDF files in .zip description | | |
|---|---|---|---|
| | **Document Description** | **Start** | **End** |
| | Preliminary Amendment | 1 | 1 |
| | Claims | 2 | 9 |
| | Applicant Arguments/Remarks Made in an Amendment | 10 | 10 |
| | Transmittal Letter | 11 | 13 |
| | Information Disclosure Statement (IDS) Form (SB08) | 14 | 15 |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| | | | | | |
|---|---|---|---|---|---|
| | Non Patent Literature | | 16 | | 23 |
| | Non Patent Literature | | 24 | | 27 |
| | Non Patent Literature | | 28 | | 33 |
| | Non Patent Literature | | 34 | | 54 |
| | Non Patent Literature | | 55 | | 60 |
| | Non Patent Literature | | 61 | | 64 |
| | Non Patent Literature | | 65 | | 71 |
| | Non Patent Literature | | 72 | | 80 |
| | Non Patent Literature | | 81 | | 92 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 6 | Petition for review by the Office of Petitions. | 111021_5981_81503_A_Petition_ACK.pdf | 163253<br>bd41bed669d04b949059b13973b968557b991e3d | no | 3 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 7 | Fee Worksheet (SB06) | fee-info.pdf | 37109<br>6862de6c5af917367d5a0b76c6d5662f376b3b6c | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 12621901 | | |

**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.**

Docket No. **5981/81503-A/GJG/ACK**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, VA 22313-1450

Date: **October 21, 2011**

Sir:

Transmitted herewith for filing is the utility patent application of:

**James Hicks et al.**
Inventor(s)

for

**VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION**
Title of Invention

including the following:

X Application ( **98** pages in total), including **69** pages of specification; **17** page(s) of claims; **1** page(s) of abstract; **0** page(s) of sequence listing; and **11** sheet(s) of drawings

**X** Oath or Declaration of Applicant(s) (_____signed **X** unsigned)

**X** Power of Attorney (_____signed **X** unsigned)

**X** Preliminary Amendment ( ___ including claim to benefit of earlier U.S. Provisional Application(s))

The following are also enclosed:

____ Assignment to _____

____ Verified Statement to establish small entity status under 37 C.F.R. §1.9 and §1.27

**X** Information Disclosure Statement, including Form PTO-1449 (Copies of citations are included: Yes **X** ; No_____)

____ Non-Publication Request (Form PTO/SB/135 must be attached)

____ Computer Readable Form (CRF) of Sequence Listing and Statement Verifying Identity of CRF and Sequence Listing

____ Certified copy of previously filed foreign application(s) as follows:

| Country | Application No. | Filing Date |
|---|---|---|
| | | |
| | | |
| | | |

____ Applicant(s) hereby claim(s) priority based upon the aforementioned foreign application(s) under 35 U.S.C. §119

___ Two copies of this Patent Application Transmittal Letter

___ Return Receipt Postcard

___ Express Mail Certificate of Mailing Label No._____

dated _____

___ Other (identify): _____

_____

_____

The filing fee is calculated as follows:

NUMBER OF CLAIMS AS FILED, MINUS ALL CLAIMS CANCELED BY ANY ACCOMPANYING PRELIMINARY AMENDMENT

| | Number Filed | | Number Extra* | | Rate Small Entity | Rate Other Entity | | Fee Small Entity | Fee Other Entity |
|---|---|---|---|---|---|---|---|---|---|
| Total Claims | **20** | -20 = | **0** | X | $26 | $52 | = | $ **0** | $ |
| Independent Claims | **3** | -3 = | **0** | X | $110 | $220 | = | $ **0** | $ |
| Pages in Excess of 100 | **98** | -100= | **0** | ÷ by 50 † X | $135 | $270 | = | $ **0** | $ |
| Multiple Dependent Claims Present: __ Yes **X** No | | | | | $195 | $390 | = | $ **0** | $ |
| Non-English Specification: __ Yes **X** No | | | | | $130. | $130. | = | $ **0** | $ |

* If less than zero, enter "0"
† Round upwards to integer, e.g. 1.1 = 2, and insert

| | Small Entity | Other Entity |
|---|---|---|
| Basic Fee | $ 165. | $ 330. |
| Search Fee | $ 270. | $ 540. |
| Examination Fee | $ 110. | $ 220. |
| **Total Fee** | $**545.00** | $ |

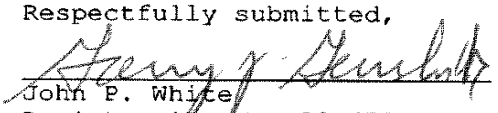
___ A check in the amount of $_____ is enclosed.

**X** Please charge Deposit Account No. **03-3125** in the amount of $ **545.00**

___ The Commissioner is hereby authorized to charge any additional fees required or credit any overpayment to Deposit Account No. 03-3125 as follows:

**X** Filing fees under 37 C.F.R. §1.16
**X** Patent application processing fees under 37 C.F.R. §1.17

Respectfully submitted,



John P. White
Registration No. 28,678
Gary J. Gershik
Registration No. 39,992
Attorneys for Applicant(s)
Cooper & Dunham LLP (Customer #23432)
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400


IPR2025-01353; IPR2025-01355
CSHL EX2002

# VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

This application claims the benefit of U.S. Provisional Application Nos. 61/510,579, filed July 22, 2011, and 61/406,067, filed October 22, 2010 the contents of each of which are hereby incorporated by reference in their entirety.

The invention disclosed herein was supported by the government under grant No. W81XWH-09-1-0591 from the Department of Defense. Accordingly, the U. S. Government has certain rights to this invention.

Throughout this application, various publications are referenced by numbers in parentheses. Full citations for these references may be found at the end of the specification immediately preceding the claims. The disclosures of these publications in their entireties are hereby incorporated by reference into this application to more fully describe the state of the art to which this invention pertains.

## Background of the Invention

Genomic copy number information is commonly obtained using whole genome amplification (WGA). The endemic problem with the WGA method is over-sampling of certain regions, yielding a non-uniform amplification of the genome (1). WGA methods begin with the step that initiates the process, a polymerase (Phi 29) makes a strand from genomic DNA utilizing a random primer coupled to an adaptor for subsequent PCR (Figure 1). If the input DNA strands are referred to as the "0-th derivative", and the first synthesized strand as a "first derivative," subsequent strands are called the (n+1)-th derivative if their template was an n-th derivative. Only strands that are 2-nd derivative or higher become amplified in the PCR step,

resulting in a 'stacking' over the regions 'chosen' by the polymerase for the first derivative.

Coverage of the genome by sequencing WGA of single cell DNA is limited by the stacking phenomenon (Figure 2). Thus it is difficult to obtain single cell measurements, particularly when based on WGA, due to distortions that originate from stochastic sampling and amplification steps. Moreover, the current method of WGA is a black box, with the unspecified reagents purchased from a vendor, which hampers optimization. Moreover, the WGA method does not extend to a method usable for single cell RNA profiling.

Ligation-mediated PCR was developed in an attempt to solve the above-identified problems inherent in WGA. In this method, adaptors are ligated to an *Mse*I restriction endonuclease digest of genomic DNA from a single cell, followed by PCR amplification using primers complementary to the adaptors. The amplified DNA is then used for CGH or DNA sequencing (2,3). However, like WGA, the method still requires an amplification step.

Parameswaran et al. (2007) and U.S. Patent No. 7,622,281 describe methods of labeling nucleic acid molecules with barcodes for the purpose of identifying the source of the nucleic acid molecules, thereby allowing for high-throughput sequencing of multiple samples (4,5). Eid et al. (2009) describe a single molecule sequencing method wherein single-molecule real time sequencing data is obtained from a DNA polymerase performing uninterrupted template-directed synthesis using four distinguishable fluorescently labeled dNTPs (6). However, these methods do not provide genomic information unaffected by amplification distortion.

Miner et al. (2004) describe a method of molecular barcoding to label template DNA prior to PCR amplification, and report that

the method allows for the identification of contaminant and redundant sequences by counting only distinctly tagged sequences (22). U.S. Patent No. 7,537,897 describes methods for molecular counting by labeling molecules of an input sample with unique oligonucleotide tags and subsequently amplifying and counting the number of different tags (23). Miner et al. and U.S. Patent No. 7,537,897 both describe labeling of input nucleic acid molecules by ligation, which has been found to be an inefficient reaction.

McCloskey et al. (2007) describe a method of molecular encoding which does not use ligation but instead uses template specific primers to barcode template DNA molecules prior to PCR amplification (24). However, such a method requires that template specific primers be made for each species of template DNA molecule studied.

As described herein, obtaining accurate genomic copy number information by high-throughput sequencing of genomic DNA prepared by WGA methods is hampered by the copy number distortions introduced by non-uniform amplification of genomic DNA. Thus, there exists a need for a method that allows for copy number determination free of distortions caused by amplification steps and which allows for accurate and efficient copy number determination of complex samples. Such a method should also be robust using existing methodologies for high volume, massively parallel sequencing.

**SUMMARY OF THE INVENTION**

A method is provided for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA;

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b), and

ii) a tag;

d) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

e) generating tag associated sequence reads by sequencing the product of step (d);

f) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a cDNA or a segment of cDNA to a location on the cDNA library; and

g) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from genomic material DNA methylation information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands to a polymerase reaction in the presence essentially unique primers, wherein the essentially unique primers are capable of hybridizing to the polynucleotide tail of the zero-th order strands, under conditions that promote the formation of only one complement, thereby generating tagged nucleic acid molecules;

d) separating the tagged nucleic acid molecules into a group consisting of hemi-methylated

tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules;

e) subjecting each group of step (d) to amplification by polymerase chain reaction (PCR);

f) generating tag associated sequence reads by sequencing the product of step (e);

g) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

h) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining DNA methylation information unaffected by amplification distortion.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands of step (b) to a polymerase reaction in the presence of primers capable of

hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

d) adding a polynucleotide tail to the first order derivative strands;

e) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (c) and (e) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of genomic material;

b) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

c) subjecting the zero-th order derivative strands of step (b) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (b) under

conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

d) subjecting the product of step (c) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (b) under conditions which promote the formation of only one complement, wherein the primers of step (d) have different nucleotide sequences than the primers of step (c), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (c) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from a mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA; and

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, wherein each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b) and

ii) a tag.

Also provided is a composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining mRNA transcripts;

b) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (c) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a kit for determining nucleic acid copy number information unaffected by amplification distortion comprising:

a) a terminal transferase, and

b) plurality of substantially unique primers,

wherein the substantially unique primers comprise substantially unique tags, and wherein the substantially unique primers are capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the terminal transferase.

## BRIEF DESCRIPTION OF THE FIGURES

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

**Figure 1. Whole Genome Amplification (WGA).**

**Figure 2. Genome coverage comparison.**

Genomic coverage comparison of 7 sequencing lanes from one single cell library versus 7 lanes from a million cell library. Stacking causes the single cell reads to be concentrated over the first derivative reads from WGA. More reads yield diminishing returns.

**Figure 3. Sequencing read counts for CpG dinucleotides within selected CpG islands**

Normal cell line SKN-1 (A,C,E) and breast cancer line MDA-MB-231 (B,D,F). Orange indicates T residues (unmethylated). Blue indicates C residues (methylated). Counts of the plus strand are above the X-axis, and counts on the minus strand are below.

**Figure 4. Comparison of array CGH with 'sequence counting' for analyzing copy number data.**

Graph shows a portion of chromosome 17 from FFPE sample JZ33 including raw (gray) ratio and segmented (green) data from array CGH, and raw sequence count data (orange) and its segmentation (blue).

**Figure 5. ROMA CGH of sKBR3 compared with profiles of 3 independent single cell isolates derived from sparse DNA sequencing.**

**Figure 6. Sequencing across the breakpoints of a homozygous deletion in cell on chromosome 5 line T47D.**

**Figure 7. Flowchart for genomic profiling by representational sequence counting**

**Figure 8. Scheme for varietal counting.** Varietal tags are shown as random nucleotide sequences ($N_{10}$)

Flowchart of one embodiment of the invention.

**Figure 9. Nick ligation protocol**

The sequence of primer A includes universal primer A (PrA; green), varietal tag (B; yellow), and oligo dT+CATG, the sequence of primer B includes universal primer B (PrB; orange) and oligo dT+CATG, PrA and PrB are universal primers, and cPrA(blue), cPrB (magenta), and cB(blue) are complementary to PrA, PrB and B, respectively.

**Figure 10. Schema for targeted sequencing of regional markers**

UfBPrA is forward primer, B (yellow) is tag, UrPrR is reverse primer, PrB (brown) is gene-specific reverse primer, Uf (orange) is universal forward primer, PrA (blue) is gene specific forward primer, Ur (magenta) is universal reverse primer.

**Figure 11. Comparison of Varietal Counting with Sequence Counting**

Figure 11A shows genome wide copy number data. Figure 11B shows partial genome copy number data.

**DETAILED DESCRIPTION OF THE INVENTION**

A method is provided for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

In an embodiment, the method further comprises estimating a genomic copy number of a region of the genome comprising more than one location on the genome by assigning as the copy

number of the region the highest count obtained in step (f) for the locations within the region.

In an embodiment, the method further comprises comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

In an embodiment, the method further comprises

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

Step (b) of the above embodiment may further comprise

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the

number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

In an embodiment of the method, the second region of the genome comprises a centromere.

In an embodiment, the method further comprises summing the counts obtained in step (f) for locations which comprise a region of the genome, and comparing the sum to a sum obtained from a reference sample for the same region of the genome, thereby estimating a relative genomic copy number of the region of the genome.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA;

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b), and

ii) a tag;

d) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

e) generating tag associated sequence reads by sequencing the product of step (d);

f) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a cDNA or a segment of cDNA to a location on the cDNA library; and

g) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

a) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

b) generating tag associated sequence reads by sequencing the product of step (b);

c) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

d) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from genomic material DNA methylation information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands to a polymerase reaction in the presence essentially unique primers, wherein the essentially unique primers are capable of hybridizing to the polynucleotide tail of the zero-th order strands, under conditions that promote the formation of only one complement, thereby generating tagged nucleic acid molecules;

d) separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules;

e) subjecting each group of step (d) to amplification by polymerase chain reaction (PCR);

f) generating tag associated sequence reads by sequencing the product of step (e);

g) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding

to a segment of the genomic material to a location on the genome; and

h) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining DNA methylation information unaffected by amplification distortion.

In an embodiment of the methods, tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

In an embodiment of the methods, tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

In an embodiment of the methods, adding a polynucleotide tail comprises the use of a terminal transferase.

In an embodiment of the methods, tagging the segments to generate tagged nucleic acid molecules comprises ligation of adaptors comprising the tags to at least one end of the segments of the genomic material.

In an embodiment of the methods, the adaptors comprising the tags are ligated to only one end of the segments of the genomic material.

In an embodiment of the methods, the tags comprise a sequence that aids PCR amplification.

In an embodiment of the methods, each tagged nucleic acid molecule comprises one tag.

In an embodiment of the methods, each tagged nucleic acid molecule comprises more than one tag.

In an embodiment of the methods, segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

In an embodiment of the methods, segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

In an embodiment of the methods, the maximum copy number of a location on a cDNA library is not less than the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library.

An embodiment of the above methods further comprises analyzing mRNA copy number.

In an embodiment of the methods, separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules is by cleavage with

methylation sensitive restriction enzymes, partitioning with antibodies, or partitioning with methyl-C binding proteins directed to methylated or hydroxymethylated cytosine.

In an embodiment of the methods, the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or prior to sequencing.

In an embodiment of the methods, each tagged nucleic acid molecule differs at more than one nucleotide.

In an embodiment of the methods, the tag sequences further comprise a sample tag.

In an embodiment of the methods, the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to sequencing.

An embodiment of the methods further comprises deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to sample tag.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from a single species.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from a single organism.

In an embodiment of the methods, tagged nucleic acid molecules are generated from a single cell.

In an embodiment, the single cell is from a needle aspirate of suspected cancer lesions.

In an embodiment, the single cell is from a core biopsy of suspected cancer lesions.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from two or more organisms.

In an embodiment, the tagged nucleic acid molecules are generated from a single cell of each organism.

In an embodiment of the methods, the single species is human.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from two or more species.

In an embodiment, the tagged nucleic acid molecules are generated from a population of microbes.

A further embodiment comprises comparing the genomic copy number information obtained for different species of the population to determine the relative count of those different species in the population.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

d) adding a polynucleotide tail to the first order derivative strands;

e) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (c) and (e) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of genomic material;

b) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

c) subjecting the zero-th order derivative strands of step (b) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (b) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

d) subjecting the product of step (c) to a polymerase reaction in the presence of primers

capable of hybridizing to the polynucleotide tail added in step (b) under conditions which promote the formation of only one complement, wherein the primers of step (d) have different nucleotide sequences than the primers of step (c), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (c) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from a mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA; and

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, wherein each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b) and

ii) a tag.

Also provided is a composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining mRNA transcripts;

b) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (c) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

In an embodiment of the above compositions, adding a polynucleotide tail comprises the use of a terminal transferase.

In an embodiment of the above compositions, each tagged nucleic acid molecule comprises one tag.

In an embodiment of the above compositions, each tagged nucleic acid molecule comprises more than one tag.

In an embodiment of the above compositions, segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

In an embodiment of the above compositions, segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

In an embodiment of the above compositions, the complexity of the tagged nucleic acid molecules is reduced by hybrid capture.

In an embodiment of the above compositions, each tagged nucleic acid molecules differs at more than one nucleotide.

In an embodiment of the above compositions, the tag sequences further comprise a sample tag.

Also provided is a composition of matter comprising at least two pools of tagged nucleic acid molecules as described herein.

In an embodiment of the above compositions, the tagged nucleic acid molecules are generated from a single cell.

In an embodiment, the single cell is from a needle aspirate of suspected cancer lesions.

In an embodiment, the single cell is from a core biopsy of suspected cancer legions.

In an embodiment of the above compositions, the tagged nucleic acid molecules are generated from two or more organisms.

Also provided is a kit for determining nucleic acid copy number information unaffected by amplification distortion comprising:

a) a terminal transferase, and

b) plurality of substantially unique primers,

wherein the substantially unique primers comprise substantially unique tags, and wherein the substantially unique primers are capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the terminal transferase.

In an embodiment, the kit further comprises a DNA polymerase having 3'-5' proofreading activity.

In an embodiment of the kits, the plurality of substantially unique primers comprises $10^n$ primers, wherein n is an integer from 2 to 9.

In an embodiment, the kits further comprise a second terminal transferase and a primer capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the second terminal transferase.

In an embodiment of the kits, the tags are six nucleotides long.

In an embodiment of the kits, the tags are 15 nucleotides long.

In an embodiment, the kits further comprise sample tags.

In an embodiment, the sample tags are 2 nucleotides long.

In an embodiment, the sample tags are 4 nucleotides long.

In an embodiment, the sample tags comprise a sample tag set.

In an embodiment of the above kits, the substantially unique tags comprise the sample tags.

**Terms**

For the purpose of this invention, different words and phrases are defined as follows:

As used herein, the term "adaptor" refers to an oligonucleotide or nucleic acid fragment or segment that can be ligated to a nucleic acid molecule of interest. For the purposes of this invention adaptors may, as options, comprise primer binding sites, recognition sites for endonucleases, common sequences, promoters, tag sequences, and sample tag sequences. Preferably, adaptors are positioned to be located on both sides (flanking) a particular nucleic acid molecule of interest. In accordance with the invention, adaptors may be added to nucleic acid molecules of interest by standard recombinant techniques (e.g. restriction digest and ligation). For example, adaptors may be added to a population of linear molecules, (e.g. a genomic DNA which has been cleaved or digested) to form a population of linear molecules containing adaptors at one and preferably both termini of all or a substantial portion of molecules. The adaptor may be entirely or substantially double stranded or entirely single stranded. A double stranded adaptor may comprise two oligonucleotides that are at least partially complementary. The adaptor may be phosphorylated or unphosphorylated on one or both strands. Adaptors may be used for DNA sequencing. Adaptors may also incorporate modified nucleotides that modify the properties of the adaptor sequence. For example, methylated cytosines may be substituted for cytosines. In an embodiment of this invention the adaptors ligated to genomic DNA to enable cluster generation on the sequencer contain cytosines which were all methylated. This modification protects such adaptors from bisulfite conversion, and is taken into account in the downstream applications and analysis of this invention.

As used herein, the term "amplifying" refers to the process of synthesizing nucleic acid molecules that are complementary to one or both strands of a template nucleic acid. Amplifying a nucleic acid molecule typically includes denaturing the

template nucleic acid, annealing primers to the template nucleic acid at a temperature that is below the melting temperatures of the primers, and enzymatically elongating from the primers to generate an amplification product. The denaturing, annealing and elongating steps each can be performed once. Generally, however, the denaturing, annealing and elongating steps are performed multiple times (e.g., polymerase chain reaction (PCR)) such that the amount of amplification product is increasing, often times exponentially, although exponential amplification is not required by the present methods. Amplification typically requires the presence of deoxyribonucleoside triphosphates, a DNA polymerase enzyme and an appropriate buffer and/or co-factors for optimal activity of the polymerase enzyme. The term "amplification product" refers to the nucleic acid sequences, which are produced from the amplifying process as defined herein.

As used herein, the term "amplification distortion" refers to non-uniform amplification of template nucleic acid molecules.

As used herein, the term "bisulfite-treatment" refers to the treatment of nucleic acid with a reagent used for the bisulfite conversion of cytosine to uracil. Examples of bisulfite conversion reagents include but are not limited to treatment with a bisulfite, a disulfite or a hydrogensulfite compound.

As used herein, the term "bisulfite-converted material" refers to a nucleic acid that has been contacted with bisulfite ion in an amount appropriate for bisulfite conversion protocols known in the art. Thus, the term "bisulfite-converted material" includes nucleic acids that have been contacted with, for example, magnesium bisulfite or sodium bisulfite, prior to treatment with base.

As used herein, the term "sequence capture" or "hybrid capture" refers to a process of hybridizing a "capture probe" to a nucleic acid having a sequence which is complementary to the sequence of the capture probe. A capture probe may be immobilized to a substrate on a solid phase microarray, wherein "substrate" refers to short nucleic acid sequences which are known. Their location on the solid phase microarray may or may not be predetermined. The capture probe comprising a "sequence complementary to the substrate" may be immobilized to the solid phase microarray by hybridizing to its complementary "substrate sequence."

As used herein, the term "clamp" refers to a nucleotide sequence of CC, CG, GC or GG which may be located at the first two nucleotide positions at the 5' end of a primer, tag, or sample tag, or the last two nucleotide positions at the 3' end of a primer, tag, or sample tag.

As used herein, the term "comparable number of locations" means that a first genomic region has a number of locations on the genome which is within 25% of the number of locations of a second genomic region. In an embodiment, the first genomic region which has a comparable number of locations to the second genomic region has a number of locations which is within 20%, 15%, 10%, 5%, 1%, or exactly the same as the number of locations of the second genomic location.

As used herein, a subsequence of a tag associated sequence read is "corresponding to" a tag when the subsequence is identical to the nucleotide sequence of the tag.

As used herein, a subsequence of a tag associated sequence read is "corresponding to" a species of a nucleic acid molecule when the subsequence is substantially identical to or substantially complementary to at least about 10, 12, 14, 16, 18, 20, or more nucleotides of the sequence of the species of

nucleic acid molecule. In an embodiment, the subsequence is identical to or fully complementary to at least about 10, 12, 14, 16, 18, 20, or more nucleotides of the sequence of the species of nucleic acid molecule.

The term "CpG site" refers to the CpG dinucleotide.

As used here, the term "CpG island" refers to a region of DNA with a high G+C content and a high frequency of CpG dinucleotides relative to the bulk genome, as further defined at the UCSC Genome Bioinformatics Site at genome.ucsc.edu/index.html?org=Human&db=hg19&hgsid=171216665, and in particular, by annotation in the USCS Genome Bioinformatics Site database (CpG Islands Track or Table: cpgIslandExt).

As used herein, the term "fully complementary" refers to the reverse complement of a nucleic acid sequence.

As used herein, the term "library" refers to a collection of nucleic acid molecules (circular or linear). In one preferred embodiment, a library is representative of all of the DNA content of an organism (such a library is referred to as a "genomic" library), or a set of nucleic acid molecules representative of all of the expressed genes (such a library is referred to as a cDNA library) in a cell, tissue, organ or organism. A library may also comprise random sequences made by de novo synthesis, mutagenesis of one or more sequences and the like. A library may be contained in one vector.

As used herein, the term "mapping" refers to identifying a location on a genome or cDNA library that has a sequence which is substantially identical to or substantially fully complementary to the subsequence of a tag associated sequence read corresponding to a species of nucleic acid molecule, and assigning the tag associated sequence read or the tagged nucleic acid molecule generating the tag associated sequence

read to the location. The nucleic acid molecule may be, but is not limited to the following: a segment of genomic material, a cDNA, a mRNA, or a segment of a cDNA.

As used herein, the term "methylation" refers to the covalent attachment of a methyl group at the C5-position of the nucleotide base cytosine within the CpG dinucleotides of genomic region of interest. The term "methylation state" or refers to the presence or absence of 5-methyl-cytosine ("5-Me") at one or a plurality of CpG dinucleotides within a DNA sequence. A methylation site is a sequence of contiguous linked nucleotides that is recognized and methylated by a sequence specific methylase. A methylase is an enzyme that methylates (i. e., covalently attaches a methyl group) one or more nucleotides at a methylation site.

As used herein, the term "mRNA transcript" refers to the nucleic acid molecule resulting from transcription of DNA.

As used herein, the term "pool" means to combine one plurality of nucleic acid molecules with at least one other plurality of nucleic acid molecules. In an embodiment, tagged nucleic acid molecules are pooled. Pooling may be done after any step in which tagged nucleic acid molecules are generated. In an embodiment, pooling is done prior to PCR and/or prior to sequencing. In an embodiment, tagged nucleic acid molecules generated from a single cell are pooled with tagged nucleic acid molecules generated from a different single cell.

As used herein, the term "probe array" refers to an array of N different nucleic acid molecules deposited on a reaction substrate which serve to interrogate mixtures of target molecules or multiple sites on a single target molecule administered to the surface of the array.

As used herein, the term "read" or "sequence read" refers to the nucleotide or base sequence information of a nucleic acid that has been generated by any sequencing method. A read therefore corresponds to the sequence information obtained from one strand of a nucleic acid fragment. For example, a DNA fragment where sequence has been generated from one strand in a single reaction will result in a single read. However, multiple reads for the same DNA strand can be generated where multiple copies of that DNA fragment exist in a sequencing project or where the strand has been sequenced multiple times. A read therefore corresponds to the purine or pyrimidine base calls or sequence determinations of a particular sequencing reaction.

As used herein, the term "reference genome" refers to a genome of the same species as that being analyzed for which genome the sequence information is known.

As used herein, the term "region of the genome" refers to a continuous genomic sequence comprising multiple discrete locations.

As used herein, the term "representation of DNA" or "representation of a genome" refers to a sampling of the DNA or genome produced by a restriction endonuclease digestion of genomic or other DNA, followed by linkage of adaptors and then amplification with primers complementary to the adaptors.

As used herein, the term "sample tag" refers to a nucleic acid having a sequence no greater than 1000 nucleotides and no less than two that may be covalently attached to each member of a plurality of tagged nucleic acid molecules or tagged reagent molecules. A "sample tag" may comprise part of a "tag."

As used herein, the term "sample tag set" refers to a plurality of unique sample tags of the same length, wherein the nucleotide sequence of each unique sample tag differs from

the nucleotide sequence of any other unique sample tag in the sample tag set at two or more corresponding positions in the sequence. A unique sample tag selected from the sample tag set may be covalently attached to each member of a unique sample consisting of a plurality of tagged nucleic acid molecules in order to identify a tagged nucleic acid molecule as a member of that unique sample.

As used herein, the term "segments of cDNA," refers to the nucleic acid molecules resulting from fragmentation of cDNA.

As used herein, the term "segments of genomic material" refers to the nucleic acid molecules resulting from fragmentation of genomic DNA.

A nucleic acid molecule which contains a sequence of nucleotides identical to a segment of a cDNA or a segment of genomic material comprises a "segment of cDNA" or a "segment of genomic material," respectively.

As used herein, the "sequence complexity" or "complexity" with regards to a population of polynucleotides refers to the number of different species of polynucleotides present in the population.

As used herein, "substantially identical" or "substantially fully complementary" sequences have at least about 80% sequence identity or complementarity, respectively, to a nucleotide sequence. Substantially identical sequences or substantially fully complementary sequences may have at least about 85%, 90%, 95% or 100% sequence identity or complementarity, respectively.

As used herein, the term "substantially unique primers" refers to a plurality of primers, wherein each primer comprises a tag, and wherein at least 50% of the tags of the plurality of

primers are unique. Preferably, the tags are at least 60%, 70%, 80%, 90%, or 100% unique tags.

As used herein, the term “substantially unique tags” refers to tags in a plurality of tags, wherein at least 50% of the tags of the plurality are unique to the plurality of tags. Preferably, substantially unique tags will be at least 60%, 70%, 80%, 90%, or 100% unique tags.

As used herein, the term "tag” refers to a nucleic acid having a sequence no greater than 1000 nucleotides and no less than two that may be covalently attached to a nucleic acid molecule or reagent molecule. A tag may comprise a part of an adaptor or a primer.

As used herein, a “tagged nucleic acid molecule” refers to a nucleic acid molecule which is covalently attached to a “tag.”

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range, and any other stated or intervening value in that stated range, is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

All combinations of the embodiments of the invention described herein are within the scope of the invention unless clearly indicated otherwise.

*Single Cell Copy Number Profiling by ‘Sequence Counting’ and Sample Tagging.*

High throughput sequencers are in effect single molecule sequencers. Because the human genome is sequenced and assembled, a single molecule sequencer is also a single molecule identifier. Therefore, sequencing is also counting, as illustrated in Figure 4 which illustrates how low coverage sequencing can yield an estimate of copy number which is comparable to array hybridization. The distinction for this is that 'sequence counting' can be done on single cells, and, using the multiplex system disclosed herein, relatively inexpensively, in effect converting the high throughput sequencer into a detector of single cells with abnormal genomes. To the sequencing technology two additional features were added: limiting dilution and sample tagging. For single-cell sequencing, the single cells or the nuclei of the single cells are first deposited into separate wells or are kept physically separated by other means, i.e. a gel. The single cells or the nuclei of cells can be deposited in a well by a cell sorter. Alternatively a collection of cells or nuclei can be diluted into a fluid that is then made semi-solid. A unique sample tag is prepared for each well for sequencing, with a separate sample tag per well. After this step, the wells are pooled and sequenced. The sample tag is used to deconvolute the information, allowing a copy number profile to be made for each individual cell.

Protocol: Isolate single tumor cells or single nuclei of the tumor cell, tag and optionally sample tag the DNA of each cell or nuclei, and amplify the tagged DNA to microgram quantities to generate tagged libraries for Solexa sequencing. The isolation of single tumor cells will occur by Fluoresent Activated Cell Sorting by total genomic DNA content, to deposit single nuclei into individual wells.

Alternatively, tumor cells can be isolated by FACS sorting with fluorescent antibodies against breast epithelium markers,

including cytokeratin 8, 18 or 19. Cells are lysed *in situ* and the genome is fragmented by heating. Libraries are constructed from the fragmented genomic DNA. Preliminary experiments were conducted using high-resolution ROMA microarray to quantify whole-genome copy number on DNA amplified using WGA from single SK-BR-3 cells (Figure 5). These experiments demonstrate the feasibility and efficiency of amplifying 6 pg of DNA into microgram quantities to determine copy number by microarray methods, which suggests that the DNA is also suitable for determining copy number by sequencing.

Since sequencing whole tumor genomes from single cells is costly and time consuming, two methods are proposed to stratify the genome and sequence for copy number: Representation Sequencing and Bin Sequencing.

*Representation Sequencing*

Representation sequencing involves restriction digesting genomic DNA into fragments and sequencing only the stratified population of restriction fragments to determine copy number. This method has the advantage of generating precise restriction fragments that can be reliably prepared as libraries from different individual tumor cells. In this protocol genomic DNA is amplified from a single tumor cell, digested with the DpnII, a second digestion is preformed with a AluI enzyme to remove highly repetitive "Alu SINE elements," sequencing adaptors (Illumina GA, Solexa) with a unique 4 nucleotide sample tag are ligated and the ligated fragments are PCR amplified to generate a sequencing library. *In silico* modeling predicts that the double digest will increase the fraction of unique sequencing reads to 75% from the 55% normally seen in unselected human genome sequencing. The sample tagged DNA libraries are then pooled, and applied to the Solexa flow-cells for cluster amplification and single

read sequencing. Single-end reads are used because they are faster and cheaper than paired-end reads and provide all of the information necessary for 'sequence counting'. Mapping algorithms then align each sequence read with a high quality score (with a read length in excess of 35bp) and an unambiguous sample tag to a specific location on the DpnII restriction map of the human genome. The result is a count of reads at places in the genome, similar to the process demonstrated in Figure 4, except that the data is segregated by sample tag, so that each well receives a copy number profile.

*Bin Sequencing*

It is clear that sequencing of random size-selected fragment from the genome is a viable alternative. In this method, the DNA from single tumor cells is amplified to generate fragments from 200-2000bp. Blunt-end cloning is then used to ligate adaptors to these fragments, followed by PCR amplification to generate libraries (standard Solexa protocol for single read genomic DNA). However, in addition to this protocol adaptor sequences with 4 nucleotide sample tags are ligated to generate unique libraries from each tumor cell genome. The sample tagged DNA libraries are then pooled, and applied to the Solexa flow-cells for cluster amplification and single read sequencing. The 35bp sequences are then mapped to specific regions of the human genome and deconvoluted using the unique sample tag to determine which sequence reads originated from which tumor cells. The crucial step in this procedure that is different from representation methods is to artificially divide the human genome into "bins" of several kilobases. Break analysis in sequence count data has another degree of freedom compared to what is practiced for array data, namely that the boundaries of the bins can be adjusted so as to minimize variance. Each bin contains a number of

random genomic sequence reads from which a mean value is computed to determine genomic copy number at low sequence coverage.

The particular use of the WGA method does not depend on absolutely uniform amplification across the genome. Useful data can be obtained even if only half of the potential fragments from a single cell were amplified in a given reaction. Additionally, redundancy of data gained from segmentation benefits in identifying breakpoints.

Representation Sequencing vs. Bin Sequencing

A comparison of the two methods reveals the number of samples that can be profiled per sequencing run. Of the 3,149,324 Dpn II fragments in size range 200-800 (size optimal for sequencing) there are 618,629 not cut with AluI and thus available to be amplified. Approximately 75% of the fragment ends, or about 450,000, yield unique mappings and thus provide locations for counting. The overall process is presented as a flowchart in Figure 7.

Assuming 15 million quality reads per Solexa sequencing lane, random sequencing yields about 55% or 8.25 million unique reads per lane. The DpnII/AluI representation should increase the number of unique reads to 11.25 million. Modeling a single flow cell lane with 100kb bins distributed across the genome, the projected results are as follows:

Table 1

| # samples | avg reads/bin/sample (random) | avg.reads/bin/sample (Dpn/Alu) |
|---|---|---|
| 1 sample per lane | 267 | 363 |
| 2 samples per lane | 133 | 181 |
| 12 samples per lane | 22 | 30 |

The number of samples (individual cells) to be sequenced per lane, is between 4 and ten. With 8 lanes on a flowcell, 30-80 samples could be profiled at the level of Figure 4 relatively inexpensively.

Finally, the whole genomic copy number (generated by sequencing) will be determined for multiple, individual tumor cells in a single solid tumor. The data will be segmented into sections that minimize variance of copy number, subject to Kolmogorov-Smirnov significance tests, to generate individual tumor copy number profiles. As stated above, an additional degree of freedom, which our algorithms do not yet incorporate, is selecting the "boundary" of the bins. The single cell tumor genome profiles will be compared to study tumor heterogeneity in single solid tumors which may elucidate the progression of genomic events during tumorigenesis.

*Varietal counting of nucleic acids*

Measuring the absolute or relative concentration of nucleic acid sequences in a sample often suffers from severe stochastic distortions due to differential amplifications or losses during processing and manipulation of the nucleic acids. There is never more information available than the number of molecules of the starting material in the sample


IPR2025-01353; IPR2025-01355
CSHL EX2002

itself. The method described herein is designed to capture the information at the earliest stage, in a form that is then resistant to distortion by further amplification.

Input nucleic acid molecules (0-th order derivatives), or first round copies of them (1-st order derivatives), present in the sample, can be randomly tagged using a large choice of distinct tags. Every tagged molecule becomes essentially unique, the combination of the information in the tag and the information in the nucleic acid, which can later be read by sequencing. The combination of a tag with a 0-th order, 1-st order, or subsequent N-th order nucleic acid molecule is called a "tagged nucleic acid molecule." After amplification or selection or other processing, the number of different tags is counted.

Within a pool of tagged nucleic acid molecules, each tagged nucleic acid molecule is likely unique in the pool when a sufficiently large number of distinct tags is used. Then in subsequent amplifications these unique tagged nucleic acid molecules are amplified to facilitate detection, but no new species of tagged nucleic acid molecules are created. Thus, distorted counts are not created provided that the number of different tagged nucleic acid molecules is counted.

In the disclosed method, each sequence read generated from a tagged nucleic acid molecule comprises two parts, or subsequences. The first part corresponds to the tag, which identifies the tag (and the sample tag if samples are pooled). The second part corresponds to a nucleic acid in the sample which is mapped to a location in the genome (or to a transcript if RNA molecules are being counted). After deconvoluting by sample tag, the number of different tagged nucleic acid molecules mapped to each location is counted. There are two separate methods of counting, and each can be combined for greater accuracy and determination of confidence.

The first counting method is called "maximum tagged nucleic acid molecule number." The number of different tagged nucleic acid molecules at a given location cannot exceed the true (absolute) copy number. Therefore, for a given region comprising multiple locations, the maximum copy number of the region is not less than the maximum number of different tagged nucleic acid molecules mapped to any location in that region. The maximum number of different tagged nucleic acid molecules centered about a moving window of a fixed number of locations can be taken as the measurement of true copy number of the window. This will provide at worst an underestimate of the maximum true copy number for that window, and never an overestimate. The set of maximum number of different tagged nucleic acid molecules is one type of copy number profile.

The counting method described above is most accurate when the processing efficiency is excellent and the number of molecules mapped to each location is less than the number of tags. Under such conditions virtually each molecule in the sample is tagged and counted, and the true count can be derived. But there is a second method.

The second counting method is called "total different tagged nucleic acid molecules," and is useful, for example, when tagging efficiency is low. The total number of different tagged nucleic acid molecules within a region of a fixed number of locations will be, within a statistical counting error, a monotonic function of true copy number, regardless of the efficiency of tagging. Low efficiency of tagging merely increases the statistical error. The relative copy number of a region can be estimated by comparing the total number of different tagged nucleic acid molecules mapped to the region to, for example, the mean value of tagged nucleic acid molecules mapped to other regions of the genome which have a comparable number of locations.

If the number of molecules at a location on a genome or cDNA library exceeds the number of tags, and tagging is too efficient, copy number information can not be recovered. Rather, the total number of tags at each location would merely be measured. There are at least three ways to remedy this condition: 1) diminish the efficiency of the first step of the process, for example by reducing the reaction time; 2) reduce the amount of sample; and 3) increase the total number of tags. For example, if the length of the tag were N nucleotides, there are $4^N$ possible tags. For N = 15, there are approximately $10^9$ tags, easily in excess of the number of molecules per location.

The collection of reads lend themselves to modeling the power of inference. For each location, there exists the number of tagged nucleic acid molecules observed, and for each species of tagged nucleic acid molecule, the number of times it was observed. With a good estimate of the input, the number of genomes, critical parameters can be accurately deduced because the vast majority of the copy numbers are two per genome. The limiting case of single cell analysis lends itself to the most accurate modeling, with the proportion of null reads enabling the inference of the probability, theta, that a single molecule at a location is detected. Theta is a function of the efficiency of tagging, the efficiency of amplification and processing, and the read depth.

Using the language invented of the WGA reaction, in the specific protocol described here, all first order derivative DNA molecules (i.e., first copies) are tagged, with no new tags introduced when subsequent order derivatives are created. This enables the "counting" of original molecules, alleles in the case of genome analysis. This method is referred to herein as "varietal counting." As long as sequence depth is sufficient, subsequent amplification or methods of trapping

that enrich for regions of interest, or the combination, cannot distort the count. Pooling from multiple cells becomes a natural extension of this method, which improves efficiency, increases throughput and decreases costs.

*Tags*

The tags are nucleotide sequences with constant portions and variable portions. The length of the constant portion and the length of the variable portion may be the same or different. The length of the tags within a plurality of tags is typically, but not necessarily, the same for all tags. In an embodiment, the total length of a tag is less than 100 nucleotides.

The constant portions are used to aid in the tagging of nucleic acid molecules. For example, the constant portion may contain a nucleotide sequence which allows hybridization to a poly-N tail added by a terminal transferase. The constant portion may also be used to aid in manipulation of the tagged nucleic acid molecules. For example, the constant portion may contain a sequence which is a primer binding site useful for PCR amplification. The constant portion may also be used to identify a sample if there are multiple samples being processed in parallel, i.e., the constant portion may contain a sample tag.

The sequence of individual tags does not need to be known prior to generating tagged nucleic acid molecules. A plurality of tags may have variable portions whose sequences are elucidated for the first time during sequencing. The variable portion of a tag may be a sequence of nucleotides where each nucleotide is individually one of A, T, C, or G, or one of any two of A, T, C, and G, or one of any three of A, T, C, and G, i.e. the variable portion of a tag may consist of only two or three species of nucleotides. Tags may be designed such that

the variable portion has a random sequence. Tags may be designed such that the variable portions are constructed from a set of dinucleotides or trinucleotides.

Tags may be associated with nucleic acid molecules by using, singly or in combination, ligation, nick-ligation, terminal transferase, hybridization, priming, and the like. For example, nucleic acid molecules may be cut with a restriction endonuclease(s), treated with a terminal transferase to add a poly-N tail, hybridized to a tag molecule, and then nick ligated to form a tagged nucleic acid molecule. Specific methods of generating tagged nucleic acid molecules are discussed in detail below.

*Tagging with the aid of a terminal transferase*

Figure 8 illustrates one embodiment of this method, where the $N_{10}$ designates an oligonucleotide having a tag component and a sample tag component. The first six nucleotides (4096 possible sequences), randomly generated, provide the tag component, and the last four nucleotides (256 choices) are chosen to encode the micro-well, and, hence, a sample tag. The lengths of the tag and the sample tag can be changed to suit needs, and are not limited to a total of ten nucleotides. The lengths are chosen for illustration. The sample tag allows pooling, with deconvolution occurring afterwards. The combination of a tag with a nucleic acid molecule from the input sample provides an essentially unique tagged nucleic acid molecule for the first derivative strand, that can only with very low probability be replicated by chance from another like molecule (a chance of roughly 1/N where N is the number of available tags). In the end, multiple copies of sequence-identical molecules in the input sample are each individually counted. In applications to mRNA counting, the length of the tag component can be increased to accommodate possibly thousands of transcripts from the same gene.

Although there are multiple embodiments of the basic method besides the one illustrated in Figure 8, the embodiment illustrated in Figure 8 has several useful design features. First, it avoids the use of ligases (an alternate method for associating a tag with a nucleic acid molecule), which has been found to be an inefficient reaction. Instead, this embodiment uses terminal transferase (TTase), which is a robust enzyme from which high efficiency is expected. Second, once the tag and sample tag are incorporated into the first derivative DNA strand, micro-wells can be pooled and the subsequent steps carried out more efficiently in larger volumes with larger numbers of molecules, possibly with the addition of carrier nucleic acids. Because of a clean-up step, removing free adaptor, no more unique first order derivative strands with tags can be created, and second order strands can be arithmetically expanded if found to be useful. Third, the method extends immediately from DNA to mRNA. Rather than cleaving and adding a polynucleotide tail with terminal transferase, the naturally occurring polyA tail of mRNAs can be utilized. Fourth, it is easy to see how to combine analysis of both RNA and DNA from the same cell with this schema. Instead of tailing DNA with A, it can be tailed with, for example, C, thus distinguishing the DNA from the RNA. DNA and RNA derivatives can either be read together in the same sequencing runs, or separately by amplification using their differing PCR primer adaptors. Fifth, by using a longer sample tag, sequences belonging to a particular cell can be preferentially amplified for deeper study when warranted.

This method can be applied to assessing DNA methylation. After the first order derivative strands are made, the duplex DNA can be pooled from multiple samples. These molecules will be hemi-methylated, and still retain epigenetic marks. Therefore they can then be separated into methylated and unmethylated samples either by cleavage with methylation sensitive

restriction enzymes, or by partitioning with antibodies or methyl-C binding proteins directed to methylated (or hydroxyl methylated) cytosine. Error rates will be measured based on expectations from regions of known methylation state.

*Tagging by ligation*

In another embodiment of the invention, adaptors with tag(s) can be ligated to sample DNA or RNA that has been converted to DNA. In the second step these molecules can be pooled (or not) and amplified for sequencing. In the third step the numbers of each different tagged nucleic acid molecule of each locus (or mappable sequence read) are compiled, giving statistics from which copy number can be inferred.

In another embodiment, to measure RNA, RNA is first converted to DNA using methods that produce a single double stranded molecule for each RNA in the sample. (For example, prime with oligo dT and reverse transcriptase, followed with random priming and Klenow polymerase.) The procedures for analyzing DNA can then be followed.

In another embodiment, the sample DNA can be sheared and the ends prepared for ligation. However, there are three major advantages to cleaving with one or more restriction endonucleases (RE). The first is that the expected number of tagged nucleic acid molecules at each location can be precisely predicted from the starting number of cells in the sample, or conversely the number of cells in the sample inferred from the expected number of tagged nucleic acid molecules per location. The second is that the expected complexity of the product is reduced, enabling copy number determination from fewer sequence reads. The third are certain advantages in the flexibility and functionality that can be built into the tags.

In an embodiment, sample nucleic acid is cleaved with a specific restriction enzyme and sample molecules are ligated to tagged (and sample tagged if appropriate) oligonucleotide adaptors which contain sequences that allow PCR amplification. There are various designs for these adaptors, and also alternatives to PCR (such as WGA), based on the functionalities built into the tags. Some of these additional functionalities are discussed below.

The ligated molecules now are essentially unique, the ligated molecules each being the combination of a substantially unique tag and DNA sequence to which they were ligated. Further amplification may distort the yields of molecules, but will not create new tags (although some may be lost by statistical processes which are readily modeled during analysis).

Through sequence error during processing, the tag may very occasionally mutate, creating the appearance of a new tagged nucleic acid molecule. But if the set of tags are sufficiently large, the tagged nucleic acid molecules will generally differ at more than one nucleotide, and so tagged nucleic acid molecules that differ by a single nucleotide can (optionally) be ignored when counting.

For single cells, individual cells are loaded at most one per microwell. The nucleic acid is then prepared in wells, and the molecules tagged. Once tagged they can be pooled for further processing, which greatly improves the uniformity and efficiency of DNA amplification methods.

In yet another embodiment, DNA is cut, and adaptors are ligated to both ends of each fragment, enabling them to be amplifiable by PCR. To avoid adaptor-adaptor self ligation, double stranded adaptors without 5’ phosphorylation can be used, so an adaptor is ligated on one strand only at each end of the cleaved sample DNA. Then the molecules are pooled, and

treated as for representations: the shorter single strand adaptor is melted off at elevated temperature but well below the Tm for the sample duplex, the ends of the double stranded DNA filled in with Klenow polymerase, unligated adaptors removed, and then the sample is PCR amplified.

In the preceding embodiment, adaptors are used not only for tagging but also for amplification. Adaptors are for identifying each molecule individually. They can also be used for amplification, as described above, or for other uses (see below). Because in the above embodiment adaptor ligation at both ends of sample fragments is required for amplification, the yield may be diminished. To improve yield, the method can work with molecules that have an adaptor at one end only.

In one embodiment, the reactions are performed as before, but tags do not contain sequences enabling PCR. Prior to PCR, but after an optional pooling, fresh adaptors (different composition and without tags) are ligated so that molecules with only a single tag adaptor can be amplified and eventually counted. If the original tag has a selection sequence, molecules with the tag adaptor before sequencing can be optionally enriched by hybrid selection.

In another embodiment, initial adaptors are designed with a phosphorylated 5′ end and a 3′ dideoxy end, so that the initial ligation occurs to the opposite strand as described in the embodiment above. (Instead of using an RE with a 5′ overhang, one which leaves a 3′ overhang may be preferred.) After removal of unligated adaptors, there follows several rounds of arithmetic amplification, followed by tailing molecules with dT using terminal transferase, and amplification of the processed sample DNA by PCR.

In another embodiment, DNA from sample is labeled as before, pooled if required, and unligated adaptors removed from the

reaction. The reaction is subject to whole genome amplification (WGA), and the product of WGA is used to make sequencing libraries. For this embodiment, the initial adaptors may also contain: (A) cleavage sites for ease of making sequencing libraries; and (B) a sequence large enough for optional enrichment by hybrid selection prior to making libraries for sequencing.

In the embodiments just described, it is assumed that the tag is double stranded. The production of a great variety of double stranded molecules is not completely straightforward. In one embodiment, single strand DNA is synthesized as type I and type II. Type II has three sets of functionalities A, B, C, and D. Functionalities B contain the tag sequence, C the adhesive ends with a RE site, and D the buffer. The B will contain some length of runs with random sequence, for example 10 nucleotides with a total complexity of $4^{10}$, or about $10^6$ possible sequences. Type I oligo will be complementary to type II at functionality A, the two joined and annealed, and then extended with polymerase to create a double stranded molecule. After cleavage with the RE site, the resulting double stranded molecule is the tag used for the subsequent tagging of the sample.

Additional methods for producing single and double stranded oligonucleotides suitable for use as tags are known in the art, for example, in U.S. Patent Nos. 5,639,603 and 7,622,281, U.S. Patent Application Publication No. 2006/0073506 and Parameswaran et al. (2007) (7,5,8,4).

*Nick-ligation*

In this embodiment of the invention, 3'OH ends of nucleic acid molecules, e.g., segments of genomic material or mRNA transcripts, are first extended with terminal transferase, adding a poly-N tail. Primer adaptors are then annealed to the


IPR2025-01353; IPR2025-01355
CSHL EX2002

poly-N tails and the resulting duplex with a nick is sealed with DNA ligase. This series of reactions is much more efficient than a single step double stranded ligation. The resulting product is amplified by PCR. This method has been successfully used with as little as 1ng of DNA. Tagged nucleic acid molecules generated using the nick-ligation method can be generated from a single cell DNA, and then pooled prior to PCR. Pooling from 100 cells yields the DNA concentration needed for successful amplification. Sample tags can be added in the nick ligation step, and this is readily achieved (see Figure 9).

*Technology for inexpensive assay of regional markers*

The above description is of methodology designed to measure copy number and mutational loads throughout much of the genome in single cells. But the application of discoveries, once made, of predictive markers for outcome or therapeutic response will not require whole genome information. Our present outcome markers are in fact copy number markers at very specific loci. Essentially the same methodology as disclosed hereinabove can be employed to make these assessments very affordable. Consider, for example, the case where it is desirable to assess a large study for the predicative power of a set of N markers, and these markers are copy number markers that can be assessed by varietal counting. In principle, sample coding, pooling, and parallel capture and sequencing can be used to assay a large number of samples in a single sequencing lane. The major difference is that we are not restricted to minute quantities of DNA, so WGA is not needed, and assays will be more robust. We have shown in our autism work that excellent copy number data can be obtained from read depth after capture. Thus we can use direct library preparation with just sample tags. The target area for capture is likely to be different and smaller than the exome. Thus we

can pool more samples per sequencing lane, and the sequencing costs for assaying hundreds of markers in hundreds of samples can be reduced. Of course, there are costs associated with the development and testing of the capture reagents, and the cost of the capture itself, and these are not negligible, but if a test were to come into widespread clinical use, it would be inexpensive, and this is one of our major objectives. There is a clear alternative, namely creating a bank of region specific PCR primers with tags and sample tags. The tags would be necessary because random unique ends created by shearing would not be available for counting based methods. This alternative plan could be less expensive and more robust than capture (see Figure 10).

*Single cell RNA measurements*

Several methods have now been published whereby researchers have profiled the RNA content of single cells by sequencing (Tang et al., 2009). This is potentially a very powerful method for analysis of tumor cells. A varietal counting method could be very useful for RNA profiling from single cells, and for the same reasons that it is useful with DNA. When variety is not introduced by nucleic acid processing varietal counting provides a way to avoid the distortions created by PCR. RNA and DNA can be analyzed from the same cells. The DNA analysis enables us to cluster cells into stromal and tumor subpopulations, and this in turn facilitates the interpretation of the RNA profiling. As an example of utility, suppose one is interested in searching thousands of circulating cells, or cells from a biopsy, or from the margins of a tumor, for rare tumor cells. DNA copy number can facilitate its identification, and the RNA from that cell can then be reliably used to infer additional properties, such as the tissue of origin. This combination is especially powerful

for the early detection of the new incidence or recurrence of cancer.

**Examples**

*Methylation sequencing*

Currently, high throughput sequencing requires physical purification of sub-genomic regions of interest in order to obtain high coverage of those regions. The method of sequence capture has been applied to examine methylation in large sets of CpG islands at the sequence level. Methylation sequencing is based on the conversion of C residues to T using bisulfite reagents, with methylated C's detected as those protected from C to T conversion. This means either that sequence capture must take place before bisulfite treatment, which requires more starting DNA than is available for patient specimens, or that sequence capture must be designed to accommodate the C to T conversion. The bisulfite based sequencing method used herein adopts the latter and so is suitable for high throughput single molecule sequencing by either the Illumina or Roche 454 methods, using as little as 100 ng of input DNA. The sequencing method depends on hybrid capture of bisulfite treated DNA followed by sequencing through the Illumina Solexa sequencer and mapping back to the genome using an algorithm developed Dr. Andrew Smith (19). It was observed that CpG islands can exhibit multiple different levels of methylation, from imprinted (allelic) to randomly partially methylated, to completely methylated or unmethylated. In particular, one observation that is important for this invention is that many islands switch from completely unmethylated in normal tissues to virtually completely methylated in matched tumor tissue ("DNA-methylation switch domains"). Figure 3 shows three examples of the simplest cases in normal cells versus the breast cancer cell line MDA-MB-231: stably unmethylated (A and B); the ALX gene, switched from unmethylated to completely

methylated (C and D); and a case, SSTR4, where only a segment of the island is switched (E and F). This complete switch makes it possible to detect rare tumor cell DNA in bodily fluids or other cell populations using nested PCR of bisulfite treated DNA and opens up the possibility of a creating a method for the early detection of recurrence in the form of circulating metastatic cells.

*Comparative genomic hybridization (CGH) by high throughput sequencing*

Until now the most economical method for high resolution genomic analysis has employed some form of microarray. The rapid rise of 'next generation' DNA sequencing is changing that landscape. Both the Illumina (Solexa) and Roche 454 instruments can provide millions of individual sequence reads randomly distributed across any input DNA (Craig et al., 2008). The depth of coverage of these technologies depends on the total length of the input DNA, from whole genome (large breadth, lower depth) to narrowly focused regions achieved by hybrid capture of specific sequences (narrow breadth, higher depth) (Albert et al., 2007; Hodges et al., 2007; Okou et al., 2007). Converting sequencing reads to copy number analysis is a matter of mapping reads and then put them into 'bins' of a certain size, depending on the purpose, counting up the numbers in the bins and performing appropriate statistical analysis to define copy number change breakpoints.

Sequence counting was applied to the degraded formalin-fixed paraffin-embedded (FFPE) DNA that was available from retrospective clinical trials with very encouraging results. Figure 4 displays a comparison of microarray and sequence counting results in a region around Her2 on chromosome 17 for a 10 year old paraffin-embedded clinical trial sample JZ33. The gray and orange points are (respectively) the raw ratio data for 244K Agilent microarray and one lane of Solexa

sequencing reads divided into 30K bins. Note that the sequencing (orange) track shows significantly less noise. The green and blue lines are the result of segmenting each data type. The segments are generally concordant, but the sequence counting method reveals a somewhat more detailed pattern. Another feature of note from Figure 4 is the robustness of our segmenting program, irrespective of the variation of noise in the two sets of raw data.

The efficacy of sequence counting on FFPE DNA and the rapidly decreasing cost of DNA sequencing in general provides an opportunity to convert our previous microarray methods to a highly flexible and scalable platform for breast cancer genome diagnostics. Both the breadth of DNA to be sequenced and the size and location (and therefore resolution) of individual bins can be easily adjusted to fit the particular loci to be assayed.

*Single cell microarray and sequencing*

A representation of DNA amplified from single cells was used for genomic profiling by 'sequence counting'. As a preliminary step in developing that technology, DNA was amplified from batches of 1, 10 and 100 cells of the cell line SKBR3 using the Whole Genome Amplification (WGA) Kit from Sigma-Aldrich and compared these amplified representations to equivalent microarrays on the same format done with a standard SKBR3 DNA prep. Individual cells were picked from adherent culture dishes using a cell transfer micropipette. After random WGA1 amplification the DNA preps were digested with DpnII and AluI, adaptors added and only DpnII fragments without AluI sites were amplified by PCR. These representations were labeled and applied to ROMA custom microarrays with 390,000 probes designed to detect DpnII fragments of appropriate size across the genome. The idea was to test how reproducible this two step process would be when

only 2-6 chromosomes are present in each single cell preparation (SKBR3 is roughly triploid).

The results are shown in Figure 5, presenting a comparison of the segmented profiles of a batch preparation of SKBR3 DNA (Figure 5A) against our normal male reference cell line (SKN-1) and three independent single cell DNAs (Figure 5 B,C,D). A visual inspection of these profiles makes it clear that after segmentation, the results for single cells are virtually identical to a standard DNA, leading us to have confidence that sequencing DNA from a representation will yield useful results. Although the raw ratio data is not presented here, we note that the standard deviation of the ratio data is much higher for the three single cell experiments (0.706, 0.725, 0.676) than for a batch DNA preparation (0.313). The extra noise does not adversely affect segmentation, although in certain regions very fine detail may be lost (red arrow in Figure 5A). This is comparable to other WGA1 experiments such as the one depicted in gray in Figure 4. In this light it is also worth noting that our particular use of the WGA1 method does not depend on absolutely uniform amplification across the genome. We benefit from the redundancy of data that we gain from segmentation to identify breakpoints. Even if only half of the potential fragments from a single cell were amplified in a given reaction, we would still obtain useful data.

*Sequencing across chromosome breakpoints*

One of the methods for the detection of rare tumor cells, in lymph nodes, blood or other bodily fluids is the amplification or sequencing of unique breakpoints resulting from deletions and/or translocations that occur during breast tumor development. Although events such as deletions and duplications are apparent from array CGH, the actual structure underlying the copy number changes is not.

The DNA sequences surrounding 22 breakpoints from two breast cancer cell lines were enriched by hybrid capture on custom microarrays as described by Hodges et al (20). Following hybrid capture, the enriched DNA was amplified and prepared for 'paired end' sequencing on the Illumina GA2 instrument. Paired end sequencing yields approximately 36 bp of sequence from each end of a single DNA molecule (after preparation for sequencing, each molecule is approximately 200 bp in length), and those sequences are then mapped back to the consensus genome. For a molecule from a normal, non-rearranged region the two matched reads should map approximately 130 bp apart. In contrast, at the edges of a deletion, the two reads from a single molecule would be separated on the map by the width of the deletion, or in the case of a translocation, would map to another chromosome. These are referred to as 'bridging fragments' because they connect regions not normally adjacent to one another. Examples of two such bridging fragments are shown at the top of Figure 6, along with the original CGH pattern, and graphs of the distance between paired ends (blue) and the total number of reads (red) across the capture region (X axis). By mapping the exact location of a breakpoint, highly sensitive and specific probes that will distinguish tumor cells from large numbers of normal cell can be prepared. Of 22 regions displaying deletions captured from cell lines T47D and MDA-MB-436, 15 behaved as simple deletions from which unique 'bridging fragments' could be identified as in Figure 6. Two additional sites 'jumped' to different chromosomes indicating translocations. The other 5 sites involved repeated sequences on one of the paired for which the identification of the exact breakpoint was not possible and which would not be suitable for creating unique probes.

*Copy number profiling using varietal tags on restriction enzyme fragments.*

As one illustration of varietal counting, copy number information was obtained using a "total different tagged nucleic acid molecules method.", DNA fragments were first associated with tags by using terminal transferase and nick-ligation (Figure 9). Genome DNA was obtained from the cell line SKBR3. DNA (2µg) was digested by NlaIII, resulting DNA fragments with 4bp, "CATG", 3' overhangs. This 3'OH end was extended with terminal transferase by adding a poly-A tail. One of primer adaptors was designed to have tags, 12 Ts and CATG at the 3' end, so it could be annealed to the 3' ends of the DNA fragments with a nick site (Figure 9, Product 3). This nick site was ligated by DNA ligase, forming duplex DNA fragments with tags. Another primer adaptor was then added to these duplex DNA fragments by one cycle of PCR (Figure 9, Step 5), allowing the resulting products of this reaction to be further amplified by PCR, followed by sequencing (Illumina/Solexa). Then, by mapping the subsequence of each tag associated sequence read to a location on the genome, the number of DNA fragments (assigned to the same genomic location) with different tags was counted.

We have previously obtained copy number profiles of SKBR3 genomic DNA from a million cell library, by sequencing and sequence counting. Figure 11 compares the use of varietal counting (a "total different tagged nucleic acid molecules method") with 'sequence counting' for analyzing copy number data, genome-wide (Figure 11A) and partial-genome (Figure 11B). The essentially equivalent result obtained using varietal counting illustrates that by adding varietal tags to DNA fragments before PCR one can use varietal counting as described herein to reliably obtain copy number information unaffected by amplification distortion.

Data analysis for the procedure described above employed a modification of the 'bin-counting' method because the NlaIII restriction sites are not distributed evenly in the genome. Therefore, in order to ensure equal weight for each of the 50K bins, the bin boundaries for this experiment were designed to encompass and equal number of NlaIII fragments sized between 100bp and 250bp, the size range from which the Illumina libraries were made.

**Discussion**

Single cell data is much more informative than data obtained from multiple cells, for example from array hybridization of multiple sections of a primary tumor. First, one can afford to analyze many more genomes by multiplex single cell sequencing than by array hybridization (on the order of a hundred single cells versus a dozen sections). Second, one is able to directly assess the exact proportion of cells in a region of a tumor that are malignant. Third, one is able to make a far more detailed analysis of the temporal order of genetic lesions, since one will not be looking at the average state of a population. Fourth, one is able to assess whether tumor heterogeneity is of the anatomically segregated or the co-habitation type, or both.

The invention described herein may be used to determine the original genome copy number (or the gene expression profile) of a sample without amplification distortion. All measurements of copy number or gene expression, especially involving multiple samples, are potential applications. Samples can be nucleic acids derived from single cells or populations of cells. The invention can be used to study cellular heterogeneity in tumors or normal populations of cells.

The invention may be used to generate a genomic profile of a cell, e.g., a cancer cell, or population of cells. From the

genomic profile it can be determined if specific genes are amplified or deleted. When the sample is derived from expressed RNA, and the regions are expressed genes, the relative counts comprise an expression profile. From the expression profile the tissue of origin of a cancer cell can be identified. When the sample is a mixture of genomes from different species, and the genomic regions distinguish the different species, the relative counts comprise a census of those different species. For example, from a census we can determine the population of microbial flora and relate the same to disease.

One of the advantages of the use of tags is that it allows the use of restriction endonuclease fragments in genome analysis. Counting methods for genome copy number determination use counts of independent reads. Two reads are not considered independent if they are the result of DNA amplification occurring during sample processing. Typically, this entails shearing sample DNA fragments, and the position of ends of the fragments are used to disambiguate PCR duplicates. If two DNA reads have the same ends, they could be PCR duplicates, and not independent, so they are counted as only one true read. One cannot disambiguate restriction fragments this way, because two reads of the same fragment will have the same ends. However, by the addition of tags, independent reads can be distinguished, if they are added before any amplification step, as they will have different tags. By contrast, if two reads of the same genome region have the same tag, they are highly likely to be PCR duplicates, and not independent.

The ability to use sequence of restriction fragments in genome analysis can be exploited to reduce the cost of determining copy number profile, much as performing ROMA facilitates genome analysis. By size selection of fragments, one can sample the genome at greater depth at fewer loci, and still obtain a high resolution copy number profile but with fewer

reads and therefore at lower cost. This is the essence of representation, and results in complexity reduction, the central element of ROMA. In ROMA, the analysis of the representation is done by hybridization to arrays, and here the representation is analyzed by sequencing.

Further complexity reduction can be achieved by sequence capture. The sequence capture can be designed specifically for the fragments that can be predicted by in silico methods to map to the regions where one wants copy number mutation. In this way one can target special regions of the genome, such as the exome, or known cancer genes. Very little sequence of the sample is then needed for copy number determination, and samples can be multiplexed at a great cost saving.

The invention disclosed herein may also be used to determine the proportion of components in a mixture. For example, suppose one possesses a set of reagents, such as antibody molecules, that bind to a set of substrates, such as different proteins, and one desires to detect the relative proportion of those substrates in a sample, such as blood. First, one covalently combines a tag having a sample tag with each reagent, where the sample tag portion identifies the reagent, and the varietal tag will be used in counting. The tagged reagents are then combined with the sample, and reagents that combine with substrates are physically separated. For example, the substrates may be biotinylated before being mixed with reagents, then reagents combining with substrates separated by avidin affinity chromatography. The tags are amplified by PCR and sequenced, effectively yielding relative counts for each reagent species and therefore each substrate. In this case, the constant portions of the tags include the sample identifier and nucleotides used for amplification.

The invention provides the following advantages: (1) the method is simple; (2) with statistical analysis the confidence

of a measurement can be determined with great accuracy; (3) this is a sequence based method, which in the long run will be cheaper than arrays; (4) long read lengths are not necessary, providing savings; (5) because tagged nucleic acid molecules are counted, many processing methods that create distortion in yields can be applied without major loss in the tagged nucleic acid molecule count; (6) among these processing methods are hybrid selection (aka capture microarrays), which restrict the sequence information to regions of importance, and thus reduce cost by reducing the total number of reads per sample needed for determination of the desired information; (7) the molecules from one sample can be combined with tags and a sample tag, creating sample tagged nucleic acid molecules which can then be pooled, and after processing and sequencing, the sample tags decoded, giving a profile for each sample; (8) pooling adds efficiency and uniformity; (9) the samples can be individual cells, yielding large numbers of single cell measurements cheaply.

The invention described herein has clinical applications as described above, including: (1) determination of the genomic heterogeneity of tumors; (2) detection of tumor cells in needle biopsies, in particular an assessment of the heterogeneity of the cells in a biopsy; (3) assessment of the spread of a cancer to regional lymph nodes; and (4) detection of malignant cells in sites such as blood and bone marrow. The presence of cells with the molecular signature of the primary in the blood or bone marrow opens up the possibility of using more general single cell/molecule assays as a first line of defense in the early detection of cancer, i.e. breast cancer in women at risk.

When there is radiological evidence of a breast mass, the patient most commonly undergoes fine needle biopsy, which is examined by the pathologist who makes an assessment of the

malignancy of the mass. Because needle aspiration destroys the histology of the biopsy, one of the pathologist's canonical tools for assessing malignancy is not available. Single cell sequencing would not only more accurately predict the presence of malignant cells, but also give what we might call "genomic staging." Is the tumor a simplex type of complex type? Is it mongenomic or polygenomic? Are there critical loci that are amplified or deleted that might immediately suggest a preoperative therapy?

The single cell/molecule approach allows the assessment of the spread of cancer to regional lymph nodes and to establish the DNA-methylation states at switch regions.

Successful application of single cell/molecule detection in the above contexts will lead to application to the detection of cancer recurrence, and the measurement of cancer load. Detection of recurrence at the earliest possible time may prolong survival for a large class of patients by earlier intervention than they would otherwise receive. Moreover, if cancer load were monitored in the patient, even before the site of the cancer growth is discernible by imaging, then one could measure the response of the cancer to a variety of therapeutic measurements. That is, the most effective choice of therapy could be determined empirically in the patient.

Additional applications exist in the diagnosis and selection of treatment for patients with auto-immune diseases, and in vitro fertilization (IVF) screening of pre-implant embryos for genetic defects such as trisomy-21.

**References**

The disclosures of the following references in their entireties are hereby incorporated by reference into this application.

1. Pinard R, de Winter A, Sarkis GJ, Gerstein MB, Tartaro KR, Plant RN, et al. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics. 2006; 7:216. PMCID: PMC1560136.

2. Klein, C. et al. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. PNAS. 1999; 96, pp. 4494–4499.

3. Stoecklein, N. et al. SCOMP Is Superior to Degenerated Oligonucleotide Primed-Polymerase Chain Reaction for Global Amplification of Minute Amounts of DNA from Microdissected Archival Tissue Samples. American Journal of Pathology. 2002; 161(1):43-51.

4. Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130.

5. U.S. Patent No. 7,622,281, Ronaghi et al.

6. Eid et al. Real-Time DNA Sequencing from Single Polymerase Molecules. Science. 2009; 323:133-138.

7. U.S. Patent No. 5,639,603, Dower et al.

8. U.S. Patent Application Publication No. 2006/0073506, Christians et al.

9. Navin N, Krasnitz A, Rodgers L, Cook K, Meth J, Kendall J, et al. Inferring tumor progression from genomic

heterogeneity. Genome Res. 2010;20(1):68-80. PMCID: 2798832.

10. Fisher B, Redmond CK, Fisher ER. Evolution of knowledge related to breast cancer heterogeneity: a 25-year retrospective. J Clin Oncol. 2008;26(13):2068-71.

11. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25. PMCID: PMC2690996.

12. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4(4):406-25.

13. De Baetselier P, Roos E, Brys L, Remels L, Gobert M, Dekegel D, et al. Nonmetastatic tumor cells acquire metastatic properties following somatic hybridization with normal cells. Cancer Metastasis Rev. 1984;3(1):5-24.

14. Duelli DM, Padilla-Nash HM, Berman D, Murphy KM, Ried T, Lazebnik Y. A virus causes cancer by inducing massive chromosomal instability through cell fusion. Curr Biol. 2007;17(5):431-7.

15. Jorgensen HF, Adie K, Chaubert P, Bird AP. Engineering a high-affinity methyl-CpG-binding protein. Nucleic Acids Res. 2006;34(13):e96. PMCID: PMC1540740.

16. Meehan RR, Lewis JD, Bird AP. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. Nucleic Acids Res. 1992;20(19):5085-92. PMCID: PMC334288.

17. Huang J., Pang J., Watanabe T., Ng HK, Ohgaki H. Whole genome amplification for array comparative genomic hybridization using DNA extracted from formalin-fixed,

paraffin-embedded histological sections. J Mol Diagn. 2009 Mar;11(2):109-16. Epub 2009 Feb 5.

18. Talseth-Palmer BA, Bowden NA, Hill A, Meldrum C, Scott RJ. Whole genome amplification and its impact on CGH array profiles. BMC Res Notes. 2008 Jul 29;1:56.

19. Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L, McCombie WR, Wigler M, Hannon GJ, Hicks JB. High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res. 2009 Sep;19(9):1593-605. Epub 2009 Jul 6.

20. Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ, McCombie WR. Genome-wide in situ exon capture for selective resequencing. Nat Genet. 2007 Dec;39(12):1522-7. Epub 2007 Nov 4.

21. genome.ucsc.edu/index.html?org=Human&db=hg19&hgsid=171216665

22. Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S.. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135.

23. U.S. Patent No. 7,537,897, issued May 26, 2009 (Brenner et al.).

24. McCloskey M.L., Stoger, R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767.

What is claimed is:

1. A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

2. The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome by assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise

a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. The method of any one of claims 4-5, wherein the second region of the genome comprises a centromere.

7. The method of claim 1, further comprising summing the counts obtained in step (f) for locations which comprise a region of the genome, and comparing the sum to a sum obtained from a reference sample for the same region of the genome, thereby estimating a relative genomic copy number of the region of the genome.

8. A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA;

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b), and

ii) a tag;

d) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

e) generating tag associated sequence reads by sequencing the product of step (d);

f) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a cDNA or a segment of cDNA to a location on the cDNA library; and

g) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

9. A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under

conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. A method for obtaining from genomic material DNA methylation information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands to a polymerase reaction in the presence essentially unique primers, wherein the essentially unique primers are capable of hybridizing to the polynucleotide tail of the zero-th order strands, under conditions that promote the formation of only one complement, thereby generating tagged nucleic acid molecules;

d) separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules;

e) subjecting each group of step (d) to amplification by polymerase chain reaction (PCR);

f) generating tag associated sequence reads by sequencing the product of step (e);

g) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

h) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining DNA methylation information unaffected by amplification distortion.

11. The method of any one of claims 1-7, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the

polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. The method of any one of claims 1-7, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. The method of any one of claims 9-12, wherein adding a polynucleotide tail comprises the use of a terminal transferase.

14. The method of any one of claims 1-7, wherein tagging the segments to generate tagged nucleic acid molecules comprises ligation of adaptors comprising the tags to at least one end of the segments of the genomic material.

15. The method of claim 14, wherein the adaptors comprising the tags are ligated to only one end of the segments of the genomic material.

16. The method of any one of claims 1-15, wherein the tags comprise a sequence that aids PCR amplification.

17. The method of any one of claims 1-16, wherein each tagged nucleic acid molecule comprises one tag.

18. The method of any one of claims 1-17, wherein each tagged nucleic acid molecule comprises more than one tag.

19. The method of any one of claims 1-7 or 10-18, wherein segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

20. The method of claim 8, wherein segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

21. The method of any one of claims 8-9, 13, 17-18 or 20, wherein the maximum copy number of a location on a cDNA library is not less than the number of tagged nucleic

acid molecules having a different tag that have been assigned to the same location on the cDNA library.

22. The method of any one of claims 1-7 or 10-19, further comprising analyzing mRNA copy number.

23. The method of any one of claims 10, 13, 16-19, or 22, wherein separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules is by cleavage with methylation sensitive restriction enzymes, partitioning with antibodies, or partitioning with methyl-C binding proteins directed to methylated or hydroxymethylated cytosine.

24. The method of any one of claims 1-23, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or prior to sequencing.

25. The method of any one of claims 1-24, wherein each tagged nucleic acid molecule differs at more than one nucleotide.

26. The method of any one of claims 1-25, wherein the tag sequences further comprise a sample tag.

27. The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to sequencing.

28. The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to sample tag.

29. The method of any one of claims 1-28, wherein the tagged nucleic acid molecules are generated from a single species.

30. The method of claim 29, wherein the tagged nucleic acid molecules are generated from a single organism.

31. The method of any one of claims 1-30, wherein tagged nucleic acid molecules are generated from a single cell.

32. The method of claim 31, wherein the single cell is from a needle aspirate of suspected cancer lesions.

33. The method of claim 31, wherein the single cell is from a core biopsy of suspected cancer lesions.

34. The method of claim 29, wherein the tagged nucleic acid molecules are generated from two or more organisms.

35. The method of claim 34, wherein the tagged nucleic acid molecules are generated from a single cell of each organism.

36. The method of any one of claims 30-35, wherein the single species is human.

37. The method of any one of claims 1-28, wherein the tagged nucleic acid molecules are generated from two or more species.

38. The method of claim 37, wherein the tagged nucleic acid molecules are generated from a population of microbes.

39. The method of claim 38, comprising comparing the genomic copy number information obtained for different species of the population to determine the relative count of those different species in the population.

40. A composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

d) adding a polynucleotide tail to the first order derivative strands;

e) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (c) and (e) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

41. A composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of genomic material;

b) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

c) subjecting the zero-th order derivative strands of step (b) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (b) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

d) subjecting the product of step (c) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (b) under conditions which promote the formation of only one complement, wherein the primers of step (d) have different nucleotide sequences than the primers of step (c), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (c) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

42. A composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged

nucleic acid molecules being produced by a process comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA; and

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, wherein each tagged nucleic acid molecule comprises

iii) one cDNA from step (a) or one segment of the cDNA from step (b) and

iv) a tag.

43. A composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining mRNA transcripts;

b) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the

polynucleotide tail added in step (c) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

44. The composition of matter of any one of claims 40-41 or 43, wherein adding a polynucleotide tail comprises the use of a terminal transferase.

45. The composition of matter of any one of claims 40-44, wherein each tagged nucleic acid molecule comprises one tag.

46. The composition of matter of any one of claims 40-44, wherein each tagged nucleic acid molecule comprises more than one tag.

47. The composition of matter of any one of claims 40-41 or 44-46, wherein segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

48. The composition of matter of any one of claims 42-46, wherein segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

49. The composition of matter of any one of claims 40-48, wherein the complexity of the tagged nucleic acid molecules is reduced by hybrid capture.

50. The composition of matter of any one of claims 40-49, wherein each tagged nucleic acid molecules differs at more than one nucleotide.

51. The composition of matter of any one of claims 40-50, wherein the tag sequences further comprise a sample tag.

52. A composition of matter comprising at least two pools of tagged nucleic acid molecules according to claim 51.

53. The composition of matter of any one of claims 40-52, wherein the tagged nucleic acid molecules are generated from a single cell.

54. The composition of matter of claim 53, wherein the single cell is from a needle aspirate of suspected cancer lesions.

55. The composition of matter of claim 53, wherein the single cell is from a core biopsy of suspected cancer legions.

56. The composition of matter of any one of claims 40-52, wherein the tagged nucleic acid molecules are generated from two or more organisms.

57. A kit for determining nucleic acid copy number information unaffected by amplification distortion comprising:

a) a terminal transferase, and

b) plurality of substantially unique primers,

wherein the substantially unique primers comprise substantially unique tags, and wherein the substantially unique primers are capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the terminal transferase.

58. The kit of claim 57, further comprising a DNA polymerase having 3'-5' proofreading activity.

59. The kit of any one of claims 57-58, wherein the plurality of substantially unique primers comprises $10^n$ primers, wherein n is an integer from 2 to 9.

60. The kit of any one of claims 57-59, further comprising a second terminal transferase and a primer capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the second terminal transferase.

61. The kit of any one of claims 57-60, wherein the tags are six nucleotides long.

62. The kit of any one of claims 57-60, wherein the tags are 15 nucleotides long.

63. The kit of any one of claims 57-62, further comprising sample tags.

64. The kit of claim 63, wherein the sample tags are 2 nucleotides long.

65. The kit of claim 63, wherein the sample tags are 4 nucleotides long.

66. The kit of any one of claims 63-65, wherein the sample tags comprise a sample tag set.

67. The kit of any one of claims 63-66, wherein the substantially unique tags comprise the sample tags.

# VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

## Abstract of the Disclosure

A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising obtaining segments of the genomic material, tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material and a tag, subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR), generating tag associated sequence reads by sequencing the product of the PCR reaction, assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome, and counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome, thereby obtaining genomic copy number information unaffected by amplification distortion.

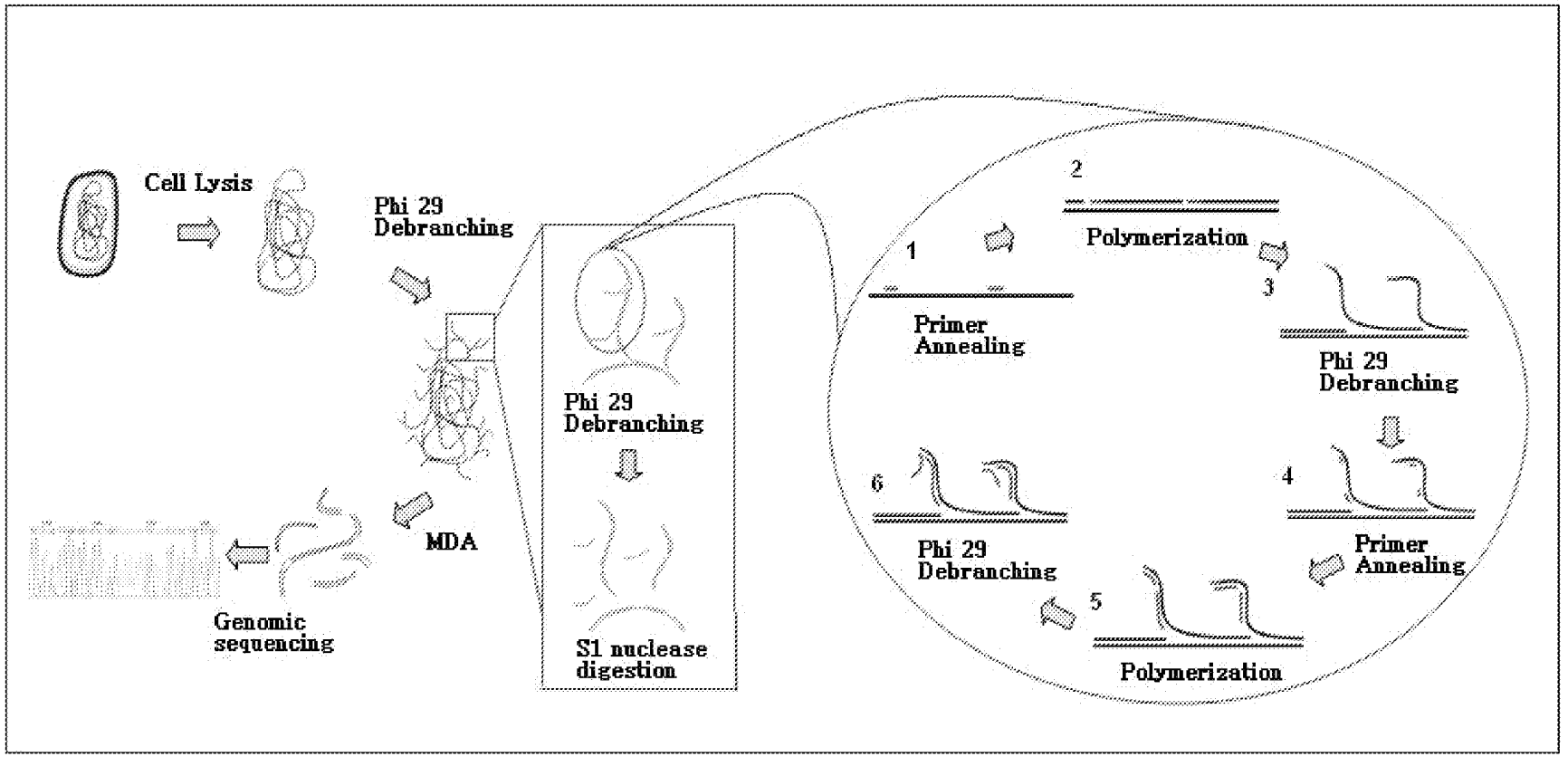



Fig. 1



IPR2025-01353; IPR2025-01355
CSHL EX2002

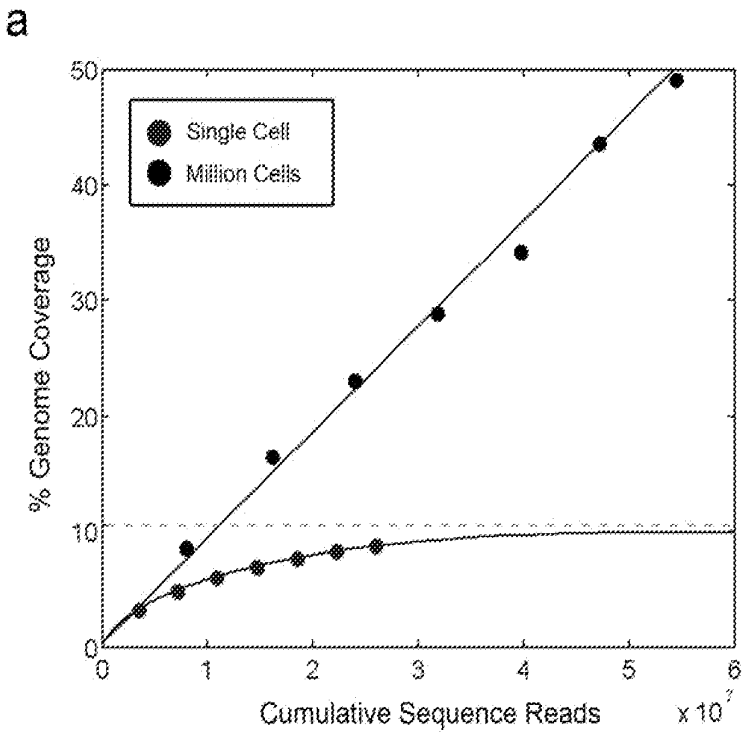



Fig. 2



IPR2025-01353; IPR2025-01355
CSHL EX2002

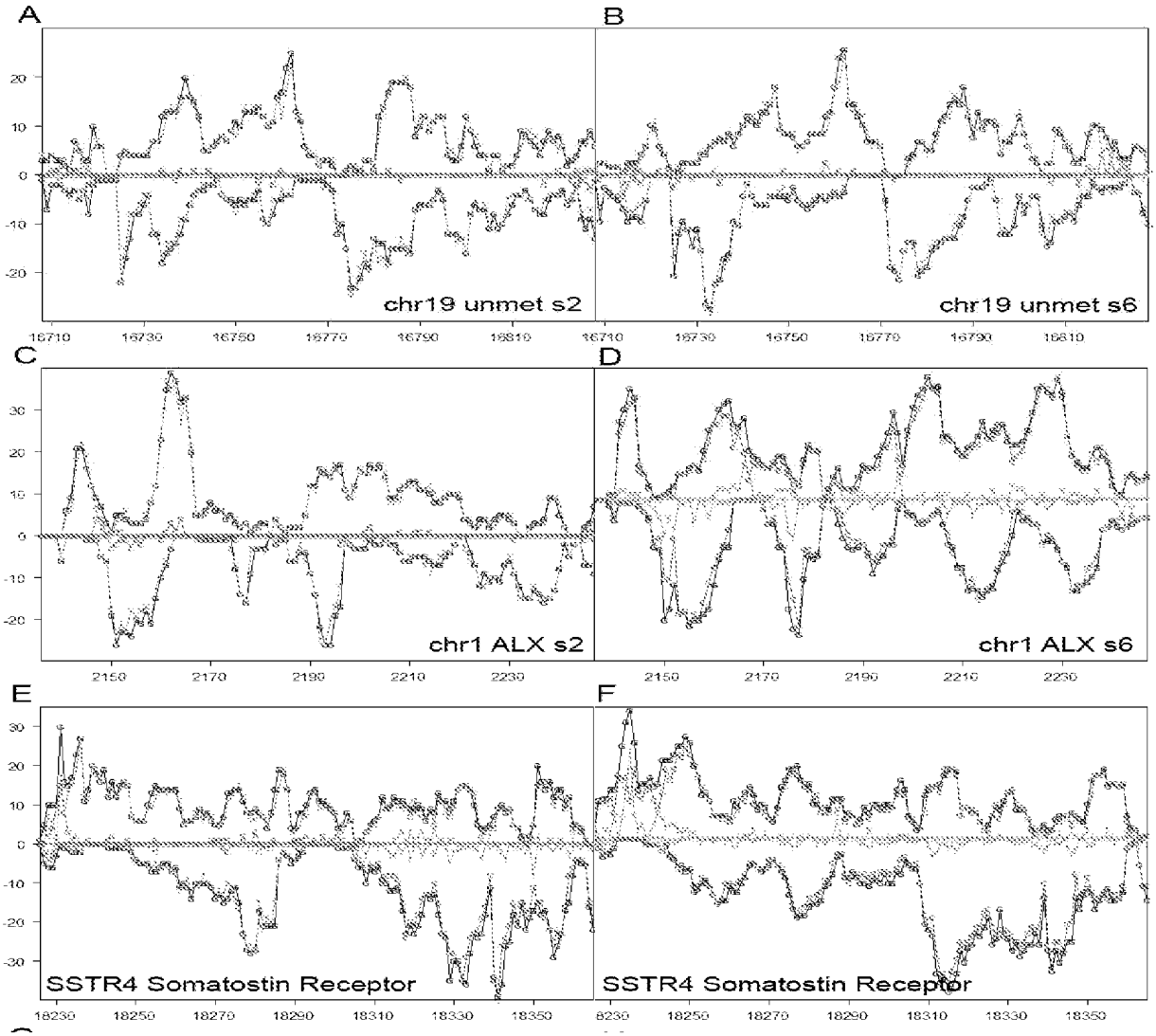

A
B
chr19 unmet s2
chr19 unmet s6
C
D
chr1 ALX s2
chr1 ALX s6
E
F
SSTR4 Somatostin Receptor
SSTR4 Somatostin Receptor

Fig. 3

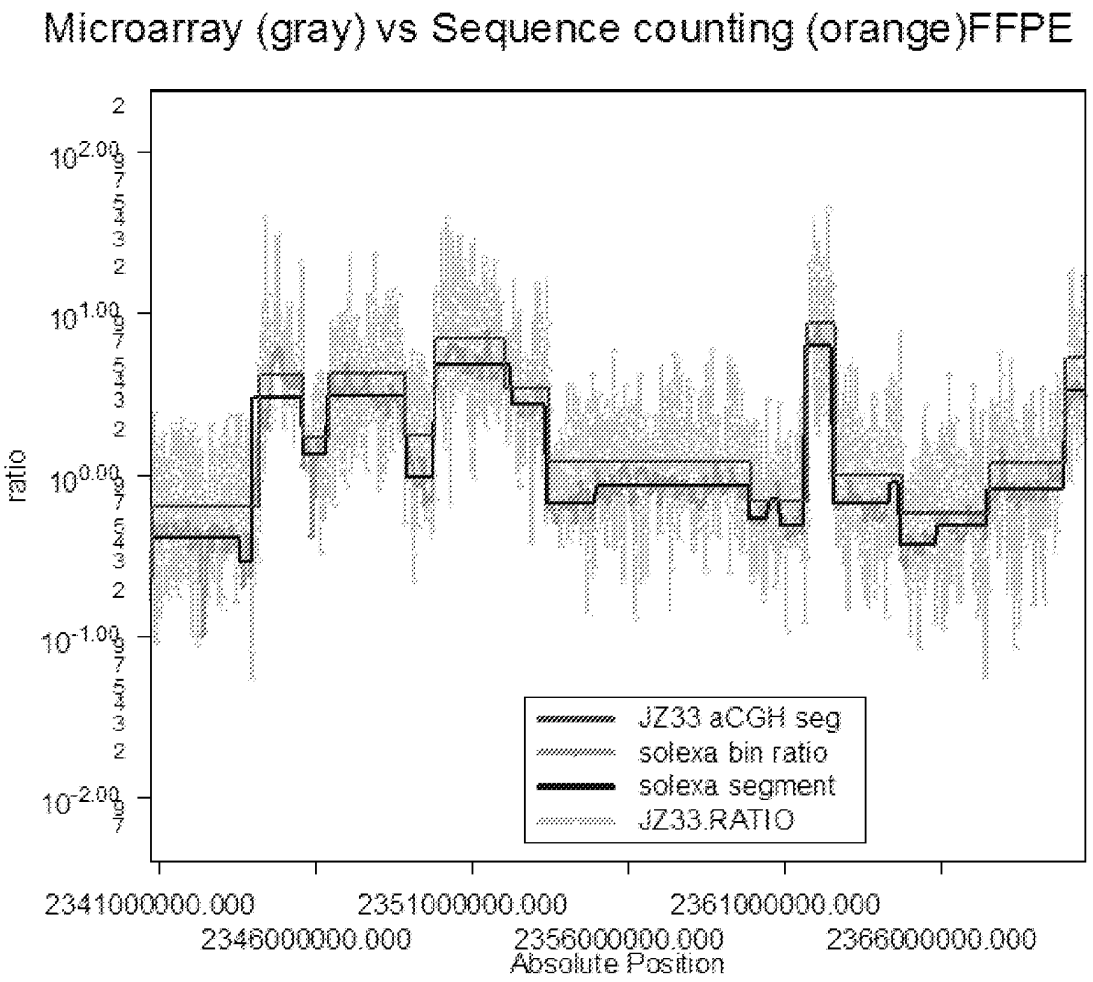



Fig. 4



IPR2025-01353; IPR2025-01355
CSHL EX2002

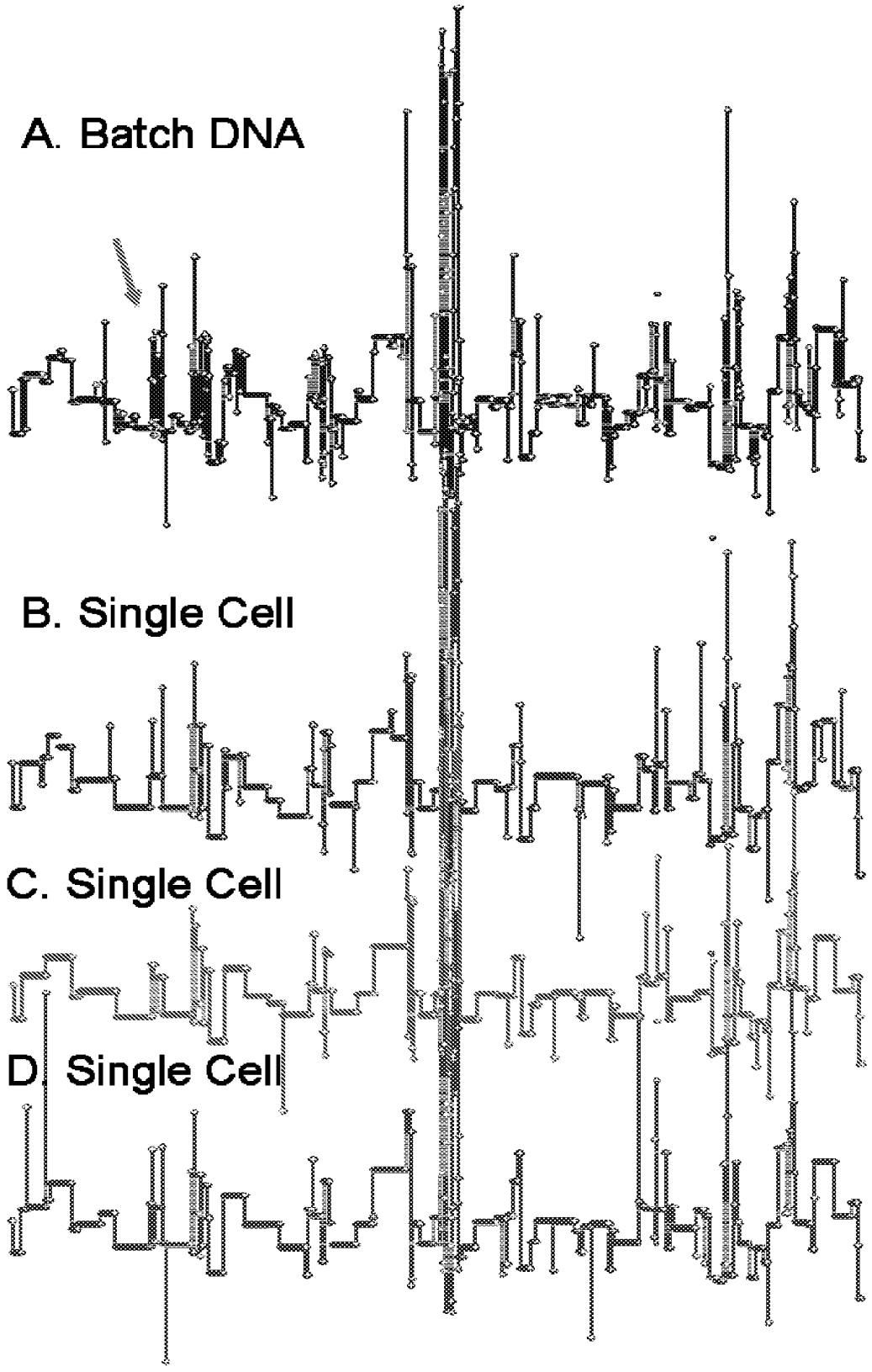



Fig. 5

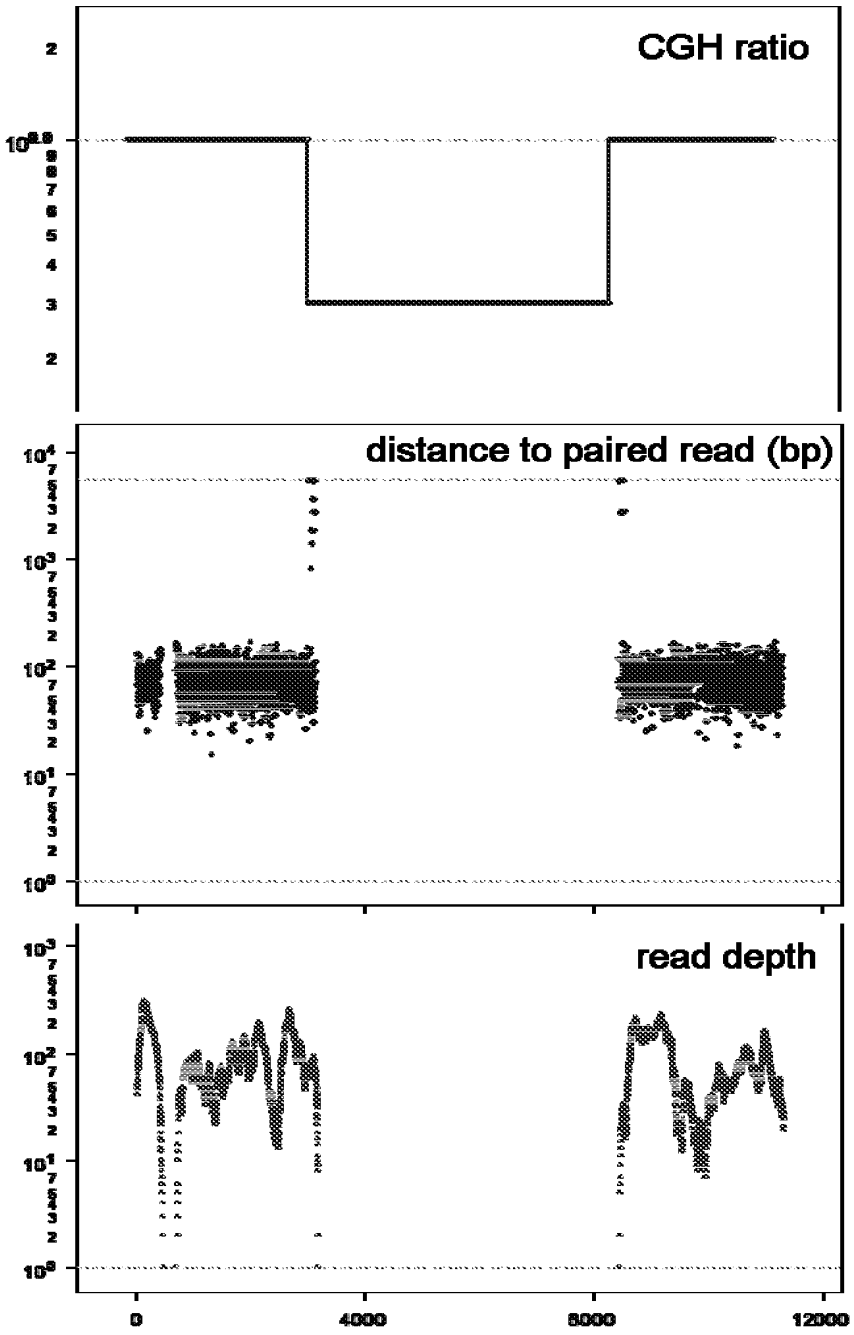



Fig. 6


IPR2025-01353; IPR2025-01355
CSHL EX2002

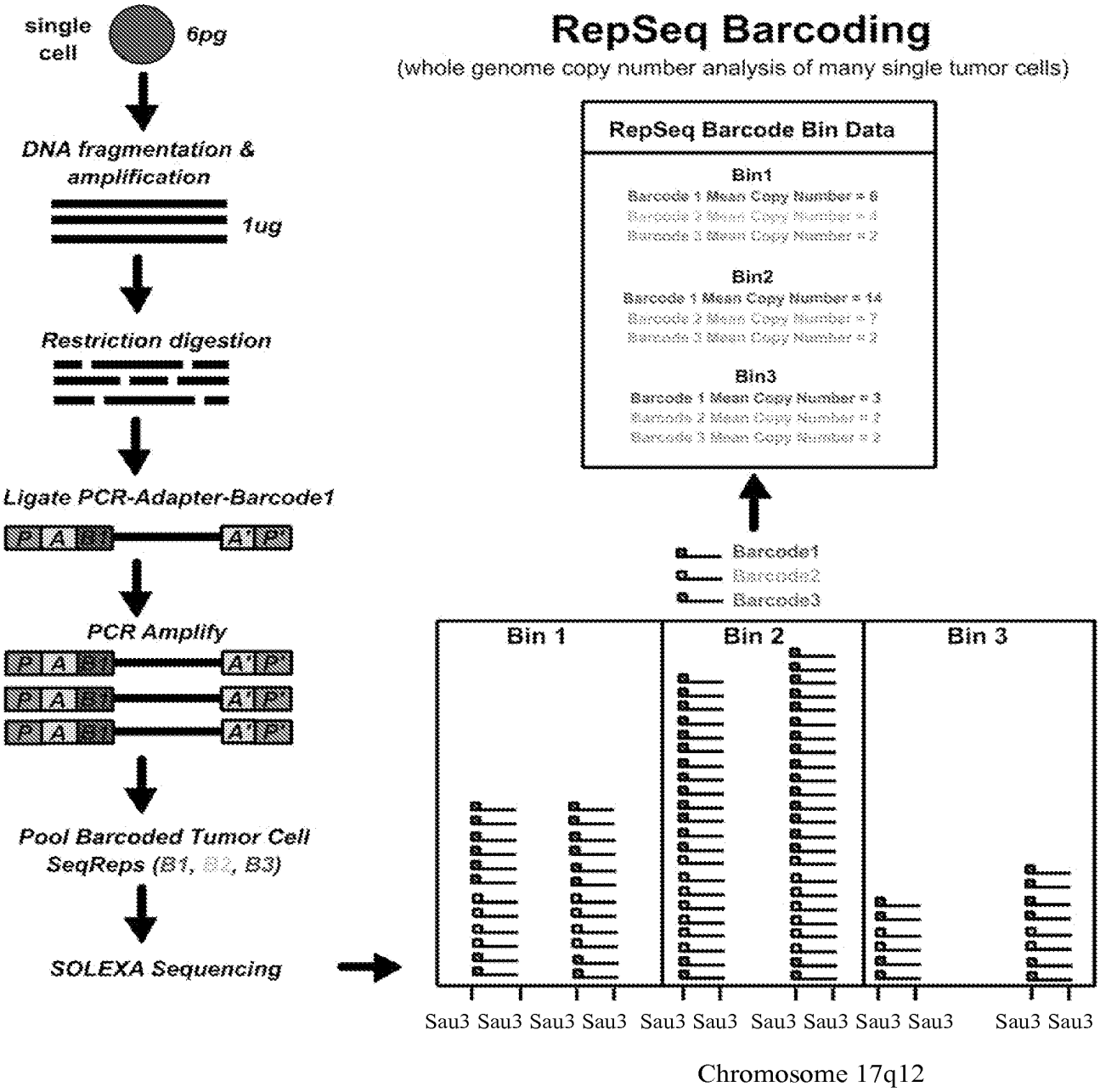



Fig. 7



IPR2025-01353; IPR2025-01355
CSHL EX2002

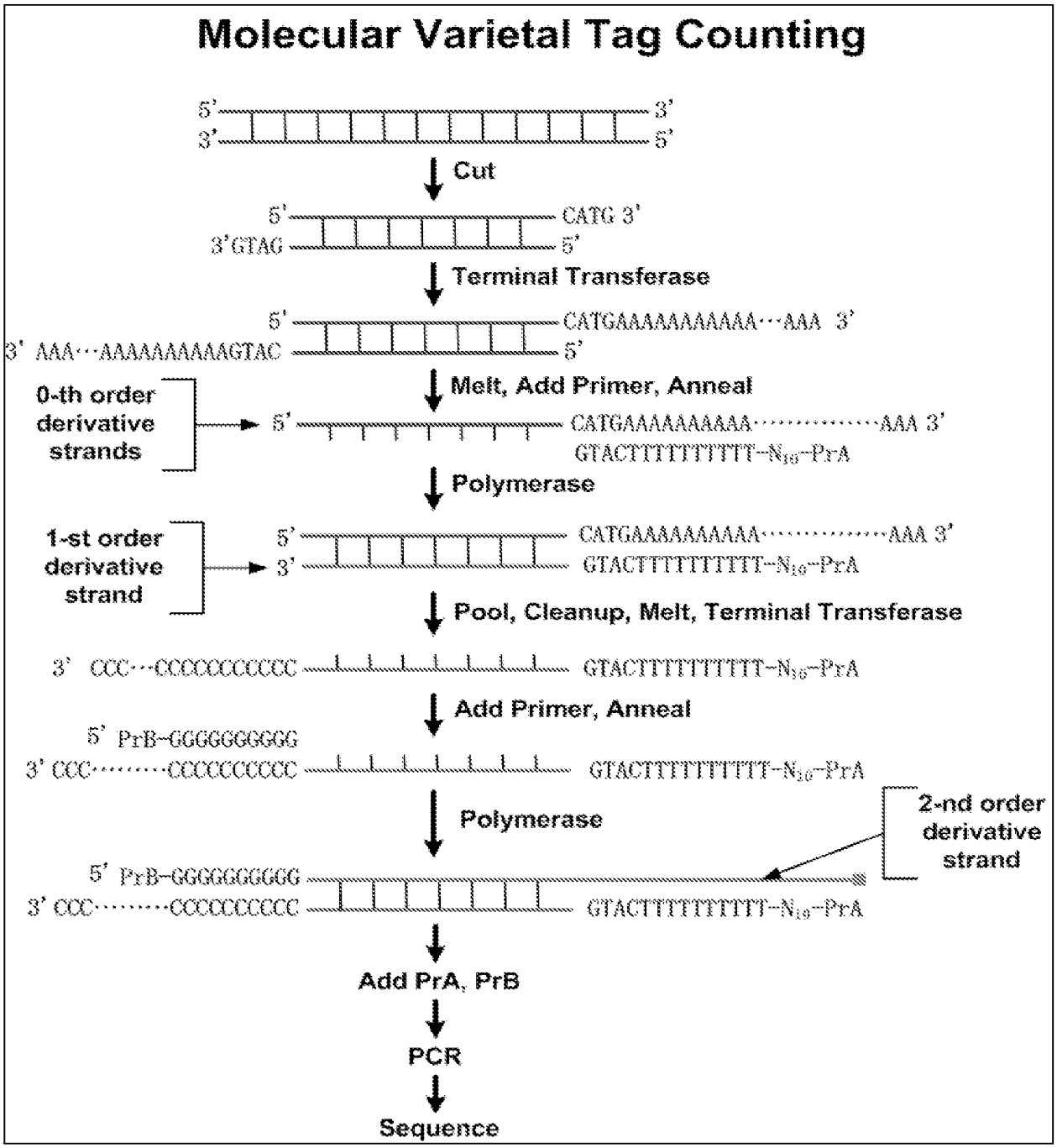

Molecular Varietal Tag Counting
5'
3'
3'
5'
Cut
5'
CATG 3'
3'GTAG
5'
Terminal Transferase
5'
CATGAAAAAAAAAA···AAA 3'
3' AAA···AAAAAAAAAGTAC
5'
Melt, Add Primer, Anneal
0-th order derivative strands
5'
CATGAAAAAAAAAA·····AAA 3'
GTACTTTTTTTTTT-N10-PrA
Polymerase
1-st order derivative strand
5'
3'
CATGAAAAAAAAAA·····AAA 3'
GTACTTTTTTTTTT-N10-PrA
Pool, Cleanup, Melt, Terminal Transferase
3' CCC···CCCCCCCCCC
GTACTTTTTTTTTT-N10-PrA
Add Primer, Anneal
5' PrB-GGGGGGGGG
3' CCC·····CCCCCCCCC
GTACTTTTTTTTTT-N10-PrA
Polymerase
2-nd order derivative strand
5' PrB-GGGGGGGGG
3' CCC·····CCCCCCCCC
GTACTTTTTTTTTT-N10-PrA
Add PrA, PrB
PCR
Sequence

Fig. 8

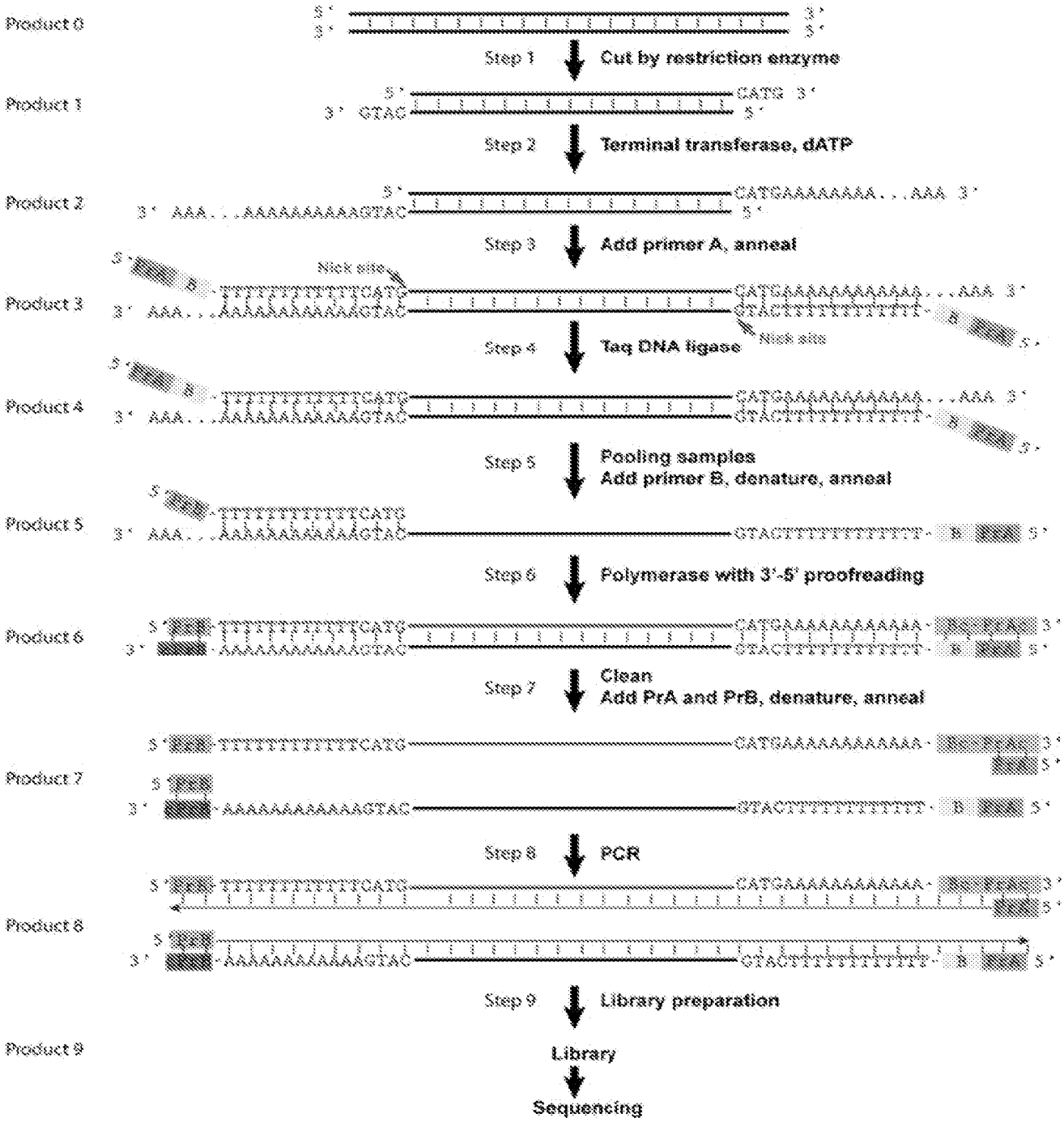

Product 0
Step 1
Cut by restriction enzyme
Product 1
Step 2
Terminal transferase, dATP
Product 2
Step 3
Add primer A, anneal
Nick site
Product 3
Step 4
Taq DNA ligase
Nick site
Product 4
Step 5
Pooling samples
Add primer B, denature, anneal
Product 5
Step 6
Polymerase with 3'-5' proofreading
Product 6
Step 7
Clean
Add PrA and PrB, denature, anneal
Product 7
Step 8
PCR
Product 8
Step 9
Library preparation
Product 9
Library
Sequencing

Fig. 9

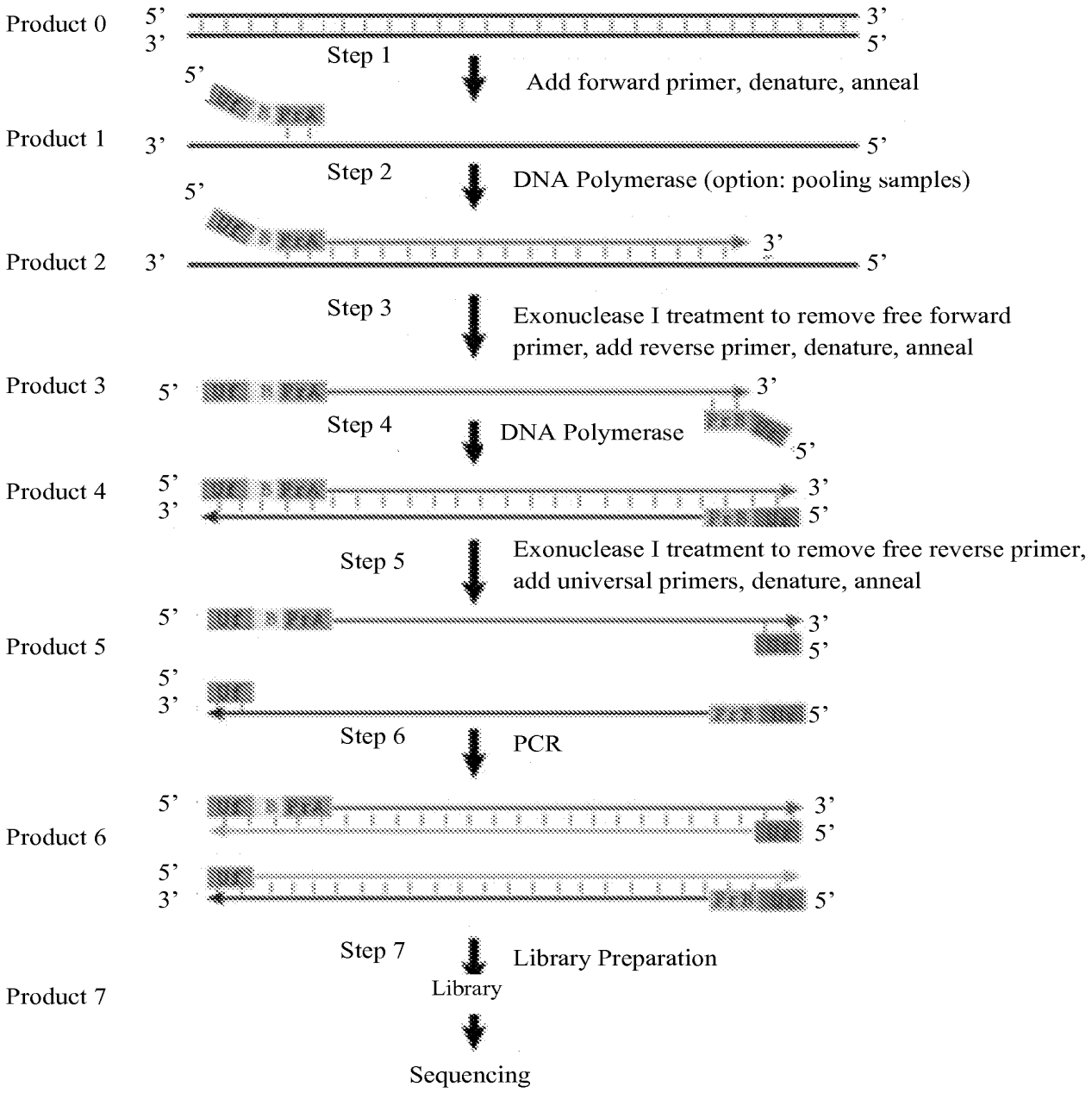

Product 0
5'
3'
3'
5'
Step 1
Add forward primer, denature, anneal
5'
Product 1
3'
5'
Step 2
DNA Polymerase (option: pooling samples)
5'
3'
Product 2
3'
5'
Step 3
Exonuclease I treatment to remove free forward primer, add reverse primer, denature, anneal
Product 3
5'
3'
Step 4
DNA Polymerase
5'
Product 4
5'
3'
3'
5'
Step 5
Exonuclease I treatment to remove free reverse primer, add universal primers, denature, anneal
5'
3'
Product 5
5'
3'
5'
Step 6
PCR
5'
3'
Product 6
5'
3'
5'
Step 7
Library Preparation
Product 7
Library
Sequencing

Fig. 10

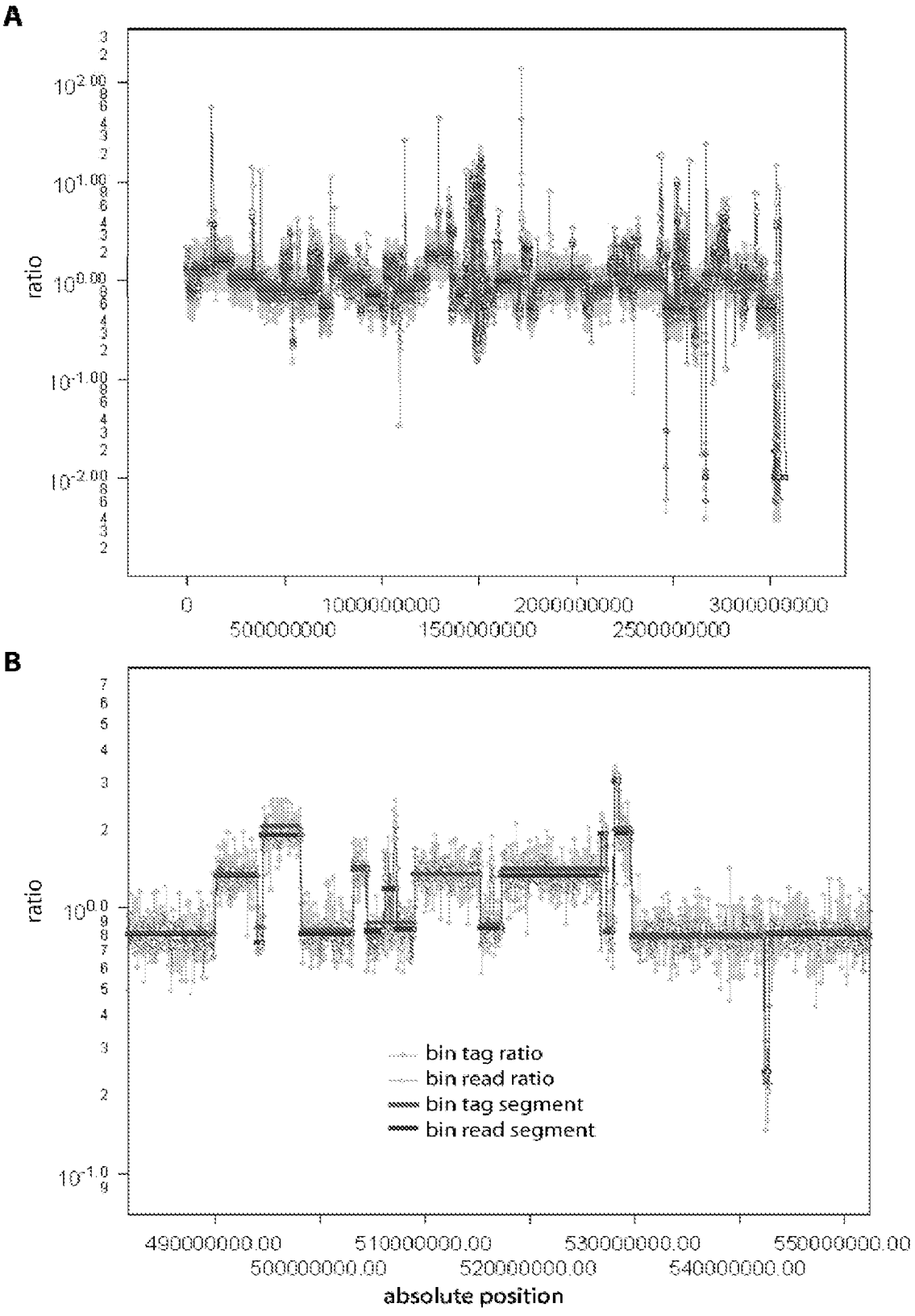

A
ratio
0
500000000
1000000000
1500000000
2000000000
2500000000
3000000000
B
ratio
bin tag ratio
bin read ratio
bin tag segment
bin read segment
490000000.00
500000000.00
510000000.00
520000000.00
530000000.00
540000000.00
550000000.00
absolute position

Fig. 11

## *DECLARATION AND POWER OF ATTORNEY*

*As a below-named inventor, I hereby declare that:*

*My residence, post office address, and citizenship are as stated below next to my name.*

*I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:*

VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

*the specification of which:*
*(check one)*

__X__ *is attached hereto.*

_____ *was filed on* _____ *as*

*Application Serial No.* _____

*and was amended* October 21, 2011
*(if applicable)*

*I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.*

*I acknowledge the duty to disclose to the U.S. Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.*

*I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International Application which designated at least one country other than the United States, listed below. I have also identified below any foreign application for patent or inventor's certificate, or PCT International Application having a filing date before that of the earliest application from which priority is claimed:*

| *Prior Foreign Application(s)* | | | *Priority Claimed* | |
|---|---|---|---|---|
| *Number* | *Country* | *Filing Date* | *Yes* | *No* |
| N/A | | | ___ | ___ |
| | | | ___ | ___ |
| | | | ___ | ___ |
| | | | ___ | ___ |
| | | | ___ | ___ |





IPR2025-01353; IPR2025-01355
CSHL EX2002

*I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below:*

| *Provisional Application No.* | *Filing Date* | *Status* |
|---|---|---|
| 61/510,579 | July 22, 2011 | Pending |
| 61/406,067 | October 22, 2010 | Pending |
| | | |
| | | |
| | | |

*I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States Application(s), or Section 365(c) of any PCT International Application(s) designating the United States listed below. Insofar as this application discloses and claims subject matter in addition to that disclosed in any such prior Application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56, which became available between the filing date(s) of such prior Application(s) and the national or PCT international filing date of this application:*

| *Application Serial No.* | *Filing Date* | *Status* |
|---|---|---|
| N/A | | |
| | | |
| | | |
| | | |
| | | |
| | | |

*And I hereby appoint*

Gary J. Gershik (Reg. No. 39,992); John P. White (Reg. No. 28,678); Christopher C. Dunham (Reg. No. 22,031); Norman H. Zivin (Reg. No. 25,385); William E. Pelton (Reg. No. 25,702); Robert D. Katz (Reg. No. 30,141); and Paul Teng (Reg. No. 40,837).

*and each of them, all c/o Cooper & Dunham LLP, 30 Rockefeller Plaza, 20th Floor, New York, New York 10112, my attorneys, each with full power of substitution and revocation, to prosecute this application, to make alterations and amendments therein, to receive the patent, to transact all business in the Patent and Trademark Office connected therewith and to file any International Applications which are based thereon under the provisions of the Patent Cooperation Treaty.*



IPR2025-01353; IPR2025-01355
CSHL EX2002

*Please address all communications, and direct all telephone calls, regarding this application to:*

Gary J. Gershik, Esq. *Reg.No.* 39,992
*Cooper & Dunham, LLP (Customer Number 23432)*
*30 Rockefeller Plaza*
*20th Floor*
*New York, New York 10112*
*Tel. (212) 278-0400*

*I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.*

*Full name of sole or first joint inventor* James Hicks

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* US

*Residence* Huntington, NY 11743

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Nicholas Navin

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* US

*Residence* 1621 Wichita St., Houston, TX 77004

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Jennifer Troge

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* US

*Residence* 100 Simcoe St., Oyster Bay, NY 11771

*Post Office Address* Same as above


IPR2025-01353; IPR2025-01355
CSHL EX2002

*Full name of additional joint inventor (if any)* Zihua Wang

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* China

*Residence* 13 Meadowlawn St., Huntington, NY 11743

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Michael Wigler

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* US

*Residence* 52 Snake Hill Road, Cold Spring Harbor, NY 11724

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* _____

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* _____

*Residence* _____

*Post Office Address* _____

*Full name of additional joint inventor (if any)* _____

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* _____

*Residence* _____

*Post Office Address* _____



IPR2025-01353; IPR2025-01355
CSHL EX2002

Dkt. 5981/81503-A/GJG/ACK

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicants : James Hicks et al.

Serial No. : Not Yet Known

Filed : Herewith

For : VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

30 Rockefeller Plaza
20th Floor
New York, New York 10112
October 21, 2011

**By EFS**
Commissioner for Patents
Alexandria, VA 22313-1450

Sir:

**PRELIMINARY AMENDMENT AND INFORMATION DISCLOSURE STATEMENT**

Applicants request that the following amendment be made in the above-identified application:


IPR2025-01353; IPR2025-01355
CSHL EX2002

**Amendments to the Claims:**

Pursuant to 37 C.F.R. § 1.121(c), this listing of claims will replace all prior versions, and listings, of claims in the application:

1. (Original) A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

2. (Original) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome by assigning as the copy number

of the region the highest count obtained in step (f) for the locations within the region.

3. (Original) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Original) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Original) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is

comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Currently Amended) The method of ~~any one of claims 4-5~~ claim 4, wherein the second region of the genome comprises a centromere.

7. (Original) The method of claim 1, further comprising summing the counts obtained in step (f) for locations which comprise a region of the genome, and comparing the sum to a sum obtained from a reference sample for the same region of the genome, thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Original) A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. (Cancelled)

11. (Currently Amended) The method of ~~any one of claims 1-7~~ claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Currently Amended) The method of ~~any one of claims 1-7~~ claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands

added in step (a) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Currently Amended) The method of ~~any one of claims 9-12~~ claim 12, wherein adding a polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Currently Amended) The method of ~~any one of claims 1-23~~ claim 1, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or prior to sequencing.

25. (Cancelled)

26. (Currently Amended) The method of ~~any one of claims 1-25~~ claim 1, wherein the tag sequences further comprise a sample tag.

27. (Original) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to sequencing.

28. (Original) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to sample tag.

29-30. (Cancelled)

31. (Currently Amended) The method of ~~any one of claims 1-30~~ claim 1, wherein tagged nucleic acid molecules are generated from a single cell.

32-40. (Cancelled)

41. (Original) A composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of genomic material;

b) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

c) subjecting the zero-th order derivative strands of step (b) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (b) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

d) subjecting the product of step (c) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (b) under conditions which promote the formation of only one complement, wherein the primers of step (d) have different nucleotide sequences than the primers of step (c), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (c) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

42-43. (Cancelled)

44. (Currently Amended) The composition of matter of ~~any one of claims 40-41 or 43~~ claim 41, wherein adding a polynucleotide tail comprises the use of a terminal transferase.

45-50. (Cancelled)

51. (Currently Amended) The composition of matter of ~~any one of claims 40-50~~ claim 41, wherein the tag sequences further comprise a sample tag.

52. (Cancelled)

53. (Currently Amended) The composition of matter of ~~any one of claims 40-52~~ claim 41, wherein the tagged nucleic acid molecules are generated from a single cell.

54-67. (Cancelled)

**REMARKS**

Claims 1-67 are present in the application as filed. By this Preliminary Amendment, Applicants have cancelled claims 8, 10, 14-23, 25, 29-30, 32-40, 42-43, 45-50, 52, and 54-67 solely to reduce filing fees and without disclaimer or prejudice to Applicants' right to pursue the subject matter of these claims in the future. In addition, Applicants have amended claims 6, 11, 12, 13, 24, 26, 31, 44, 51, and 53 to remove multiple dependencies solely to reduce filing fees. Accordingly, claims 1-7, 9, 11-13, 24, 26-28, 31, 41, 44, 51, and 53 will be pending upon entry of this Preliminary Amendment.

**INFORMATION DISCLOSURE STATEMENT**

In accordance with their duty of disclosure under 37 C.F.R. §1.56, Applicants would like to direct the Examiner's attention to the following documents which are listed on Form PTO-1449 (Substitute), attached hereto as **Exhibit A,** and are also listed below.

This Information Disclosure Statement is being submitted pursuant to 37 C.F.R. §1.97(b)(3) before the mailing of a first Office Action on the merits. Thus, this Information Disclosure Statement should be entered and considered.

1. U.S. Patent No. 5,639,603, issued June 17, 1997 (Dower et al.);

2. U.S. Patent No. 7,537,897, issued May 26, 2009 (Brenner et al.).

3. U.S. Patent No. 7,622,281, issued November 24, 2009 (Ronaghi et al.);

4. U.S. Patent Application Publication No. 2006/0073506, published April 6, 2006 (Christians et al.);

5. Casbon J.A., Osborne R.J., Brenner S., Lichtenstein C.P. A method for counting PCT template molecules with application to next-generation sequencing. Nucleic Acids Research Advance Access, published April 13, 2011. **(Exhibit 1)**;

6. Cheung V.G., Nelson S.F. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. PNAS, vol. 93:14676-14679, 1996. **(Exhibit 2)**;

7. Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," Science, vol. 323 pp. 133-138 (2009); **(Exhibit 3)**;

8. Supporting Online Material for Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," accessible at www.sciencemag.org/cgi/content/full/1162986/DC1, published November 20, 2008; **(Exhibit 4)**;

9. Fu G.K., Hu J., Wang P., Fodor S.P.A. Counting individual DNA molecules by the stochastic attachment of diverse labels. PNAS, vol. 108:9026-9031, 2011. **(Exhibit 5)**;

10. Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S.. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135. **(Exhibit 6)**;

11. McCloskey M.L., Stoger, R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767. **(Exhibit 7)**;

12. Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130. **(Exhibit 8)**;

13. Paunio T., Reima I., Syvanen A. Preimplanation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. Clinical Chemistry 42:9:1382-1390 (1996). **(Exhibit 9)**; and

14. Technology Backgrounder: Single Molecule Real Time (SMRT™) DNA Sequencing. Pacific Biosciences 2009. **(Exhibit 10)**;

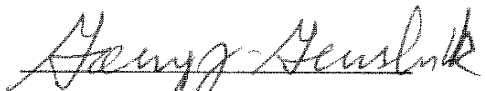
Applicants request that the Examiner review the references and make them of record in the subject application.

If a telephone interview would be of assistance in advancing prosecution of the subject application, Applicants' undersigned

attorney invites the Examiner to telephone him at the number provided below.

No fee, other than the total filing fee of $625.00, is deemed necessary in connection with the filing of this Preliminary Amendment. If any additional fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,



Gary J. Gershik
Registration No. 39,992
Attorney for Applicants
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400


IPR2025-01353; IPR2025-01355
CSHL EX2002

# SCORE Placeholder Sheet for IFW Content

Application Number: 13278333 Document Date: 10/21/2011

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

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 developed by SIRA.

At the time of document entry (noted above):

- Examiners may access SCORE content via the eDAN interface using the Supplemental Content tab.
- Other USPTO employees can bookmark the current SCORE URL (http://es/ScoreAccessWeb/).
- External customers may access SCORE content via the Public and Private PAIR interfaces using the Supplemental Content tab.

Form Revision Date: May 1, 2009



IPR2025-01353; IPR2025-01355
CSHL EX2002

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<br>Substitute for Form PTO-875 | Application or Docket Number<br>13/278,333 | Filing Date<br>10/21/2011 | ☐ 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

| 10/21/2011 | (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)) | * 20 | Minus | ** 20 | = 0 | X $30 = | 0 |
| Independent (37 CFR 1.16(h)) | * 3 | Minus | ***3 | = 0 | X $125 = | 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.

LDRC
/ANDREW JAMES JR/

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.*


IPR2025-01353; IPR2025-01355
CSHL EX2002

# MULTIPLE DEPENDENT CLAIM FEE CALCULATION SHEET

Substitute for Form PTO-1360
(For use with Form PTO/SB/06)

| Application Number | Filing Date |
|---|---|
| 13278333 | |

Applicant(s) James Hicks

* May be used for additional claims or amendments

| CLAIMS | AS FILED | | AFTER FIRST AMENDMENT | | AFTER SECOND AMENDMENT | |
|---|---|---|---|---|---|---|
| | Indep | Depend | Indep | Depend | Indep | Depend |
| 1 | 1 | | | | | |
| 2 | | 1 | | | | |
| 3 | | 1 | | | | |
| 4 | | 1 | | | | |
| 5 | | 1 | | | | |
| 6 | | 1 | | | | |
| 7 | | 1 | | | | |
| 8 | --- | --- | | | | |
| 9 | 1 | | | | | |
| 10 | --- | --- | | | | |
| 11 | | 1 | | | | |
| 12 | | 1 | | | | |
| 13 | | 1 | | | | |
| 14 | --- | --- | | | | |
| 15 | --- | --- | | | | |
| 16 | --- | --- | | | | |
| 17 | --- | --- | | | | |
| 18 | --- | --- | | | | |
| 19 | --- | --- | | | | |
| 20 | --- | --- | | | | |
| 21 | --- | --- | | | | |
| 22 | --- | --- | | | | |
| 23 | --- | --- | | | | |
| 24 | | 1 | | | | |
| 25 | --- | --- | | | | |
| 26 | | 1 | | | | |
| 27 | | 1 | | | | |
| 28 | | 1 | | | | |
| 29 | --- | --- | | | | |
| 30 | --- | --- | | | | |
| 31 | | 1 | | | | |
| 32 | --- | --- | | | | |
| 33 | --- | --- | | | | |
| 34 | --- | --- | | | | |
| 35 | --- | --- | | | | |
| 36 | --- | --- | | | | |
| 37 | --- | --- | | | | |
| 38 | --- | --- | | | | |
| 39 | --- | --- | | | | |
| 40 | --- | --- | | | | |
| 41 | 1 | | | | | |
| 42 | --- | --- | | | | |
| 43 | --- | --- | | | | |
| 44 | | 1 | | | | |
| 45 | --- | --- | | | | |
| 46 | --- | --- | | | | |
| 47 | --- | --- | | | | |
| 48 | --- | --- | | | | |
| 49 | --- | --- | | | | |
| 50 | --- | --- | | | | |
| Total Indep | 3 | | 0 | | 0 | |
| Total Depend | 17 | ↵ | 0 | ↵ | 0 | ↵ |
| Total Claims | 20 | | 0 | | 0 | |

| | * | | * | | * | |
|---|---|---|---|---|---|---|
| | Indep | Depend | Indep | Depend | Indep | Depend |
| 51 | | 1 | | | | |
| 52 | --- | --- | | | | |
| 53 | | 1 | | | | |
| 54 | --- | --- | | | | |
| 55 | --- | --- | | | | |
| 56 | --- | --- | | | | |
| 57 | --- | --- | | | | |
| 58 | --- | --- | | | | |
| 59 | --- | --- | | | | |
| 60 | --- | --- | | | | |
| 61 | --- | --- | | | | |
| 62 | --- | --- | | | | |
| 63 | --- | --- | | | | |
| 64 | --- | --- | | | | |
| 65 | --- | --- | | | | |
| 66 | --- | --- | | | | |
| 67 | --- | --- | | | | |
| 68 | | | | | | |
| 69 | | | | | | |
| 70 | | | | | | |
| 71 | | | | | | |
| 72 | | | | | | |
| 73 | | | | | | |
| 74 | | | | | | |
| 75 | | | | | | |
| 76 | | | | | | |
| 77 | | | | | | |
| 78 | | | | | | |
| 79 | | | | | | |
| 80 | | | | | | |
| 81 | | | | | | |
| 82 | | | | | | |
| 83 | | | | | | |
| 84 | | | | | | |
| 85 | | | | | | |
| 86 | | | | | | |
| 87 | | | | | | |
| 88 | | | | | | |
| 89 | | | | | | |
| 90 | | | | | | |
| 91 | | | | | | |
| 92 | | | | | | |
| 93 | | | | | | |
| 94 | | | | | | |
| 95 | | | | | | |
| 96 | | | | | | |
| 97 | | | | | | |
| 98 | | | | | | |
| 99 | | | | | | |
| 100 | | | | | | |

| PATENT APPLICATION FEE DETERMINATION RECORD<br>Substitute for Form PTO-875 | Application or Docket Number<br>13/278,333 |
|---|---|

## APPLICATION AS FILED - PART I

| | (Column 1) | (Column 2) | SMALL ENTITY | | OR | OTHER THAN SMALL ENTITY | |
|---|---|---|---|---|---|---|---|
| FOR | NUMBER FILED | NUMBER EXTRA | RATE($) | FEE($) | | RATE($) | FEE($) |
| BASIC FEE (37 CFR 1.16(a), (b), or (c)) | N/A | N/A | N/A | 95 | | N/A | |
| SEARCH FEE (37 CFR 1.16(k), (i), or (m)) | N/A | N/A | N/A | 310 | | N/A | |
| EXAMINATION FEE (37 CFR 1.16(o), (p), or (q)) | N/A | N/A | N/A | 125 | | N/A | |
| TOTAL CLAIMS (37 CFR 1.16(i)) | 20 minus 20 = | * | x 30 = | 0.00 | OR | | |
| INDEPENDENT CLAIMS (37 CFR 1.16(h)) | 3 minus 3 = | * | x 125 = | 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 | 530 | | TOTAL | |

## APPLICATION AS AMENDED - PART II

| | | (Column 1) | | (Column 2) | (Column 3) | SMALL ENTITY | | OR | OTHER THAN SMALL ENTITY | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | RATE($) | ADDITIONAL FEE($) | | RATE($) | ADDITIONAL FEE($) |
| AMENDMENT A | Total (37 CFR 1.16(i)) | * | Minus | ** | = | x = | | OR | x = | |
| | Independent (37 CFR 1.16(h)) | * | Minus | *** | = | x = | | OR | x = | |
| | Application Size Fee (37 CFR 1.16(s)) | | | | | | | | | |
| | FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | OR | | |
| | | | | | | TOTAL ADD'L FEE | | OR | TOTAL ADD'L FEE | |

| | | (Column 1) | | (Column 2) | (Column 3) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | RATE($) | ADDITIONAL FEE($) | | RATE($) | ADDITIONAL FEE($) |
| AMENDMENT B | Total (37 CFR 1.16(i)) | * | Minus | ** | = | x = | | OR | x = | |
| | Independent (37 CFR 1.16(h)) | * | Minus | *** | = | x = | | OR | x = | |
| | Application Size Fee (37 CFR 1.16(s)) | | | | | | | | | |
| | FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | OR | | |
| | | | | | | TOTAL ADD'L FEE | | OR | TOTAL ADD'L FEE | |

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.


IPR2025-01353; IPR2025-01355
CSHL EX2002

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 |
|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/ACK |

**CONFIRMATION NO. 4469**

23432
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

**FORMALITIES LETTER**

*OC000000050897307*

Date Mailed: 11/15/2011

# 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 oath or declaration is unsigned.

The application is informal since it does not comply with the regulations for the reason(s) indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- A replacement abstract not exceeding 150 words in length and commencing on a separate sheet in compliance with 37 CFR 1.72(b) and 37 CFR 1.121 is required.
- 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.

**To Download Patentin Software, visit http://www.uspto.gov/web/patents/software.htm**

**For questions regarding compliance to these requirements, please contact:**

- **For Rules Interpretation, call (571) 272-0623**
- **For Patentin Software Program Help, call Patent EBC at 1-866-217-9197 between the hours of 6 a.m. and 12 midnight, Monday through Friday, EST.**
- **Send e-mail correspondence for Patentin Software Program Help to ebc@uspto.gov**

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.

The applicant needs to satisfy supplemental fees problems indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- A surcharge (for late submission of filing fee, search fee, examination fee or oath or declaration) as set forth in 37 CFR 1.16(f) of $**65** for a small entity in compliance with 37 CFR 1.27, must be submitted.

**SUMMARY OF FEES DUE:**

Total fee(s) required within **TWO MONTHS** from the date of this Notice is $**65** for a small entity

- $**65** Surcharge.

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.
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 |
|---|---|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | 1636 | 530 | 5981/81503-A/GJG/ACK | 20 | 3 |

**CONFIRMATION NO. 4469**

23432
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

**FILING RECEIPT**

*OC00000050897306*

Date Mailed: 11/15/2011

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**

**Applicant(s)**

James Hicks, Huntington, NY;
Nicholas Navin, Houston, TX;
Jennifer Troge, Oyster Bay, NY;
Zihua Wang, Huntington, NY;
Michael Wigler, Cold Spring Harbor, NY;

**Power of Attorney:** None

**Domestic Priority data as claimed by applicant**

This appln claims benefit of 61/510,579 07/22/2011
and claims benefit of 61/406,067 10/22/2010

**Foreign Applications** (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.)

**Projected Publication Date:** To Be Determined - pending completion of Missing Parts

**Non-Publication Request:** No

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

**Title**

VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

**Preliminary Class**

435

## 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 (1-866-999-4158).

## 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 USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage, facilitate, and accelerate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

DocCode – SEQ.TXT

# SCORE Placeholder Sheet for IFW Content

Application Number: 13278333 Document Date: 01/16/2012

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.

- Sequence Listing

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 developed by SIRA.

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://es/ScoreAccessWeb/).
- External customers may access SCORE content via the Public and Private PAIR interfaces.

Form Revision Date: February 8, 2006


IPR2025-01353; IPR2025-01355
CSHL EX2002

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicant(s) : James Hicks et al.

U.S. Serial No. : 13/278,333

Filing Date : October 21, 2011

Confirmation No. : 4469

For : VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

30 Rockefeller Plaza, 20th Floor
New York, New York 10112
January 16, 2012

**BY EFS**
Commissioner for Patents
Alexandria, VA

**AMENDMENT IN RESPONSE TO NOVEMBER 15, 2011**
**NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION**

This Amendment is submitted in response to the November 15, 2011 Notice to File Missing Parts of Nonprovisional Application issued in connection with the above-identified application. The November 15, 2011 Notice, a copy of which is attached hereto as **Exhibit A**, indicates that a response is due two (2) months from the mailing date of the Notice, i.e., by Tuesday, January 17, 2012 (since January 15, 2012 falls on Sunday and January 16, 2012 falls on a Federal holiday within the District of Columbia). Accordingly, this Amendment is being timely filed.

Please amend the subject application as follows:

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

A **Sequence Listing** is referred to on page 4 of this paper and filed herewith as a **text file.**

A **Statement of Compliance under 37 C.F.R. § 1.825(a)** is referred to on page 6 of this paper and attached hereto as **Exhibit B.**

**Remarks** begin on page 5 of this paper.

**AMENDMENTS TO THE SPECIFICATION**

Please insert the following new section heading and paragraph on page 1, immediately after line 18 and immediately before the section titled, "Background of the Invention":

**Reference to a Sequence Listing**

This application incorporates-by-reference nucleotide and/or amino acid sequences which are present in the file named "120116_5981_81503_A_Sequence_Listing_GC.txt," which is 2.48 kilobytes in size, and which was created January 13, 2012 in the IBM-PC machine format, having an operating system compatibility with MS-Windows, which is contained in the text file filed January 16, 2012 as part of this application.

Please replace the paragraph on page 87, which sets forth an Abstract of the Disclosure, with the following amended paragraph:

A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising obtaining segments of the genomic material, tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material and a tag, subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification ~~by polymerase chain reaction (PCR)~~, generating tag associated sequence reads by sequencing the product of the PCR reaction, assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome, and counting the number of tagged nucleic acid molecules ~~having a different tag that have been~~ assigned to the same location on the genome having a different tag, thereby obtaining genomic copy number information unaffected by amplification distortion.

**SEQUENCE LISTING**

Please amend the subject application to include the Sequence Listing filed herewith as a text file named "120116_5981_81503_A_Sequence_Listing_GC.txt."

**REMARKS**

***Amendments to the Specification***

By this Amendment, applicants have amended the specification to insert a section heading and paragraph indicating the name, size, date of creation, and date of filing of the text file named "120116_5981_81503_A_Sequence_Listing_GC.txt" which corresponds to the Sequence Listing.

The November 15, 2011 Notice indicates that a replacement abstract not exceeding 150 words in length and commencing on a separate sheet in compliance with 37 C.F.R. §§ 1.72(b) and 1.121 is required. In response, by this Amendment, applicants have amended the Abstract of the Disclosure, which commences on page 87 of the subject application. Applicants maintain that the Abstract as amended does not exceed 150 words in length and commences on a separate sheet.

***Sequence Listing***

The November 15, 2011 Notice indicates that the subject application does not comply with the requirements of 37 C.F.R. §§ 1.821 and/or 1.825. Specifically, the Notice indicates that applicants must submit:

(1) an initial computer readable form (CRF) copy and initial paper or compact disc copy of the Sequence Listing;
(2) an amendment specifically directing its entry into the application; and
(3) a statement that the content of the sequence listing information recorded in CRF is identical to the written (on paper or compact disc) sequence listing and, where applicable, includes no new matter.

With this Amendment, applicants have:

(1) submitted a text file containing the Sequence Listing named "120116_5981_81503_A_Sequence_Listing_GC.txt," which will serve as both the CRF and paper copy of the

Sequence Listing;

(2) amended the subject application to include the accompanying text file containing the Sequence Listing; and

(3) attached hereto as **Exhibit B** a Statement of Compliance under 37 C.F.R. § 1.825(a) certifying that the Sequence Listing submitted herewith as a text file does not introduce any new matter into the subject application.

The subject matter in the Sequence Listing submitted herewith corresponds to the subject matter in Figures 8 and 9 of the subject application as filed.

In accordance with the guidance set forth in section I1 of "Legal Framework for EFS-Web (06April11)" published by the U.S. Patent and Trademark Office at www.uspto.gov/patents/process/file/efs/guidance/New_legal_framework.jsp, applicants are not submitting herewith a statement under 37 C.F.R. § 1.821(f), since the text file submitted herewith will serve as both the "paper copy" and the "computer readable form" of the Sequence Listing.

***Declaration***

The November 15, 2011 Notice indicates that a signed oath or declaration of the inventors must be submitted. In response, applicants attach hereto as **Exhibit C** a signed Declaration and Power of Attorney pursuant to 37 C.F.R. § 1.53(f). In compliance with 37 C.F.R. § 1.63, the Declaration refers to the serial number and filing date of the subject application.

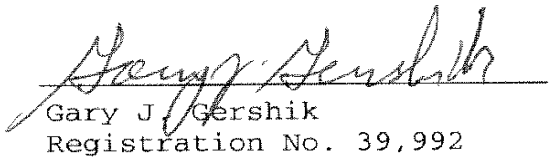
The November 15, 2011 Notice also indicates that a $65.00 surcharge for late filing of the inventors' Declaration is required for a small entity. Accordingly, authorization is hereby given to charge the amount of SIXTY-FIVE DOLLARS ($65.00) to Deposit Account No. 03-3125.

Applicants maintain that the amendments presented hereinabove raise no issue of new matter and respectfully request that this Amendment be entered.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicants' undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee, other than the above-mentioned $65.00 surcharge for late filing of the Declaration, is deemed necessary in connection with the filing of this Amendment. However, if any additional fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

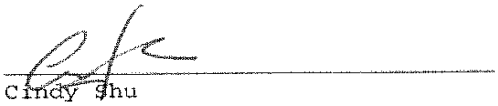
Respectfully submitted,



Gary J. Gershik
Registration No. 39,992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on January 16, 2012.



Cindy Shu



IPR2025-01353; IPR2025-01355
CSHL EX2002

5981/81503-A GJG/REB

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 |
|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/ACK |

CONFIRMATION NO: 4469

23432
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

RECEIVED
COOPER & DUNHAM
NOV 22 2011
DOCKET CLERK

FORMALITIES LETTER

*OC000000050897307*

Date Mailed: 11/15/2011

MISSING parts DUE 1-15
3MO 2-15
4MO 3-15
5MO 4-15
6MO 5-15
TRK 6-15

## 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 oath or declaration is unsigned.

The application is informal since it does not comply with the regulations for the reason(s) indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- A replacement abstract not exceeding 150 words in length and commencing on a separate sheet in compliance with 37 CFR 1.72(b) and 37 CFR 1.121 is required.
- 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.

**To Download Patentin Software, visit http://www.uspto.gov/web/patents/software.htm**



Applicant(s): James Hicks et al.
U.S. Serial No.: 13/278,333
Filing Date: October 21, 2011
**Exhibit A**



IPR2025-01353; IPR2025-01355
CSHL EX2002

**For questions regarding compliance to these requirements, please contact:**

- **For Rules Interpretation, call (571) 272-0623**
- **For Patentin Software Program Help, call Patent EBC at 1-866-217-9197 between the hours of 6 a.m. and 12 midnight, Monday through Friday, EST.**
- **Send e-mail correspondence for Patentin Software Program Help to ebc@uspto.gov**

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.

The applicant needs to satisfy supplemental fees problems indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- A surcharge (for late submission of filing fee, search fee, examination fee or oath or declaration) as set forth in 37 CFR 1.16(f) of **$65** for a small entity in compliance with 37 CFR 1.27, must be submitted.

**SUMMARY OF FEES DUE:**

Total fee(s) required within **TWO MONTHS** from the date of this Notice is **$65** for a small entity

- **$65** Surcharge.

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.
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

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicant(s) : James Hicks et al.

U.S. Serial No. : 13/278,333

Filing Date : October 21, 2011

Confirmation No. : 4469

For : VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

30 Rockefeller Plaza, 20th Floor
New York, New York 10112
January 16, 2012

**BY EFS**
Commissioner for Patents
Alexandria, VA

**STATEMENT OF COMPLIANCE UNDER 37 C.F.R § 1.825(a) IN CONNECTION WITH ABOVE-IDENTIFIED APPLICATION**

In accordance with 37 C.F.R. § 1.825(a), I hereby certify that the content of the Sequence Listing submitted herewith as a text file named "120116_5981_81503_A_Sequence_Listing_GC.txt" does not introduce any new matter into the subject application.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment or both, under Section 1001 of Title 18 of the United States Code and that any such willful false statement may jeopardize the validity of the application or any patent issued therefrom.

Respectfully submitted,

Regina Cheng
c/o Cooper & Dunham LLP
30 Rockefeller Plaza, 20th Floor
New York, New York 10112
(212) 278-0400

*Docket Number:* 5981/81503-A/GJG/ACK

## *DECLARATION AND POWER OF ATTORNEY*

*As a below-named inventor, I hereby declare that:*

*My residence, post office address, and citizenship are as stated below next to my name.*

*I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:*

VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

*the specification of which:*
*(check one)*

____ *is attached hereto.*

_X_ *was filed on* October 21, 2011 *as*

*Application Serial No.* 13/278,333

*and was amended* October 21, 2011
*(if applicable)*

*I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.*

*I acknowledge the duty to disclose to the U.S. Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.*

*I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International Application which designated at least one country other than the United States, listed below. I have also identified below any foreign application for patent or inventor's certificate, or PCT International Application having a filing date before that of the earliest application from which priority is claimed:*

| *Prior Foreign Application(s)* | | | *Priority Claimed* | |
|---|---|---|---|---|
| *Number* | *Country* | *Filing Date* | *Yes* | *No* |
| N/A | | | — | — |
| | | | — | — |
| | | | — | — |
| | | | — | — |
| | | | — | — |

*JPW Rev. 8/13/08*


IPR2025-01353; IPR2025-01355
CSHL EX2002

*I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below:*

| *Provisional Application No.* | *Filing Date* | *Status* |
|---|---|---|
| 61/510,579 | July 22, 2011 | Pending |
| 61/406,067 | October 22, 2010 | Pending |
| | | |
| | | |
| | | |

*I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States Application(s), or Section 365(c) of any PCT International Application(s) designating the United States listed below. Insofar as this application discloses and claims subject matter in addition to that disclosed in any such prior Application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56, which became available between the filing date(s) of such prior Application(s) and the national or PCT international filing date of this application:*

| *Application Serial No.* | *Filing Date* | *Status* |
|---|---|---|
| N/A | | |
| | | |
| | | |
| | | |
| | | |
| | | |

*And I hereby appoint*

Gary J. Gershik (Reg. No. 39,992); John P. White (Reg. No. 28,678); Christopher C. Dunham (Reg. No. 22,031); Norman H. Zivin (Reg. No. 25,385); William E. Pelton (Reg. No. 25,702); Robert D. Katz (Reg. No. 30,141); and Paul Teng (Reg. No. 40,837).

*and each of them, all c/o Cooper & Dunham LLP, 30 Rockefeller Plaza, 20th Floor, New York, New York 10112, my attorneys, each with full power of substitution and revocation, to prosecute this application, to make alterations and amendments therein, to receive the patent, to transact all business in the Patent and Trademark Office connected therewith and to file any International Applications which are based thereon under the provisions of the Patent Cooperation Treaty.*

*Please address all communications, and direct all telephone calls, regarding this application to:*

Gary J. Gershik, Esq. *Reg.No.* 39,992
*Cooper & Dunham, LLP (Customer Number 23432)*
*30 Rockefeller Plaza*
*20th Floor*
*New York, New York 10112*
*Tel. (212) 278-0400*

*I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.*

*Full name of sole or first joint inventor* James Hicks
*Inventor's signature* _____ *Date of signature* 11-16-11
*Citizenship* US
*Residence* 3 Great Meadow Road, Locust Valley, NY 11560
*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Nicholas Navin
*Inventor's signature* _____ *Date of signature* _____
*Citizenship* US
*Residence* 1621 Wichita St., Houston, TX 77004
*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Jennifer Troge
*Inventor's signature* _____ *Date of signature* 11.16.11
*Citizenship* US
*Residence* 100 Simcoe St., Oyster Bay, NY 11771
*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Zihua Wang

*Inventor's signature* _____ *Date of signature* 11/16/11

*Citizenship* China

*Residence* 13 Meadowlawn St., Huntington, NY 11743

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Michael Wigler

*Inventor's signature* _____ *Date of signature* 11/16/2011

*Citizenship* US

*Residence* 52 Snake Hill Road, Cold Spring Harbor, NY 11724

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* _____

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* _____

*Residence* _____

*Post Office Address* _____

*Full name of additional joint inventor (if any)* _____

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* _____

*Residence* _____

*Post Office Address* _____

*Docket Number:* 5981/81503-A/GJG/ACK

## *DECLARATION AND POWER OF ATTORNEY*

*As a below-named inventor, I hereby declare that:*

*My residence, post office address, and citizenship are as stated below next to my name.*

*I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:*

VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

*the specification of which:*
*(check one)*

____ *is attached hereto.*

_X_ *was filed on* October 21, 2011 *as*

*Application Serial No.* 13/278,333

*and was amended* October 21, 2011
*(if applicable)*

*I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.*

*I acknowledge the duty to disclose to the U.S. Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.*

*I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International Application which designated at least one country other than the United States, listed below. I have also identified below any foreign application for patent or inventor's certificate, or PCT International Application having a filing date before that of the earliest application from which priority is claimed:*

| *Prior Foreign Application(s)* | | | *Priority Claimed* | |
|---|---|---|---|---|
| *Number* | *Country* | *Filing Date* | *Yes* | *No* |
| N/A | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

*JPW Rev. 8/13/08*

*I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below:*

| *Provisional Application No.* | *Filing Date* | *Status* |
|---|---|---|
| 61/510,579 | July 22, 2011 | Pending |
| 61/406,067 | October 22, 2010 | Pending |
| | | |
| | | |
| | | |

*I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States Application(s), or Section 365(c) of any PCT International Application(s) designating the United States listed below. Insofar as this application discloses and claims subject matter in addition to that disclosed in any such prior Application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56, which became available between the filing date(s) of such prior Application(s) and the national or PCT international filing date of this application:*

| *Application Serial No.* | *Filing Date* | *Status* |
|---|---|---|
| N/A | | |
| | | |
| | | |
| | | |
| | | |
| | | |

*And I hereby appoint*

Gary J. Gershik (Reg. No. 39,992); John P. White (Reg. No. 28,678); Christopher C. Dunham (Reg. No. 22,031); Norman H. Zivin (Reg. No. 25,385); William E. Pelton (Reg. No. 25,702); Robert D. Katz (Reg. No. 30,141); and Paul Teng (Reg. No. 40,837).

*and each of them, all c/o Cooper & Dunham LLP, 30 Rockefeller Plaza, 20th Floor, New York, New York 10112, my attorneys, each with full power of substitution and revocation, to prosecute this application, to make alterations and amendments therein, to receive the patent, to transact all business in the Patent and Trademark Office connected therewith and to file any International Applications which are based thereon under the provisions of the Patent Cooperation Treaty.*

*Please address all communications, and direct all telephone calls, regarding this application to:*

Gary J. Gershik, Esq. *Reg.No.* 39,992
*Cooper & Dunham, LLP (Customer Number 23432)*
*30 Rockefeller Plaza*
*20th Floor*
*New York, New York 10112*
*Tel. (212) 278-0400*

*I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.*

*Full name of sole or first joint inventor* James Hicks

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* US

*Residence* 3 Great Meadow Road, Locust Valley, NY 11560

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Nicholas Navin

*Inventor's signature* Nicholas E Navin *Date of signature* 11-14-2011

*Citizenship* US

*Residence* 1621 Wichita St., Houston, TX 77004

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Jennifer Troge

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* US

*Residence* 100 Simcoe St., Oyster Bay, NY 11771

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Zihua Wang

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* China

*Residence* 13 Meadowlawn St., Huntington, NY 11743

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* Michael Wigler

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* US

*Residence* 52 Snake Hill Road, Cold Spring Harbor, NY 11724

*Post Office Address* Same as above

*Full name of additional joint inventor (if any)* _____

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* _____

*Residence* _____

*Post Office Address* _____

*Full name of additional joint inventor (if any)* _____

*Inventor's signature* _____ *Date of signature* _____

*Citizenship* _____

*Residence* _____

*Post Office Address* _____

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 13278333 |
| **Filing Date:** | 21-Oct-2011 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | John P. White/Cindy Shu |
| **Attorney Docket Number:** | 5981/81503-A/GJG/ACK |

Filed as Small Entity

## Utility under 35 USC 111(a) Filing Fees

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| Late filing fee for oath or declaration | 2051 | 1 | 65 | 65 |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| **Extension-of-Time:** | | | | |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Miscellaneous:** | | | | |
| | | | **Total in USD ($)** | **65** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 11842890 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Cindy Shu |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 5981/81503-A/GJG/ACK |
| **Receipt Date:** | 16-JAN-2012 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 15:58:25 |
| **Application Type:** | Utility under 35 USC 111(a) |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | Deposit Account |
| Payment was successfully received in RAM | $65 |
| RAM confirmation Number | 11054 |
| Deposit Account | 033125 |
| Authorized User | |

## File Listing:

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



IPR2025-01353, IPR2025-01355
CSHL EX2002

| | | | | | |
|---|---|---|---|---|---|
| 1 | | 120116_5981_81503_A_Amend_in_Resp_to_Notice_to_File_Missing_Parts_GC.pdf | 1074302<br>172fec166d23b15bd555c4b2da8cea6a379e8291 | yes | 18 |

| Multipart Description/PDF files in .zip description | | |
|---|---|---|
| **Document Description** | **Start** | **End** |
| Applicant Response to Pre-Exam Formalities Notice | 1 | 1 |
| Specification | 2 | 2 |
| Abstract | 3 | 3 |
| Applicant Arguments/Remarks Made in an Amendment | 4 | 7 |
| Applicant Response to Pre-Exam Formalities Notice | 8 | 9 |
| CRF Statement Paper and CRF are the same | 10 | 10 |
| Oath or Declaration filed | 11 | 18 |

| | | | | | |
|---|---|---|---|---|---|
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Sequence Listing (Text File) | 120116_5981_81503_A_Sequence_Listing_GC.txt | 2546 | no | 0 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Fee Worksheet (SB06) | fee-info.pdf | 30513<br>0c3b31fd550a7f4d6f36ba383e683732456bbaee | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 1107361 | | |

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

**New Applications Under 35 U.S.C. 111**

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

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

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

**New International Application Filed with the USPTO as a Receiving Office**

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

```
============================================================
Sequence Listing was accepted.
If you need help call the Patent Electronic Business Center at (866)
217-9197 (toll free).
Reviewer: Durreshwar Anjum
Timestamp: [year=2012; month=1; day=23; hr=15; min=31; sec=33; ms=760; ]
============================================================
```

**Validated By CRFValidator v 1.0.3**

Application No: 13278333 Version No: 1.0

**Input Set:**

**Output Set:**

**Started:** 2012-01-16 15:59:12.676
**Finished:** 2012-01-16 15:59:13.316
**Elapsed:** 0 hr(s) 0 min(s) 0 sec(s) 640 ms
**Total Warnings:** 8
**Total Errors:** 0
**No. of SeqIDs Defined:** 8
**Actual SeqID Count:** 8

| Error code | Error Description |
|---|---|
| W 213 | Artificial or Unknown found in <213> in SEQ ID (1) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (2) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (3) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (4) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (5) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (6) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (7) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (8) |


IPR2025-01353; IPR2025-01355
CSHL EX2002

SEQUENCE LISTING

<110> Hicks, James
Navin, Nicholas
Troge, Jennifer
Wang, Zihua
Wigler, Michael

<120> VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

<130> 5981/81503-A/GJG/REB

<140> 13278333
<141> 2012-01-16

<150> 61/510,579
<151> 2011-07-22

<150> 61/406,067
<151> 2010-10-22

<160> 8

<170> PatentIn version 3.3

<210> 1
<211> 15
<212> DNA
<213> Artificial

<220>
<223> Portion of 3' end of nucleotide sequence after Terminal Transferase step.

<400> 1
catgaaaaaa aaaaa 15

<210> 2
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Nucleotide sequence comprising a sequence complementary to portion of 3' end of a nucleotide sequence, and a varietal tag.

<220>
<221> misc_feature
<222> (1)..(10)
<223> n is a, t, g or c

<400> 2
nnnnnnnnnn ttttttttt catg 24

<210> 3
<211> 11
<212> DNA
<213> Artificial

<220>
<223> Portion of 3' end of nuceotide sequence after Pool, Cleanup, Melt, Terminal Transferase.

<400> 3
ccccccccccc 11

<210> 4
<211> 10
<212> DNA
<213> Artificial

<220>
<223> Portion of a primer complementary to portion of 3' end of nucleotide sequence.

<400> 4
gggggggggg 10

<210> 5
<211> 12
<212> DNA
<213> Artificial

<220>
<223> Portion of 3' end of nucleotide sequence after Terminal Transferase step.

<400> 5
catgaaaaaa aa 12

<210> 6
<211> 16
<212> DNA
<213> Artificial

<220>
<223> Portion of 3' end of nucleotide sequence after Terminal Transferase step.

<400> 6
catgaaaaaa aaaaaa 16

<210> 7
<211> 16
<212> DNA

<213> Artificial

<220>
<223> Portion of primer complementary to portion of 3' end of
nucleotide sequence.

<400> 7
tttttttttt ttcatg 16

<210> 8
<211> 14
<212> DNA
<213> Artificial

<220>
<223> Portion of 3' end of nucleotide segment after Terminal
Transferase step.

<400> 8
catgaaaaaa aaaa 14

| PATENT APPLICATION FEE DETERMINATION RECORD<br>Substitute for Form PTO-875 | Application or Docket Number<br>13/278,333 |
|---|---|

## APPLICATION AS FILED - PART I

| | (Column 1) | (Column 2) | SMALL ENTITY | | OR | OTHER THAN SMALL ENTITY | |
|---|---|---|---|---|---|---|---|
| FOR | NUMBER FILED | NUMBER EXTRA | RATE($) | FEE($) | | RATE($) | FEE($) |
| BASIC FEE (37 CFR 1.16(a), (b), or (c)) | N/A | N/A | N/A | 95 | | N/A | |
| SEARCH FEE (37 CFR 1.16(k), (i), or (m)) | N/A | N/A | N/A | 310 | | N/A | |
| EXAMINATION FEE (37 CFR 1.16(o), (p), or (q)) | N/A | N/A | N/A | 125 | | N/A | |
| TOTAL CLAIMS (37 CFR 1.16(i)) | 20 minus 20 = | * | x 30 = | 0.00 | OR | | |
| INDEPENDENT CLAIMS (37 CFR 1.16(h)) | 3 minus 3 = | * | x 125 = | 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 | 530 | | TOTAL | |

## APPLICATION AS AMENDED - PART II

| | | (Column 1) | | (Column 2) | (Column 3) | SMALL ENTITY | | OR | OTHER THAN SMALL ENTITY | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | RATE($) | ADDITIONAL FEE($) | | RATE($) | ADDITIONAL FEE($) |
| AMENDMENT A | Total (37 CFR 1.16(i)) | * | Minus | ** | = | x = | | OR | x = | |
| | Independent (37 CFR 1.16(h)) | * | Minus | *** | = | x = | | OR | x = | |
| | Application Size Fee (37 CFR 1.16(s)) | | | | | | | | | |
| | FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | OR | | |
| | | | | | | TOTAL ADD'L FEE | | OR | TOTAL ADD'L FEE | |

| | | (Column 1) | | (Column 2) | (Column 3) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | RATE($) | ADDITIONAL FEE($) | | RATE($) | ADDITIONAL FEE($) |
| AMENDMENT B | Total (37 CFR 1.16(i)) | * | Minus | ** | = | x = | | OR | x = | |
| | Independent (37 CFR 1.16(h)) | * | Minus | *** | = | x = | | OR | x = | |
| | Application Size Fee (37 CFR 1.16(s)) | | | | | | | | | |
| | FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | OR | | |
| | | | | | | TOTAL ADD'L FEE | | OR | TOTAL ADD'L FEE | |

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.


IPR2025-01353; IPR2025-01355
CSHL EX2002

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 |
|---|---|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | 1636 | 595 | 5981/81503-A/GJG/GC | 20 | 3 |

**CONFIRMATION NO. 4469**

**UPDATED FILING RECEIPT**

23432
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

*OC000000052170931*

Date Mailed: 01/27/2012

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**

**Applicant(s)**

James Hicks, Locust Valley, NY;
Nicholas Navin, Houston, TX;
Jennifer Troge, Oyster Bay, NY;
Zihua Wang, Huntington, NY;
Michael Wigler, Cold Spring Harbor, NY;

**Power of Attorney:**

Christopher Dunham--22031
Norman Zivin--25385
William Pelton--25702
John White--28678
Robert Katz--30141
Gary Gershik--39992
Paul Teng--40837

**Domestic Priority data as claimed by applicant**

This appln claims benefit of 61/510,579 07/22/2011
and claims benefit of 61/406,067 10/22/2010

**Foreign Applications** (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.)

**Projected Publication Date:** To Be Determined - pending completion of Security Review

**Non-Publication Request:** No

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

**Title**

VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

**Preliminary Class**

435

## 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 (1-866-999-4158).

## 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 USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage, facilitate, and accelerate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

**Form PTO-1449 (Substitute)** **U.S. Department of Commerce**
**Patent and Trademark Office**

**INFORMATION DISCLOSURE STATEMENT**
**(Use several sheets if necessary)**

| | |
|---|---|
| **Application Number** | 13/278,333 |
| **Filing Date** | October 21, 2011 |
| **First Named Inventor** | James Hicks |
| **Art Unit** | |
| **Examiner Name** | |
| **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| | 1 | 6,172,214 | 01-09-2001 | Brenner et al. |
| | 6 | 2006/0035258 | 02-16-2006 | Tadakamalla et al. |
| | 7 | 2007/0172873 | 07-26-2007 | Brenner et al. |
| | 8 | 2009/0137402 | 05-28-2009 | Wang et al. |
| | 9 | 2010/0203538 | 08-12-2010 | Dube et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |

**EXAMINER SIGNATURE** **DATE CONSIDERED**

***EXAMINER:** Initial if citation 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 of 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.


IPR2025-01353, IPR2025-01355
CSHL EX2002

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | **Application Number** | 13/278,333 |
| | **Filing Date** | October 21, 2011 |
| | **First Named Inventor** | James Hicks |
| | **Art Unit** | |
| | **Examiner Name** | |
| | **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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] |
|---|---|---|---|
| | 2 | Islam et al., "Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq" Genome Research (2001) 21:1160-1167 | |
| | 3 | Kivioja et al., "Counting absolute numbers of molecules using unique molecular identifiers" Nature Methods (2012) 9(1):72-74 and Online Methods (2 pages) | |
| | 4 | Shiroguchi et al., "Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes" PNAS (2012) 109(4):1347-1352 | |
| | 5 | International Search Report, issued in connection with PCT International Application No. PCT/US2011/057350 | |
| | 10 | Castle et al., "DNA copy number, including telomeres and mitochondria, assayed using next-generation sequencing" BMC Genomics (2010) 11(244):1-11 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

# PATENT COOPERATION TREATY

# PCT

## INTERNATIONAL SEARCH REPORT

**(PCT Article 18 and Rules 43 and 44)**

| Applicant's or agent's file reference | **FOR FURTHER ACTION** | see Form PCT/ISA/220 as well as, where applicable, item 5 below. |
|---|---|---|
| 81503-A-PCT/GJG/WS | | |
| International application No. | International filing date *(day/month/year)* | (Earliest) Priority Date *(day/month/year)* |
| **PCT/US2011/057350** | **21 OCTOBER 2011 (21.10.2011)** | 22 OCTOBER 2010 (22.10.2010) |

Applicant

**COLD SPRING HARBOR LABORATORY et al**

This International search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of **4** sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language**, the international search was carried out on the basis of:
      ☒ the international application in the language in which it was filed
      ☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))
   b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).
   c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.
2. ☒ **Certain claims were found unsearchable** (See Box No. II)
3. ☐ **Unity of invention is lacking** (See Box No. III)
4. With regard to the **title,**
   ☒ the text is approved as submitted by the applicant.
   ☐ the text has been established by this Authority to read as follows:
5. With regard to the **abstract,**
   ☒ the text is approved as submitted by the applicant.
   ☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.
6. With regard to the drawings,
   a. the figure of the **drawings** to be published with the abstract is Figure No. _____
      ☐ as suggested by the applicant.
      ☐ as selected by this Authority, because the applicant failed to suggest a figure.
      ☐ as selected by this Authority, because this figure better characterizes the invention.
   b. ☒ none of the figure is to be published with the abstract.

Form PCT/ISA/210 (first sheet) (July 2009)



IPR2025-01353; IPR2025-01355
CSHL EX2002
U.S. Serial No. 13/278,339
IDS - Exhibit 4

**INTERNATIONAL SEARCH REPORT**

International application No.
**PCT/US2011/057350**

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 15,27,28,30,32-35,38,39,52,54,55,64,65
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

The subject matter of said claims is so unclear and indefinite that no meaningful opinion could be made because these claims are drafted in reference to multiple dependent claims such as claims 11-14, 16-19, 21-26, 29, 31, 36, 37, 45-51, 53, 56, 60-63, 66 and 67 (see the category Box No. VII below).

3. ☒ Claims Nos.: 11-14,16-19,21-26,29,31,36,37,45-51,53,56,60-63,66,67
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required addtional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
☐ No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

International application No.
**PCT/US2011/057350**

**A. CLASSIFICATION OF SUBJECT MATTER**

***C12Q 1/68(2006.01)i, C12N 15/11(2006.01)i, C12Q 1/48(2006.01)i***

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

C12Q 1/68; C40B 20/00; C12P 1/34; C12P 19/34; G01N 33/50

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

NCBI PubMed, eKOMPASS (KIPO internal), 'genomic copy number, tagging, amplification, mapping, next-generation sequencing, etc.'

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | CASTLE, J. C. et al., 'DNA copy number, including telomeres and mitochondria,assayed using next-generation sequencing' BMC Genomics, 2010 Apr. 16, Vol. 11, No. 244, pp1-11 - see the whole document, especially Abstract; Application 1 in p2; Methods in p6 | 1-10,20,40-44 ,57-59 |
| A | US 2007/0172873 A1 (BRENNER, S. et al.) 26 Jul. 2007 - see the whole document, especially Abstract; Claims 1, 17-18 | 1-10,20,40-44 ,57-59 |
| A | US 2010/0203538 A1 (DUBE, S. et al.) 12 Aug. 2010 - see the whole document, especially Abstract; Claims 1, 21-23 | 1-10,20,40-44 ,57-59 |
| A | US 2006/0035258 A1 (TADAKAMALLA, R. P. et al.) 16 Feb. 2006 - see the whole document, especially Abstract; Claims 1, 12-13, 15 | 1-10,20,40-44 ,57-59 |
| A | US 7,700,324 B1 (ISSA, J. et al.) 20 Apr. 2010 - see the whole document, especially Abstract; Claims 1, 21-23 | 1-10,20,40-44 ,57-59 |
| A | US 2009/0137402 A1 (WANG, S. M. et al.) 28 May. 2009 - see the whole document, especially Abstract; Claims 1-2, 4 | 1-10,20,40-44 ,57-59 |

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 21 JUNE 2012 (21.06.2012) | **21 JUNE 2012 (21.06.2012)** |

| Name and mailing address of the ISA/KR | Authorized officer |
|---|---|
| Korean Intellectual Property Office<br>189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea<br>Facsimile No. 82-42-472-7140 | Lee, Jun Hyuk<br>Telephone No. 82-42-481-8115 |

Form PCT/ISA/210 (second sheet) (July 2009)



IPR2025-01353; IPR2025-01355
CSHL EX2002

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.
**PCT/US2011/057350**

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| US 2007-0172873 A1 | 26.07.2007 | EP 1848827 A2 | 31.10.2007 |
| | | EP 1856293 A2 | 21.11.2007 |
| | | US 2005-0181408 A1 | 18.08.2005 |
| | | US 2006-0177832 A1 | 10.08.2006 |
| | | US 2006-0177833 A1 | 10.08.2006 |
| | | US 7217522 B2 | 15.05.2007 |
| | | US 7393665 B2 | 01.07.2008 |
| | | US 7407757 B2 | 05.08.2008 |
| | | US 7537897 B2 | 26.05.2009 |
| | | WO 2006-086209 A2 | 17.08.2006 |
| | | WO 2006-086209 A3 | 17.08.2006 |
| | | WO 2006-086210 A2 | 17.08.2006 |
| | | WO 2006-086210 A3 | 17.08.2006 |
| | | WO 2006-099604 A2 | 21.09.2006 |
| | | WO 2006-099604 A3 | 21.09.2006 |
| | | WO 2007-087312 A2 | 02.08.2007 |
| | | WO 2007-087312 A3 | 02.08.2007 |
| US 2010-0203538 A1 | 12.08.2010 | WO 2010-088288 A2 | 05.08.2010 |
| | | WO 2010-088288 A3 | 06.01.2011 |
| | | WO 2010-088288 A3 | 05.08.2010 |
| US 2006-0035258 A1 | 16.02.2006 | US 2006-0194243 A1 | 31.08.2006 |
| US 7700324 B1 | 20.04.2010 | AU 2000-14537 A1 | 22.05.2000 |
| | | AU 2000-14537 B2 | 08.07.2004 |
| | | CA 2350085 A1 | 11.05.2000 |
| | | EP 1127157 A1 | 29.08.2001 |
| | | EP 1127157 A4 | 30.10.2002 |
| | | JP 2002-533061 A | 08.10.2002 |
| | | JP 2002-533061 T | 08.10.2002 |
| | | US 2011-0020807 A1 | 27.01.2011 |
| | | US 8133986 B2 | 13.03.2012 |
| | | WO 00-26401 A1 | 11.05.2000 |
| US 2009-0137402 A1 | 28.05.2009 | None | |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 13738414 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Cindy Shu |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 13-SEP-2012 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 16:39:41 |
| **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 | | 120913_5981_81503_A_SIDS_REB.pdf | 912563<br>4688c04d4af836666bc96404ec6e86a84bf2de56 | yes | 5 |



IPR2025-01353; IPR2025-01355
CSHL EX2002

| | Multipart Description/PDF files in .zip description | | | | |
|---|---|---|---|---|---|
| | **Document Description** | | **Start** | **End** | |
| | Transmittal Letter | | 1 | 3 | |
| | Information Disclosure Statement (IDS) Form (SB08) | | 4 | 5 | |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Non Patent Literature | 120913_5981_81503_A_SIDS_Exhibit_1_REB.pdf | 1315205<br>d97ecb016ded37510194e9cfdc5c6da00e7f2894 | no | 8 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Non Patent Literature | 120913_5981_81503_A_SIDS_Exhibit_3_REB.pdf | 1410899<br>11e1d25b645a00d6f515d98dbcc3b64d0157e9fe | no | 6 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 4 | Non Patent Literature | 120913_5981_81503_A_SIDS_Exhibit_4_REB.pdf | 692639<br>67f86b140b49f79c037416c1bd1fb46fcf7f3307 | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 5 | Non Patent Literature | 120913_5981_81503_A_SIDS_Exhibit_5_REB.pdf | 1080365<br>b8f27ddeeb1188f1d394a7fc2cf996bd71ab1ab6 | no | 11 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 6 | Non Patent Literature | 120913_5981_81503_A_SIDS_Exhibit_2_REB.pdf | 5178984<br>c3985b5dcbacb605e369fbcd53dd5c9213232cec | no | 5 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 10590655 | | |

**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.**

Docket No. 5981/81503-A/GJG/REB

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicant(s) : James Hicks et al.

U.S. Serial No. : 13/278,333

Filing Date : October 21, 2011

Confirmation No. : 4469

For : VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

30 Rockefeller Plaza, 20th Floor
New York, New York 10112
September 13, 2012

**BY EFS**
Commissioner for Patents
Alexandria, VA

**SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT**

In accordance with their duty of disclosure under 37 C.F.R. §1.56, Applicants would like to direct the Examiner's attention to the following documents which are listed on Form PTO-1449 (Substitute), attached hereto as **Exhibit A,** and are also listed below.

This Information Disclosure Statement is being submitted pursuant to 37 C.F.R. §1.97(b)(3) before the mailing of a first Office Action on the merits. Thus, this Supplemental Information Disclosure Statement should be entered and considered.

Reference item **1** is a U.S. Patent and reference items **6-9** are U.S. Patent Application Publications. Pursuant to C.F.R. 1.98(a)(2), copies of these reference items are not being submitted.

Copies of reference items **2-5** and **10** are attached hereto as **Exhibits 1-5**, respectively.

1. U.S. Patent No. 6,172,214, issued January 9, 2001 (Brenner et al.);


IPR2025-01353; IPR2025-01355
CSHL EX2002

2. Islam et al., "Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq" Genome Research (2001) 21:1160-1167 **(Exhibit 1)**;

3. Kivioja et al., "Counting absolute numbers of molecules using unique molecular identifiers" Nature Methods (2012) 9(1):72-74 and Online Methods (2 pages) **(Exhibit 2)**;

4. Shiroguchi et al., "Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes" PNAS (2012) 109(4):1347-1352 **(Exhibit 3)**;

5. International Search Report, issued in connection with PCT International Application No. PCT/US2011/057350 **(Exhibit 4)**;

6. U.S. Patent Application Publication No. 2006/0035258, published February 16, 2006 (Tadakamalla et al.);

7. U.S. Patent Application Publication No. 2007/0172873, published July 26, 2007 (Brenner et al.);

8. U.S. Patent Application Publication No. 2009/0137402, published May 28, 2009 (Wang et al.);

9. U.S. Patent Application Publication No. 2010/0203538, published August 12, 2010 (Dube et al.); and

10. Castle et al., "DNA copy number, including telomeres and mitochondria, assayed using next-generation sequencing" BMC Genomics (2010) 11(244):1-11 **(Exhibit 5)**.

The Examiner is respectfully requested to make of record each item listed on the Form PTO-1449 Substitute by initialing and dating the attached Form PTO-1449 Substitute, and returning a copy of the initialed and dated form to applicants' undersigned attorney.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicants' undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this Supplemental Information Disclosure Statement. However, if any additional fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39,992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**
I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

9-13-12
Robert E. Bolcome III Date
Reg. No. 69,032

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | **Application Number** | 13/278,333 |
| | **Filing Date** | October 21, 2011 |
| | **First Named Inventor** | James Hicks |
| | **Art Unit** | |
| | **Examiner Name** | |
| | **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| | 1 | 7,700,324 | 04-20-2010 | Issa et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | 3 | WO 2012/054873 | 04-26-2012 | Hicks et al. | |
| | | | | | |
| | | | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|
| | |

*EXAMINER: Initial if citation 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 of 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | **Application Number** | 13/278,333 |
| | **Filing Date** | October 21, 2011 |
| | **First Named Inventor** | James Hicks |
| | **Art Unit** | |
| | **Examiner Name** | |
| | **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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] |
|---|---|---|---|
| | 2 | International Preliminary Report on Patentability, issued April 23, 2013 in connection with PCT International Application No. PCT/US2011/057350 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

**PATENT COOPERATION TREATY**

# PCT

## INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY
(Chapter I of the Patent Cooperation Treaty)

(PCT Rule 44*bis*)

| | | |
|---|---|---|
| Applicant's or agent's file reference<br>81503-A-PCT/GJG/WS | **FOR FURTHER ACTION** | See item 4 below |
| International application No.<br>PCT/US2011/057350 | International filing date *(day/month/year)*<br>21 October 2011 (21.10.2011) | Priority date *(day/month/year)*<br>22 October 2010 (22.10.2010) |
| International Patent Classification (8th edition unless older edition indicated)<br>See relevant information in Form PCT/ISA/237 | | |
| Applicant<br>COLD SPRING HARBOR LABORATORY | | |

1. This international preliminary report on patentability (Chapter I) is issued by the International Bureau on behalf of the International Searching Authority under Rule 44 *bis*.1(a).

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

   In the attached sheets, any reference to the written opinion of the International Searching Authority should be read as a reference to the international preliminary report on patentability (Chapter I) instead.

3. This report contains indications relating to the following items:

| | | |
|---|---|---|
| ☒ | Box No. I | Basis of the report |
| ☐ | Box No. II | Priority |
| ☒ | Box No. III | Non-establishment of opinion with regard to novelty, inventive step and industrial applicability |
| ☐ | Box No. IV | Lack of unity of invention |
| ☒ | Box No. V | Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement |
| ☐ | Box No. VI | Certain documents cited |
| ☒ | Box No. VII | Certain defects in the international application |
| ☐ | Box No. VIII | Certain observations on the international application |

4. The International Bureau will communicate this report to designated Offices in accordance with Rules 44*bis*.3(c) and 93*bis*.1 but not, except where the applicant makes an express request under Article 23(2), before the expiration of 30 months from the priority date (Rule 44*bis* .2).

| | |
|---|---|
| | Date of issuance of this report<br>23 April 2013 (23.04.2013) |
| The International Bureau of WIPO<br>34, chemin des Colombettes<br>1211 Geneva 20, Switzerland<br>Facsimile No. +41 22 338 82 70 | Authorized officer<br>Yukari Nakamura<br>e-mail: pt07.pct@wipo.int |

Form PCT/IB/373 (January 2004)

# PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
GERSHIK GARY J.

COOPER & DUNHAM LLP 30 ROCKEFELLER PLAZA
NEW YORK NY 10112 USA

# PCT

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

**(PCT Rule 43bis.1)**

Date of mailing *(day/month/year)* **21 JUNE 2012 (21.06.2012)**

Applicant's or agent's file reference
81503-A-PCT/GJG/WS

**FOR FURTHER ACTION**
See paragraph 2 below

| International application No. | International filing date *(day/month/year)* | Priority date *(day/month/year)* |
|---|---|---|
| **PCT/US2011/057350** | **21 OCTOBER 2011 (21.10.2011)** | 22 OCTOBER 2010 (22.10.2010) |

International Patent Classification (IPC) or both national classification and IPC

***C12Q 1/68(2006.01)i, C12N 15/11(2006.01)i, C12Q 1/48(2006.01)i***

Applicant

**COLD SPRING HARBOR LABORATORY et al**

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☒ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☒ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1bis(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.
For further options, see Form PCT/ISA/220.

| Name and mailing address of the ISA/KR | Date of completion of this opinion | Authorized officer |
|---|---|---|
| Korean Intellectual Property Office<br>189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea<br>Facsimile No. 82-42-472-7140 | 21 JUNE 2012 (21.06.2012) | Lee, Jun Hyuk<br>Telephone No.82-42-481-8115 |

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
**PCT/US2011/057350**

**Box No. I Basis of this opinion**

1. With regard to the **language**, this opinion has been established on the basis of :

   ☒ the international application in the language in which it was filed

   ☐ a translation of the international application into _____ , which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

   a. a sequence listing filed or furnished
      ☐ on paper
      ☐ in electronic form

   b. time of filing or furnishing
      ☐ contained in the international application as filed.
      ☐ filed together with the international application in electronic form.
      ☐ furnished subsequently to this Authority for the purposes of search.

4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additioanl copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
**PCT/US2011/057350**

**Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability**

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application

☒ claims Nos. 11-19,21-39,45-56,60-67

because:

☐ the said international application, or the said claims Nos. _____ relate to the following subject matter which does not require an international search *(specify)*:

☒ the description, claims or drawings *(indicate particular elements below)* or said claims Nos. see below are so unclear that no meaningful opinion could be formed *(specify)*:

The subject matter of claims 15, 27, 28, 30, 32-35, 38, 39, 52, 54, 55, 64 and 65 is so unclear and indefinite that no meaningful opinion could be made because these claims are drafted in reference to multiple dependent claims such as claims 11-14, 16-19, 21-26, 29, 31, 36, 37, 45-51, 53, 56, 60-63, 66 and 67.

☐ the claims, or said claims Nos. _____ are so inadequately supported by the description that no meaningful opinion could be formed *(specify)*:

☒ no international search report has been established for said claims Nos. 11-19,21-39,45-56,60-67

☐ a meaningful opinion could not be formed without the sequence listing; the applicant did not, within the prescribed time limit:

☐ furnish a sequence listing on paper complying with the standard provided for in Annex C of the Administrative Istructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.

☐ furnish a sequence listing in electronic form complying with the standard provided for in Annex C of the Administrative Istructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.

☐ pay the required late furnishing fee for the furnishing of a sequence listing in response to an invitation under Rule 13ter.1(a) or (b).

☐ See Supplemental Box for further details.

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
**PCT/US2011/057350**

**Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

| | | | |
|---|---|---|---|
| Novelty (N) | Claims | 1-10,20,40-44,57-59 | YES |
| | Claims | NONE | NO |
| Inventive step (IS) | Claims | 1-10,20,40-44,57-59 | YES |
| | Claims | NONE | NO |
| Industrial applicability (IA) | Claims | 1-10,20,40-44,57-59 | YES |
| | Claims | NONE | NO |

2. Citations and explanations :

Reference is made to the following documents:
D1: BMC Genomics, 2010 Apr. 16, Vol. 11, No. 244, pp1-11

D1 discloses a method for assaying DNA copy number using next-generation sequencing.

**1. Novelty and Inventive step**

Claims 1-7 are directed to a method for obtaining from genomic material genomic copy number information unaffected by amplification distortion. Claims 8, 9 and 20 are directed to a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion. Claim 10 is directed to a method for obtaining from genomic material DNA methylation information unaffected by amplification distortion. Claims 40-44 are directed to a composition of matter derived from genomic material or mRNA transcripts. Claims 57-59 are directed to a kit for determining nucleic acid copy number information unaffected by amplification distortion.

The subject matter of the present claims is different from D1, the most relevant state of the art, in that the tagged nucleic acid molecules, which comprise one segment of the genomic material and a tag, and the tag associated sequence reads are used to obtain genomic (or mRNA transcripts; or nucleic acid) copy number information unaffected by amplification distortion. There is no other prior art documents which directly indicates or fairly suggests the above distinguishing feature. And the present invention does not seem to be obviously achievable by a person skilled in the art.

Continued on Supplemental Box

**Box No. VII Certain defects in the international application**

The following defects in the form or contents of the international application have been noted:

Claims 11-14, 16-19, 21-26, 29, 31, 36, 37, 45-51, 53, 56, 60-63, 66 and 67 do not comply with PCT Rule 6.4(a) because multiple dependent claims should not serve as a basis for any other multiple dependent claim.

**Supplemental Box**

In case **the space in any of the preceding boxes is not sufficient.**
Continuation of :

Box V

Accordingly, claims 1-10, 20, 40-44 and 57-59 meet the requirements of PCT Article 33(2) and 33(3) with respect to novelty and inventive step.

## 2. Industrial Applicability

Claims 1-10, 20, 40-44 and 57-59 of the present invention are considered industrially applicable under Article 33(4) PCT.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2012 (26.04.2012)

PCT

(10) International Publication Number
WO 2012/054873 A2

(51) International Patent Classification:
*C12Q 1/68* (2006.01) *C12Q 1/48* (2006.01)
*C12N 15/11* (2006.01)

(21) International Application Number:
PCT/US2011/057350

(22) International Filing Date:
21 October 2011 (21.10.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

| | | |
|---|---|---|
| 61/406,067 | 22 October 2010 (22.10.2010) | US |
| 61/510,579 | 22 July 2011 (22.07.2011) | US |
| 13/278,333 | 21 October 2011 (21.10.2011) | US |

(71) Applicant *(for all designated States except US)*: **COLD SPRING HARBOR LABORATORY** [US/US]; One Bungtown Road, Cold Spring Harbor, NY 11724 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HICKS, James** [US/US]; 3 Great Meadow Road, Locust Valley, New York 11560 (US). **NAVIN, Nicholas** [US/US]; 1621 Wichita St., Houston, TX 77004 (US). **TROGE, Jennifer** [US/US]; 100 Simcoe St., Oyster Bay, NY 11771 (US). **WANG, Zihua** [CN/US]; 13 Meadowlawn St., Huntington, NY 11743 (US). **WIGLER, Michael** [US/US]; 52 Snake Hill Road, Cold Spring Harbor, NY 11724 (US).

(74) Agent: **GERSHIK, Gary, J.**; Cooper & Dunham Llp, 30 Rockefeller Plaza, New York, NY 10112 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published**:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

Applicants: James Hicks et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
IDS Exhibit 2

(54) **Title**: VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

(57) **Abstract**: A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising obtaining segments of the genomic material, tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material and a tag, subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR), generating tag associated sequence reads by sequencing the product of the PCR reaction, assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome, and counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome, thereby obtaining genomic copy number information unaffected by amplification distortion.

WO 2012/054873 A2

# VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

This application claims priority of U.S. Serial No. 13/278,333, filed October 21, 2011 and U.S. Provisional Applications Nos. 61/510,579, filed July 22, 2011, and 61/406,067, filed October 22, 2010 the contents of each of which are hereby incorporated by reference in their entirety.

The invention disclosed herein was supported by the government under grant No. W81XWH-09-1-0591 from the Department of Defense. Accordingly, the U. S. Government has certain rights to this invention.

Throughout this application, various publications are referenced by numbers in parentheses. Full citations for these references may be found at the end of the specification immediately preceding the claims. The disclosures of these publications in their entireties are hereby incorporated by reference into this application to more fully describe the state of the art to which this invention pertains.

## Background of the Invention

Genomic copy number information is commonly obtained using whole genome amplification (WGA). The endemic problem with the WGA method is over-sampling of certain regions, yielding a non-uniform amplification of the genome (1). WGA methods begin with the step that initiates the process, a polymerase (Phi 29) makes a strand from genomic DNA utilizing a random primer coupled to an adaptor for subsequent PCR (Figure 1). If the input DNA strands are referred to as the "0-th derivative", and the first synthesized strand as a "first derivative," subsequent strands are called the (n+1)-th derivative if their template was an n-th derivative. Only strands that are 2-nd derivative or higher become amplified in the PCR step, resulting in a 'stacking' over the regions 'chosen' by the polymerase for the first derivative.

Coverage of the genome by sequencing WGA of single cell DNA is limited by the stacking phenomenon (Figure 2). Thus it is difficult to obtain single cell measurements, particularly when based on WGA, due to distortions that originate from stochastic sampling and amplification steps. Moreover, the current method of WGA is a black box, with the unspecified reagents purchased from a vendor, which hampers optimization. Moreover, the WGA method does not extend to a method usable for single cell RNA profiling.

Ligation-mediated PCR was developed in an attempt to solve the above-identified problems inherent in WGA. In this method, adaptors are ligated to an *Mse*I restriction endonuclease digest of genomic DNA from a single cell, followed by PCR amplification using primers complementary to the adaptors. The amplified DNA is then used for CGH or DNA sequencing (2,3). However, like WGA, the method still requires an amplification step.

Parameswaran et al. (2007) and U.S. Patent No. 7,622,281 describe methods of labeling nucleic acid molecules with barcodes for the purpose of identifying the source of the nucleic acid molecules, thereby allowing for high-throughput sequencing of multiple samples (4,5). Eid et al. (2009) describe a single molecule sequencing method wherein single-molecule real time sequencing data is obtained from a DNA polymerase performing uninterrupted template-directed synthesis using four distinguishable fluorescently labeled dNTPs (6). However, these methods do not provide genomic information unaffected by amplification distortion.

Miner et al. (2004) describe a method of molecular barcoding to label template DNA prior to PCR amplification, and report that the method allows for the identification of contaminant and redundant sequences by counting only distinctly tagged sequences (22). U.S. Patent No. 7,537,897 describes methods for molecular counting by labeling molecules of an input sample with unique oligonucleotide tags and subsequently amplifying and counting the number of different tags (23). Miner et al. and U.S. Patent No. 7,537,897 both describe labeling of input nucleic acid molecules by ligation, which has been found to be an inefficient reaction.

McCloskey et al. (2007) describe a method of molecular encoding which does not use ligation but instead uses template specific primers to barcode template DNA molecules prior to PCR amplification (24). However, such a method requires that template specific primers be made for each species of template DNA molecule studied.

As described herein, obtaining accurate genomic copy number information by high-throughput sequencing of genomic DNA prepared by WGA methods is hampered by the copy number distortions introduced by non-uniform amplification of genomic DNA. Thus, there exists a need for a method that allows for copy number determination free of distortions caused by amplification steps and which allows for accurate and efficient copy number determination of complex samples. Such a method should also be robust using existing methodologies for high volume, massively parallel sequencing.


IPR2025-01353; IPR2025-01355
CSHL EX2002

## SUMMARY OF THE INVENTION

A method is provided for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA;



IPR2025-01353; IPR2025-01355
CSHL EX2002

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b), and

ii) a tag;

d) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

e) generating tag associated sequence reads by sequencing the product of step (d);

f) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a cDNA or a segment of cDNA to a location on the cDNA library; and

g) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;



IPR2025-01353; IPR2025-01355
CSHL EX2002

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from genomic material DNA methylation information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands to a polymerase reaction in the presence essentially unique primers, wherein the essentially unique primers are capable of hybridizing to the polynucleotide tail of the zero-th order strands, under conditions that promote the formation of only one complement, thereby generating tagged nucleic acid molecules;

d) separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules;

e) subjecting each group of step (d) to amplification by polymerase chain reaction (PCR);

f) generating tag associated sequence reads by sequencing the product of step (e);

g) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

h) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining DNA methylation information unaffected by amplification distortion.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;


IPR2025-01353; IPR2025-01355
CSHL EX2002

c) subjecting the zero-th order derivative strands of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

d) adding a polynucleotide tail to the first order derivative strands;

e) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (c) and (e) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of genomic material;

b) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

c) subjecting the zero-th order derivative strands of step (b) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (b) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

d) subjecting the product of step (c) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (b) under conditions which promote the formation of only one complement, wherein the primers of step (d) have different


IPR2025-01353; IPR2025-01355
CSHL EX2002

nucleotide sequences than the primers of step (c), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (c) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from a mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA; and

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, wherein each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b) and

ii) a tag.

Also provided is a composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining mRNA transcripts;

b) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the



IPR2025-01353; IPR2025-01355
CSHL EX2002

polynucleotide tail added in step (c) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a kit for determining nucleic acid copy number information unaffected by amplification distortion comprising:

a) a terminal transferase, and

b) plurality of substantially unique primers,

wherein the substantially unique primers comprise substantially unique tags, and wherein the substantially unique primers are capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the terminal transferase.



IPR2025-01353; IPR2025-01355
CSHL EX2002

## BRIEF DESCRIPTION OF THE FIGURES

**Figure 1. Whole Genome Amplification (WGA).**

**Figure 2. Genome coverage comparison.**

Genomic coverage comparison of 7 sequencing lanes from one single cell library versus 7 lanes from a million cell library. Stacking causes the single cell reads to be concentrated over the first derivative reads from WGA. More reads yield diminishing returns.

**Figure 3. Sequencing read counts for CpG dinucleotides within selected CpG islands**

Normal cell line SKN-1 (A,C,E) and breast cancer line MDA-MB-231 (B,D,F). Orange indicates T residues (unmethylated). Blue indicates C residues (methylated). Counts of the plus strand are above the X-axis, and counts on the minus strand are below.

**Figure 4. Comparison of array CGH with 'sequence counting' for analyzing copy number data.**

Graph shows a portion of chromosome 17 from FFPE sample JZ33 including raw (gray) ratio and segmented (green) data from array CGH, and raw sequence count data (orange) and its segmentation (blue).

**Figure 5. ROMA CGH of sKBR3 compared with profiles of 3 independent single cell isolates derived from sparse DNA sequencing.**

**Figure 6. Sequencing across the breakpoints of a homozygous deletion in cell on chromosome 5 line T47D.**

**Figure 7. Flowchart for genomic profiling by representational sequence counting**

**Figure 8. Scheme for varietal counting.** Varietal tags are shown as random nucleotide sequences ($N_{10}$)

Flowchart of one embodiment of the invention.

**Figure 9. Nick ligation protocol**

The sequence of primer A includes universal primer A (PrA), varietal tag (B), and oligo dT+CATG, the sequence of primer B includes universal primer B (PrB) and oligo

dT+CATG, PrA and PrB are universal primers, and cPrA, cPrB, and cB are complementary to PrA, PrB and B, respectively.

**Figure 10. Schema for targeted sequencing of regional markers**

UfBPrA is forward primer, B is tag, UrPrR is reverse primer, PrB is gene-specific reverse primer, Uf is universal forward primer, PrA is gene specific forward primer, Ur is universal reverse primer.

**Figure 11. Comparison of Varietal Counting with Sequence Counting**

Figure 11A shows genome wide copy number data. Figure 11B shows partial genome copy number data.

## DETAILED DESCRIPTION OF THE INVENTION

A method is provided for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

In an embodiment, the method further comprises estimating a genomic copy number of a region of the genome comprising more than one location on the genome by assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

In an embodiment, the method further comprises comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

In an embodiment, the method further comprises

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

Step (b) of the above embodiment may further comprise

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

In an embodiment of the method, the second region of the genome comprises a centromere.

In an embodiment, the method further comprises summing the counts obtained in step (f) for locations which comprise a region of the genome, and comparing the sum to a sum obtained from a reference sample for the same region of the genome, thereby estimating a relative genomic copy number of the region of the genome.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA;

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b), and

ii) a tag;

d) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

e) generating tag associated sequence reads by sequencing the product of step (d);

f) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a cDNA or a segment of cDNA to a location on the cDNA library; and

g) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the

formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

a) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

b) generating tag associated sequence reads by sequencing the product of step (b);

c) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

d) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

Also provided is a method for obtaining from genomic material DNA methylation information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;


IPR2025-01353; IPR2025-01355
CSHL EX2002

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands to a polymerase reaction in the presence essentially unique primers, wherein the essentially unique primers are capable of hybridizing to the polynucleotide tail of the zero-th order strands, under conditions that promote the formation of only one complement, thereby generating tagged nucleic acid molecules;

d) separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules;

e) subjecting each group of step (d) to amplification by polymerase chain reaction (PCR);

f) generating tag associated sequence reads by sequencing the product of step (e);

g) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

h) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining DNA methylation information unaffected by amplification distortion.

In an embodiment of the methods, tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;


IPR2025-01353; IPR2025-01355
CSHL EX2002

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

In an embodiment of the methods, tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

In an embodiment of the methods, adding a polynucleotide tail comprises the use of a terminal transferase.

In an embodiment of the methods, tagging the segments to generate tagged nucleic acid molecules comprises ligation of adaptors comprising the tags to at least one end of the segments of the genomic material.

In an embodiment of the methods, the adaptors comprising the tags are ligated to only one end of the segments of the genomic material.

In an embodiment of the methods, the tags comprise a sequence that aids PCR amplification.

In an embodiment of the methods, each tagged nucleic acid molecule comprises one tag.

In an embodiment of the methods, each tagged nucleic acid molecule comprises more than one tag.

In an embodiment of the methods, segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

In an embodiment of the methods, segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

In an embodiment of the methods, the maximum copy number of a location on a cDNA library is not less than the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library.

An embodiment of the above methods further comprises analyzing mRNA copy number.

In an embodiment of the methods, separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules is by cleavage with methylation sensitive restriction enzymes, partitioning with antibodies, or partitioning with methyl-C binding proteins directed to methylated or hydroxymethylated cytosine.


IPR2025-01353; IPR2025-01355
CSHL EX2002

In an embodiment of the methods, the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or prior to sequencing.

In an embodiment of the methods, each tagged nucleic acid molecule differs at more than one nucleotide.

In an embodiment of the methods, the tag sequences further comprise a sample tag.

In an embodiment of the methods, the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to sequencing.

An embodiment of the methods further comprises deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to sample tag.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from a single species.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from a single organism.

In an embodiment of the methods, tagged nucleic acid molecules are generated from a single cell.

In an embodiment, the single cell is from a needle aspirate of suspected cancer lesions.

In an embodiment, the single cell is from a core biopsy of suspected cancer lesions.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from two or more organisms.

In an embodiment, the tagged nucleic acid molecules are generated from a single cell of each organism.

In an embodiment of the methods, the single species is human.

In an embodiment of the methods, the tagged nucleic acid molecules are generated from two or more species.

In an embodiment, the tagged nucleic acid molecules are generated from a population of microbes.

A further embodiment comprises comparing the genomic copy number information obtained for different species of the population to determine the relative count of those different species in the population.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

d) adding a polynucleotide tail to the first order derivative strands;

e) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (c) and (e) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of genomic material;


IPR2025-01353; IPR2025-01355
CSHL EX2002

b) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

c) subjecting the zero-th order derivative strands of step (b) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (b) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

d) subjecting the product of step (c) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (b) under conditions which promote the formation of only one complement, wherein the primers of step (d) have different nucleotide sequences than the primers of step (c), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (c) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

Also provided is a composition of matter derived from a mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA; and

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, wherein each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b) and

ii) a tag.



IPR2025-01353; IPR2025-01355
CSHL EX2002

Also provided is a composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining mRNA transcripts;

b) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (c) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

In an embodiment of the above compositions, adding a polynucleotide tail comprises the use of a terminal transferase.

In an embodiment of the above compositions, each tagged nucleic acid molecule comprises one tag.

In an embodiment of the above compositions, each tagged nucleic acid molecule comprises more than one tag.

In an embodiment of the above compositions, segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

In an embodiment of the above compositions, segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.



IPR2025-01353; IPR2025-01355
CSHL EX2002

In an embodiment of the above compositions, the complexity of the tagged nucleic acid molecules is reduced by hybrid capture.

In an embodiment of the above compositions, each tagged nucleic acid molecules differs at more than one nucleotide.

In an embodiment of the above compositions, the tag sequences further comprise a sample tag.

Also provided is a composition of matter comprising at least two pools of tagged nucleic acid molecules as described herein.

In an embodiment of the above compositions, the tagged nucleic acid molecules are generated from a single cell.

In an embodiment, the single cell is from a needle aspirate of suspected cancer lesions.

In an embodiment, the single cell is from a core biopsy of suspected cancer legions.

In an embodiment of the above compositions, the tagged nucleic acid molecules are generated from two or more organisms.

Also provided is a kit for determining nucleic acid copy number information unaffected by amplification distortion comprising:

a) a terminal transferase, and

b) plurality of substantially unique primers,

wherein the substantially unique primers comprise substantially unique tags, and wherein the substantially unique primers are capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the terminal transferase.

In an embodiment, the kit further comprises a DNA polymerase having 3'-5' proofreading activity.

In an embodiment of the kits, the plurality of substantially unique primers comprises $10^n$ primers, wherein n is an integer from 2 to 9.



IPR2025-01353; IPR2025-01355
CSHL EX2002

In an embodiment, the kits further comprise a second terminal transferase and a primer capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the second terminal transferase.

In an embodiment of the kits, the tags are six nucleotides long.

In an embodiment of the kits, the tags are 15 nucleotides long.

In an embodiment, the kits further comprise sample tags.

In an embodiment, the sample tags are 2 nucleotides long.

In an embodiment, the sample tags are 4 nucleotides long.

In an embodiment, the sample tags comprise a sample tag set.

In an embodiment of the above kits, the substantially unique tags comprise the sample tags.

**Terms**

For the purpose of this invention, different words and phrases are defined as follows:

As used herein, the term "adaptor" refers to an oligonucleotide or nucleic acid fragment or segment that can be ligated to a nucleic acid molecule of interest. For the purposes of this invention adaptors may, as options, comprise primer binding sites, recognition sites for endonucleases, common sequences, promoters, tag sequences, and sample tag sequences. Preferably, adaptors are positioned to be located on both sides (flanking) a particular nucleic acid molecule of interest. In accordance with the invention, adaptors may be added to nucleic acid molecules of interest by standard recombinant techniques (e.g. restriction digest and ligation). For example, adaptors may be added to a population of linear molecules, (e.g. a genomic DNA which has been cleaved or digested) to form a population of linear molecules containing adaptors at one and preferably both termini of all or a substantial portion of molecules. The adaptor may be entirely or substantially double stranded or entirely single stranded. A double stranded adaptor may comprise two oligonucleotides that are at least partially complementary. The adaptor may be phosphorylated or unphosphorylated on one or both strands. Adaptors may be used for DNA sequencing. Adaptors may also incorporate modified nucleotides that modify the properties of the adaptor sequence. For example, methylated cytosines may be substituted for cytosines. In an embodiment of this invention the

adaptors ligated to genomic DNA to enable cluster generation on the sequencer contain cytosines which were all methylated. This modification protects such adaptors from bisulfite conversion, and is taken into account in the downstream applications and analysis of this invention.

As used herein, the term "amplifying" refers to the process of synthesizing nucleic acid molecules that are complementary to one or both strands of a template nucleic acid. Amplifying a nucleic acid molecule typically includes denaturing the template nucleic acid, annealing primers to the template nucleic acid at a temperature that is below the melting temperatures of the primers, and enzymatically elongating from the primers to generate an amplification product. The denaturing, annealing and elongating steps each can be performed once. Generally, however, the denaturing, annealing and elongating steps are performed multiple times (e.g., polymerase chain reaction (PCR)) such that the amount of amplification product is increasing, often times exponentially, although exponential amplification is not required by the present methods. Amplification typically requires the presence of deoxyribonucleoside triphosphates, a DNA polymerase enzyme and an appropriate buffer and/or co-factors for optimal activity of the polymerase enzyme. The term "amplification product" refers to the nucleic acid sequences, which are produced from the amplifying process as defined herein.

As used herein, the term "amplification distortion" refers to non-uniform amplification of template nucleic acid molecules.

As used herein, the term "bisulfite-treatment" refers to the treatment of nucleic acid with a reagent used for the bisulfite conversion of cytosine to uracil. Examples of bisulfite conversion reagents include but are not limited to treatment with a bisulfite, a disulfite or a hydrogensulfite compound.

As used herein, the term "bisulfite-converted material" refers to a nucleic acid that has been contacted with bisulfite ion in an amount appropriate for bisulfite conversion protocols known in the art. Thus, the term "bisulfite-converted material" includes nucleic acids that have been contacted with, for example, magnesium bisulfite or sodium bisulfite, prior to treatment with base.

As used herein, the term "sequence capture" or "hybrid capture" refers to a process of hybridizing a "capture probe" to a nucleic acid having a sequence which is complementary to

the sequence of the capture probe. A capture probe may be immobilized to a substrate on a solid phase microarray, wherein "substrate" refers to short nucleic acid sequences which are known. Their location on the solid phase microarray may or may not be predetermined. The capture probe comprising a "sequence complementary to the substrate" may be immobilized to the solid phase microarray by hybridizing to its complementary "substrate sequence."

As used herein, the term "clamp" refers to a nucleotide sequence of CC, CG, GC or GG which may be located at the first two nucleotide positions at the 5' end of a primer, tag, or sample tag, or the last two nucleotide positions at the 3' end of a primer, tag, or sample tag.

As used herein, the term "comparable number of locations" means that a first genomic region has a number of locations on the genome which is within 25% of the number of locations of a second genomic region. In an embodiment, the first genomic region which has a comparable number of locations to the second genomic region has a number of locations which is within 20%, 15%, 10%, 5%, 1%, or exactly the same as the number of locations of the second genomic location.

As used herein, a subsequence of a tag associated sequence read is "corresponding to" a tag when the subsequence is identical to the nucleotide sequence of the tag.

As used herein, a subsequence of a tag associated sequence read is "corresponding to" a species of a nucleic acid molecule when the subsequence is substantially identical to or substantially complementary to at least about 10, 12, 14, 16, 18, 20, or more nucleotides of the sequence of the species of nucleic acid molecule. In an embodiment, the subsequence is identical to or fully complementary to at least about 10, 12, 14, 16, 18, 20, or more nucleotides of the sequence of the species of nucleic acid molecule.

The term "CpG site" refers to the CpG dinucleotide.

As used here, the term "CpG island" refers to a region of DNA with a high G+C content and a high frequency of CpG dinucleotides relative to the bulk genome, as further defined at the UCSC Genome Bioinformatics Site at genome.ucsc.edu/index.html?org=Human&db=hg19&hgsid=171216665, and in particular, by annotation in the USCS Genome Bioinformatics Site database (CpG Islands Track or Table: cpgIslandExt).

As used herein, the term "fully complementary" refers to the reverse complement of a nucleic acid sequence.

As used herein, the term "library" refers to a collection of nucleic acid molecules (circular or linear). In one preferred embodiment, a library is representative of all of the DNA content of an organism (such a library is referred to as a "genomic" library), or a set of nucleic acid molecules representative of all of the expressed genes (such a library is referred to as a cDNA library) in a cell, tissue, organ or organism. A library may also comprise random sequences made by de novo synthesis, mutagenesis of one or more sequences and the like. A library may be contained in one vector.

As used herein, the term "mapping" refers to identifying a location on a genome or cDNA library that has a sequence which is substantially identical to or substantially fully complementary to the subsequence of a tag associated sequence read corresponding to a species of nucleic acid molecule, and assigning the tag associated sequence read or the tagged nucleic acid molecule generating the tag associated sequence read to the location. The nucleic acid molecule may be, but is not limited to the following: a segment of genomic material, a cDNA, a mRNA, or a segment of a cDNA.

As used herein, the term "methylation" refers to the covalent attachment of a methyl group at the C5-position of the nucleotide base cytosine within the CpG dinucleotides of genomic region of interest. The term "methylation state" or refers to the presence or absence of 5-methyl-cytosine ("5- Me") at one or a plurality of CpG dinucleotides within a DNA sequence. A methylation site is a sequence of contiguous linked nucleotides that is recognized and methylated by a sequence specific methylase. A methylase is an enzyme that methylates (i. e., covalently attaches a methyl group) one or more nucleotides at a methylation site.

As used herein, the term "mRNA transcript" refers to the nucleic acid molecule resulting from transcription of DNA.

As used herein, the term "pool" means to combine one plurality of nucleic acid molecules with at least one other plurality of nucleic acid molecules. In an embodiment, tagged nucleic acid molecules are pooled. Pooling may be done after any step in which tagged nucleic acid molecules are generated. In an embodiment, pooling is done prior to PCR and/or prior to sequencing. In an embodiment, tagged nucleic acid molecules generated from a single cell

are pooled with tagged nucleic acid molecules generated from a different single cell.

As used herein, the term "probe array" refers to an array of N different nucleic acid molecules deposited on a reaction substrate which serve to interrogate mixtures of target molecules or multiple sites on a single target molecule administered to the surface of the array.

As used herein, the term "read" or "sequence read" refers to the nucleotide or base sequence information of a nucleic acid that has been generated by any sequencing method. A read therefore corresponds to the sequence information obtained from one strand of a nucleic acid fragment. For example, a DNA fragment where sequence has been generated from one strand in a single reaction will result in a single read. However, multiple reads for the same DNA strand can be generated where multiple copies of that DNA fragment exist in a sequencing project or where the strand has been sequenced multiple times. A read therefore corresponds to the purine or pyrimidine base calls or sequence determinations of a particular sequencing reaction.

As used herein, the term "reference genome" refers to a genome of the same species as that being analyzed for which genome the sequence information is known.

As used herein, the term "region of the genome" refers to a continuous genomic sequence comprising multiple discrete locations.

As used herein, the term "representation of DNA" or "representation of a genome" refers to a sampling of the DNA or genome produced by a restriction endonuclease digestion of genomic or other DNA, followed by linkage of adaptors and then amplification with primers complementary to the adaptors.

As used herein, the term "sample tag" refers to a nucleic acid having a sequence no greater than 1000 nucleotides and no less than two that may be covalently attached to each member of a plurality of tagged nucleic acid molecules or tagged reagent molecules. A "sample tag" may comprise part of a "tag."

As used herein, the term "sample tag set" refers to a plurality of unique sample tags of the same length, wherein the nucleotide sequence of each unique sample tag differs from the nucleotide sequence of any other unique sample tag in the sample tag set at two or more

corresponding positions in the sequence. A unique sample tag selected from the sample tag set may be covalently attached to each member of a unique sample consisting of a plurality of tagged nucleic acid molecules in order to identify a tagged nucleic acid molecule as a member of that unique sample.

As used herein, the term "segments of cDNA," refers to the nucleic acid molecules resulting from fragmentation of cDNA.

As used herein, the term "segments of genomic material" refers to the nucleic acid molecules resulting from fragmentation of genomic DNA.

A nucleic acid molecule which contains a sequence of nucleotides identical to a segment of a cDNA or a segment of genomic material comprises a "segment of cDNA" or a "segment of genomic material," respectively.

As used herein, the "sequence complexity" or "complexity" with regards to a population of polynucleotides refers to the number of different species of polynucleotides present in the population.

As used herein, "substantially identical" or "substantially fully complementary" sequences have at least about 80% sequence identity or complementarity, respectively, to a nucleotide sequence. Substantially identical sequences or substantially fully complementary sequences may have at least about 85%, 90%, 95% or 100% sequence identity or complementarity, respectively.

As used herein, the term "substantially unique primers" refers to a plurality of primers, wherein each primer comprises a tag, and wherein at least 50% of the tags of the plurality of primers are unique. Preferably, the tags are at least 60%, 70%, 80%, 90%, or 100% unique tags.

As used herein, the term "substantially unique tags" refers to tags in a plurality of tags, wherein at least 50% of the tags of the plurality are unique to the plurality of tags. Preferably, substantially unique tags will be at least 60%, 70%, 80%, 90%, or 100% unique tags.

As used herein, the term "tag" refers to a nucleic acid having a sequence no greater than 1000 nucleotides and no less than two that may be covalently attached to a nucleic acid molecule or reagent molecule. A tag may comprise a part of an adaptor or a primer.


IPR2025-01353; IPR2025-01355
CSHL EX2002

As used herein, a "tagged nucleic acid molecule" refers to a nucleic acid molecule which is covalently attached to a "tag."

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range, and any other stated or intervening value in that stated range, is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

All combinations of the embodiments of the invention described herein are within the scope of the invention unless clearly indicated otherwise.

*Single Cell Copy Number Profiling by 'Sequence Counting' and Sample Tagging.*

High throughput sequencers are in effect single molecule sequencers. Because the human genome is sequenced and assembled, a single molecule sequencer is also a single molecule identifier. Therefore, sequencing is also counting, as illustrated in Figure 4 which illustrates how low coverage sequencing can yield an estimate of copy number which is comparable to array hybridization. The distinction for this is that 'sequence counting' can be done on single cells, and, using the multiplex system disclosed herein, relatively inexpensively, in effect converting the high throughput sequencer into a detector of single cells with abnormal genomes. To the sequencing technology two additional features were added: limiting dilution and sample tagging. For single-cell sequencing, the single cells or the nuclei of the single cells are first deposited into separate wells or are kept physically separated by other means, i.e. a gel. The single cells or the nuclei of cells can be deposited in a well by a cell sorter. Alternatively a collection of cells or nuclei can be diluted into a fluid that is then made semi-solid. A unique sample tag is prepared for each well for sequencing, with a separate sample tag per well. After this step, the wells are pooled and sequenced. The sample tag is used to deconvolute the information, allowing a copy number profile to be made for each individual cell.

Protocol: Isolate single tumor cells or single nuclei of the tumor cell, tag and optionally sample tag the DNA of each cell or nuclei, and amplify the tagged DNA to microgram


IPR2025-01353; IPR2025-01355
CSHL EX2002

quantities to generate tagged libraries for Solexa sequencing. The isolation of single tumor cells will occur by Fluoresent Activated Cell Sorting by total genomic DNA content, to deposit single nuclei into individual wells.

Alternatively, tumor cells can be isolated by FACS sorting with fluorescent antibodies against breast epithelium markers, including cytokeratin 8, 18 or 19. Cells are lysed *in situ* and the genome is fragmented by heating. Libraries are constructed from the fragmented genomic DNA. Preliminary experiments were conducted using high-resolution ROMA microarray to quantify whole-genome copy number on DNA amplified using WGA from single SK-BR-3 cells (Figure 5). These experiments demonstrate the feasibility and efficiency of amplifying 6 pg of DNA into microgram quantities to determine copy number by microarray methods, which suggests that the DNA is also suitable for determining copy number by sequencing.

Since sequencing whole tumor genomes from single cells is costly and time consuming, two methods are proposed to stratify the genome and sequence for copy number: Representation Sequencing and Bin Sequencing.

*Representation Sequencing*

Representation sequencing involves restriction digesting genomic DNA into fragments and sequencing only the stratified population of restriction fragments to determine copy number. This method has the advantage of generating precise restriction fragments that can be reliably prepared as libraries from different individual tumor cells. In this protocol genomic DNA is amplified from a single tumor cell, digested with the DpnII, a second digestion is preformed with a AluI enzyme to remove highly repetitive "Alu SINE elements," sequencing adaptors (Illumina GA, Solexa) with a unique 4 nucleotide sample tag are ligated and the ligated fragments are PCR amplified to generate a sequencing library. *In silico* modeling predicts that the double digest will increase the fraction of unique sequencing reads to 75% from the 55% normally seen in unselected human genome sequencing. The sample tagged DNA libraries are then pooled, and applied to the Solexa flow-cells for cluster amplification and single read sequencing. Single-end reads are used because they are faster and cheaper than paired-end reads and provide all of the information necessary for 'sequence counting'. Mapping algorithms then align each sequence read with a high quality score (with a read length in excess of 35bp) and an unambiguous sample tag to a specific location on the DpnII restriction map of the human genome. The result is a count of reads at places in the genome,

similar to the process demonstrated in Figure 4, except that the data is segregated by sample tag, so that each well receives a copy number profile.

*Bin Sequencing*

It is clear that sequencing of random size-selected fragment from the genome is a viable alternative. In this method, the DNA from single tumor cells is amplified to generate fragments from 200-2000bp. Blunt-end cloning is then used to ligate adaptors to these fragments, followed by PCR amplification to generate libraries (standard Solexa protocol for single read genomic DNA). However, in addition to this protocol adaptor sequences with 4 nucleotide sample tags are ligated to generate unique libraries from each tumor cell genome. The sample tagged DNA libraries are then pooled, and applied to the Solexa flow-cells for cluster amplification and single read sequencing. The 35bp sequences are then mapped to specific regions of the human genome and deconvoluted using the unique sample tag to determine which sequence reads originated from which tumor cells. The crucial step in this procedure that is different from representation methods is to artificially divide the human genome into "bins" of several kilobases. Break analysis in sequence count data has another degree of freedom compared to what is practiced for array data, namely that the boundaries of the bins can be adjusted so as to minimize variance. Each bin contains a number of random genomic sequence reads from which a mean value is computed to determine genomic copy number at low sequence coverage.

The particular use of the WGA method does not depend on absolutely uniform amplification across the genome. Useful data can be obtained even if only half of the potential fragments from a single cell were amplified in a given reaction. Additionally, redundancy of data gained from segmentation benefits in identifying breakpoints.

Representation Sequencing vs. Bin Sequencing

A comparison of the two methods reveals the number of samples that can be profiled per sequencing run. Of the 3,149,324 Dpn II fragments in size range 200-800 (size optimal for sequencing) there are 618,629 not cut with AluI and thus available to be amplified. Approximately 75% of the fragment ends, or about 450,000, yield unique mappings and thus provide locations for counting. The overall process is presented as a flowchart in Figure 7.

Assuming 15 million quality reads per Solexa sequencing lane, random sequencing yields about 55% or 8.25 million unique reads per lane. The DpnII/AluI representation should

increase the number of unique reads to 11.25 million. Modeling a single flow cell lane with 100kb bins distributed across the genome, the projected results are as follows:

Table 1

| # samples | avg reads/bin/sample (random) | avg.reads/bin/sample (Dpn/Alu) |
|---|---|---|
| 1 sample per lane | 267 | 363 |
| 2 samples per lane | 133 | 181 |
| 12 samples per lane | 22 | 30 |

The number of samples (individual cells) to be sequenced per lane, is between 4 and ten. With 8 lanes on a flowcell, 30-80 samples could be profiled at the level of Figure 4 relatively inexpensively.

Finally, the whole genomic copy number (generated by sequencing) will be determined for multiple, individual tumor cells in a single solid tumor. The data will be segmented into sections that minimize variance of copy number, subject to Kolmogorov-Smirnov significance tests, to generate individual tumor copy number profiles. As stated above, an additional degree of freedom, which our algorithms do not yet incorporate, is selecting the "boundary" of the bins. The single cell tumor genome profiles will be compared to study tumor heterogeneity in single solid tumors which may elucidate the progression of genomic events during tumorigenesis.

*Varietal counting of nucleic acids*



IPR2025-01353; IPR2025-01355
CSHL EX2002

Measuring the absolute or relative concentration of nucleic acid sequences in a sample often suffers from severe stochastic distortions due to differential amplifications or losses during processing and manipulation of the nucleic acids. There is never more information available than the number of molecules of the starting material in the sample itself. The method described herein is designed to capture the information at the earliest stage, in a form that is then resistant to distortion by further amplification.

Input nucleic acid molecules (0-th order derivatives), or first round copies of them (1-st order derivatives), present in the sample, can be randomly tagged using a large choice of distinct tags. Every tagged molecule becomes essentially unique, the combination of the information in the tag and the information in the nucleic acid, which can later be read by sequencing. The combination of a tag with a 0-th order, 1-st order, or subsequent N-th order nucleic acid molecule is called a "tagged nucleic acid molecule." After amplification or selection or other processing, the number of different tags is counted.

Within a pool of tagged nucleic acid molecules, each tagged nucleic acid molecule is likely unique in the pool when a sufficiently large number of distinct tags is used. Then in subsequent amplifications these unique tagged nucleic acid molecules are amplified to facilitate detection, but no new species of tagged nucleic acid molecules are created. Thus, distorted counts are not created provided that the number of different tagged nucleic acid molecules is counted.

In the disclosed method, each sequence read generated from a tagged nucleic acid molecule comprises two parts, or subsequences. The first part corresponds to the tag, which identifies the tag (and the sample tag if samples are pooled). The second part corresponds to a nucleic acid in the sample which is mapped to a location in the genome (or to a transcript if RNA molecules are being counted). After deconvoluting by sample tag, the number of different tagged nucleic acid molecules mapped to each location is counted. There are two separate methods of counting, and each can be combined for greater accuracy and determination of confidence.

The first counting method is called "maximum tagged nucleic acid molecule number." The number of different tagged nucleic acid molecules at a given location cannot exceed the true (absolute) copy number. Therefore, for a given region comprising multiple locations, the maximum copy number of the region is not less than the maximum number of different tagged nucleic acid molecules mapped to any location in that region. The maximum number

of different tagged nucleic acid molecules centered about a moving window of a fixed number of locations can be taken as the measurement of true copy number of the window. This will provide at worst an underestimate of the maximum true copy number for that window, and never an overestimate. The set of maximum number of different tagged nucleic acid molecules is one type of copy number profile.

The counting method described above is most accurate when the processing efficiency is excellent and the number of molecules mapped to each location is less than the number of tags. Under such conditions virtually each molecule in the sample is tagged and counted, and the true count can be derived. But there is a second method.

The second counting method is called "total different tagged nucleic acid molecules," and is useful, for example, when tagging efficiency is low. The total number of different tagged nucleic acid molecules within a region of a fixed number of locations will be, within a statistical counting error, a monotonic function of true copy number, regardless of the efficiency of tagging. Low efficiency of tagging merely increases the statistical error. The relative copy number of a region can be estimated by comparing the total number of different tagged nucleic acid molecules mapped to the region to, for example, the mean value of tagged nucleic acid molecules mapped to other regions of the genome which have a comparable number of locations.

If the number of molecules at a location on a genome or cDNA library exceeds the number of tags, and tagging is too efficient, copy number information can not be recovered. Rather, the total number of tags at each location would merely be measured. There are at least three ways to remedy this condition: 1) diminish the efficiency of the first step of the process, for example by reducing the reaction time; 2) reduce the amount of sample; and 3) increase the total number of tags. For example, if the length of the tag were N nucleotides, there are $4^N$ possible tags. For N = 15, there are approximately $10^9$ tags, easily in excess of the number of molecules per location.

The collection of reads lend themselves to modeling the power of inference. For each location, there exists the number of tagged nucleic acid molecules observed, and for each species of tagged nucleic acid molecule, the number of times it was observed. With a good estimate of the input, the number of genomes, critical parameters can be accurately deduced because the vast majority of the copy numbers are two per genome. The limiting case of single cell analysis lends itself to the most accurate modeling, with the proportion of null

reads enabling the inference of the probability, theta, that a single molecule at a location is detected. Theta is a function of the efficiency of tagging, the efficiency of amplification and processing, and the read depth.

Using the language invented of the WGA reaction, in the specific protocol described here, all first order derivative DNA molecules (i.e., first copies) are tagged, with no new tags introduced when subsequent order derivatives are created. This enables the "counting" of original molecules, alleles in the case of genome analysis. This method is referred to herein as "varietal counting." As long as sequence depth is sufficient, subsequent amplification or methods of trapping that enrich for regions of interest, or the combination, cannot distort the count. Pooling from multiple cells becomes a natural extension of this method, which improves efficiency, increases throughput and decreases costs.

*Tags*

The tags are nucleotide sequences with constant portions and variable portions. The length of the constant portion and the length of the variable portion may be the same or different. The length of the tags within a plurality of tags is typically, but not necessarily, the same for all tags. In an embodiment, the total length of a tag is less than 100 nucleotides.

The constant portions are used to aid in the tagging of nucleic acid molecules. For example, the constant portion may contain a nucleotide sequence which allows hybridization to a poly-N tail added by a terminal transferase. The constant portion may also be used to aid in manipulation of the tagged nucleic acid molecules. For example, the constant portion may contain a sequence which is a primer binding site useful for PCR amplification. The constant portion may also be used to identify a sample if there are multiple samples being processed in parallel, i.e., the constant portion may contain a sample tag.

The sequence of individual tags does not need to be known prior to generating tagged nucleic acid molecules. A plurality of tags may have variable portions whose sequences are elucidated for the first time during sequencing. The variable portion of a tag may be a sequence of nucleotides where each nucleotide is individually one of A, T, C, or G, or one of any two of A, T, C, and G, or one of any three of A, T, C, and G, i.e. the variable portion of a tag may consist of only two or three species of nucleotides. Tags may be designed such that the variable portion has a random sequence. Tags may be designed such that the variable portions are constructed from a set of dinucleotides or trinucleotides.

Tags may be associated with nucleic acid molecules by using, singly or in combination, ligation, nick-ligation, terminal transferase, hybridization, priming, and the like. For example, nucleic acid molecules may be cut with a restriction endonuclease(s), treated with a terminal transferase to add a poly-N tail, hybridized to a tag molecule, and then nick ligated to form a tagged nucleic acid molecule. Specific methods of generating tagged nucleic acid molecules are discussed in detail below.

*Tagging with the aid of a terminal transferase*

Figure 8 illustrates one embodiment of this method, where the $N_{10}$ designates an oligonucleotide having a tag component and a sample tag component. The first six nucleotides (4096 possible sequences), randomly generated, provide the tag component, and the last four nucleotides (256 choices) are chosen to encode the micro-well, and, hence, a sample tag. The lengths of the tag and the sample tag can be changed to suit needs, and are not limited to a total of ten nucleotides. The lengths are chosen for illustration. The sample tag allows pooling, with deconvolution occurring afterwards. The combination of a tag with a nucleic acid molecule from the input sample provides an essentially unique tagged nucleic acid molecule for the first derivative strand, that can only with very low probability be replicated by chance from another like molecule (a chance of roughly 1/N where N is the number of available tags). In the end, multiple copies of sequence-identical molecules in the input sample are each individually counted. In applications to mRNA counting, the length of the tag component can be increased to accommodate possibly thousands of transcripts from the same gene.

Although there are multiple embodiments of the basic method besides the one illustrated in Figure 8, the embodiment illustrated in Figure 8 has several useful design features. First, it avoids the use of ligases (an alternate method for associating a tag with a nucleic acid molecule), which has been found to be an inefficient reaction. Instead, this embodiment uses terminal transferase (TTase), which is a robust enzyme from which high efficiency is expected. Second, once the tag and sample tag are incorporated into the first derivative DNA strand, micro-wells can be pooled and the subsequent steps carried out more efficiently in larger volumes with larger numbers of molecules, possibly with the addition of carrier nucleic acids. Because of a clean-up step, removing free adaptor, no more unique first order derivative strands with tags can be created, and second order strands can be arithmetically expanded if found to be useful. Third, the method extends immediately from DNA to

mRNA. Rather than cleaving and adding a polynucleotide tail with terminal transferase, the naturally occurring polyA tail of mRNAs can be utilized. Fourth, it is easy to see how to combine analysis of both RNA and DNA from the same cell with this schema. Instead of tailing DNA with A, it can be tailed with, for example, C, thus distinguishing the DNA from the RNA. DNA and RNA derivatives can either be read together in the same sequencing runs, or separately by amplification using their differing PCR primer adaptors. Fifth, by using a longer sample tag, sequences belonging to a particular cell can be preferentially amplified for deeper study when warranted.

This method can be applied to assessing DNA methylation. After the first order derivative strands are made, the duplex DNA can be pooled from multiple samples. These molecules will be hemi-methylated, and still retain epigenetic marks. Therefore they can then be separated into methylated and unmethylated samples either by cleavage with methylation sensitive restriction enzymes, or by partitioning with antibodies or methyl-C binding proteins directed to methylated (or hydroxyl methylated) cytosine. Error rates will be measured based on expectations from regions of known methylation state.

*Tagging by ligation*

In another embodiment of the invention, adaptors with tag(s) can be ligated to sample DNA or RNA that has been converted to DNA. In the second step these molecules can be pooled (or not) and amplified for sequencing. In the third step the numbers of each different tagged nucleic acid molecule of each locus (or mappable sequence read) are compiled, giving statistics from which copy number can be inferred.

In another embodiment, to measure RNA, RNA is first converted to DNA using methods that produce a single double stranded molecule for each RNA in the sample. (For example, prime with oligo dT and reverse transcriptase, followed with random priming and Klenow polymerase.) The procedures for analyzing DNA can then be followed.

In another embodiment, the sample DNA can be sheared and the ends prepared for ligation. However, there are three major advantages to cleaving with one or more restriction endonucleases (RE). The first is that the expected number of tagged nucleic acid molecules at each location can be precisely predicted from the starting number of cells in the sample, or conversely the number of cells in the sample inferred from the expected number of tagged nucleic acid molecules per location. The second is that the expected complexity of the


IPR2025-01353; IPR2025-01355
CSHL EX2002

product is reduced, enabling copy number determination from fewer sequence reads. The third are certain advantages in the flexibility and functionality that can be built into the tags.

In an embodiment, sample nucleic acid is cleaved with a specific restriction enzyme and sample molecules are ligated to tagged (and sample tagged if appropriate) oligonucleotide adaptors which contain sequences that allow PCR amplification. There are various designs for these adaptors, and also alternatives to PCR (such as WGA), based on the functionalities built into the tags. Some of these additional functionalities are discussed below.

The ligated molecules now are essentially unique, the ligated molecules each being the combination of a substantially unique tag and DNA sequence to which they were ligated. Further amplification may distort the yields of molecules, but will not create new tags (although some may be lost by statistical processes which are readily modeled during analysis).

Through sequence error during processing, the tag may very occasionally mutate, creating the appearance of a new tagged nucleic acid molecule. But if the set of tags are sufficiently large, the tagged nucleic acid molecules will generally differ at more than one nucleotide, and so tagged nucleic acid molecules that differ by a single nucleotide can (optionally) be ignored when counting.

For single cells, individual cells are loaded at most one per microwell. The nucleic acid is then prepared in wells, and the molecules tagged. Once tagged they can be pooled for further processing, which greatly improves the uniformity and efficiency of DNA amplification methods.

In yet another embodiment, DNA is cut, and adaptors are ligated to both ends of each fragment, enabling them to be amplifiable by PCR. To avoid adaptor-adaptor self ligation, double stranded adaptors without 5' phosphorylation can be used, so an adaptor is ligated on one strand only at each end of the cleaved sample DNA. Then the molecules are pooled, and treated as for representations: the shorter single strand adaptor is melted off at elevated temperature but well below the Tm for the sample duplex, the ends of the double stranded DNA filled in with Klenow polymerase, unligated adaptors removed, and then the sample is PCR amplified.

In the preceding embodiment, adaptors are used not only for tagging but also for amplification. Adaptors are for identifying each molecule individually. They can also be used

for amplification, as described above, or for other uses (see below). Because in the above embodiment adaptor ligation at both ends of sample fragments is required for amplification, the yield may be diminished. To improve yield, the method can work with molecules that have an adaptor at one end only.

In one embodiment, the reactions are performed as before, but tags do not contain sequences enabling PCR. Prior to PCR, but after an optional pooling, fresh adaptors (different composition and without tags) are ligated so that molecules with only a single tag adaptor can be amplified and eventually counted. If the original tag has a selection sequence, molecules with the tag adaptor before sequencing can be optionally enriched by hybrid selection.

In another embodiment, initial adaptors are designed with a phosphorylated 5' end and a 3' dideoxy end, so that the initial ligation occurs to the opposite strand as described in the embodiment above. (Instead of using an RE with a 5' overhang, one which leaves a 3' overhang may be preferred.) After removal of unligated adaptors, there follows several rounds of arithmetic amplification, followed by tailing molecules with dT using terminal transferase, and amplification of the processed sample DNA by PCR.

In another embodiment, DNA from sample is labeled as before, pooled if required, and unligated adaptors removed from the reaction. The reaction is subject to whole genome amplification (WGA), and the product of WGA is used to make sequencing libraries. For this embodiment, the initial adaptors may also contain: (A) cleavage sites for ease of making sequencing libraries; and (B) a sequence large enough for optional enrichment by hybrid selection prior to making libraries for sequencing.

In the embodiments just described, it is assumed that the tag is double stranded. The production of a great variety of double stranded molecules is not completely straightforward. In one embodiment, single strand DNA is synthesized as type I and type II. Type II has three sets of functionalities A, B, C, and D. Functionalities B contain the tag sequence, C the adhesive ends with a RE site, and D the buffer. The B will contain some length of runs with random sequence, for example 10 nucleotides with a total complexity of $4^{10}$, or about $10^6$ possible sequences. Type I oligo will be complementary to type II at functionality A, the two joined and annealed, and then extended with polymerase to create a double stranded molecule. After cleavage with the RE site, the resulting double stranded molecule is the tag used for the subsequent tagging of the sample.

Additional methods for producing single and double stranded oligonucleotides suitable for use as tags are known in the art, for example, in U.S. Patent Nos. 5,639,603 and 7,622,281, U.S. Patent Application Publication No. 2006/0073506 and Parameswaran et al. (2007) (7,5,8,4).

*Nick-ligation*

In this embodiment of the invention, 3'OH ends of nucleic acid molecules, e.g., segments of genomic material or mRNA transcripts, are first extended with terminal transferase, adding a poly-N tail. Primer adaptors are then annealed to the poly-N tails and the resulting duplex with a nick is sealed with DNA ligase. This series of reactions is much more efficient than a single step double stranded ligation. The resulting product is amplified by PCR. This method has been successfully used with as little as 1ng of DNA. Tagged nucleic acid molecules generated using the nick-ligation method can be generated from a single cell DNA, and then pooled prior to PCR. Pooling from 100 cells yields the DNA concentration needed for successful amplification. Sample tags can be added in the nick ligation step, and this is readily achieved (see Figure 9).

*Technology for inexpensive assay of regional markers*

The above description is of methodology designed to measure copy number and mutational loads throughout much of the genome in single cells. But the application of discoveries, once made, of predictive markers for outcome or therapeutic response will not require whole genome information. Our present outcome markers are in fact copy number markers at very specific loci. Essentially the same methodology as disclosed hereinabove can be employed to make these assessments very affordable. Consider, for example, the case where it is desirable to assess a large study for the predicative power of a set of N markers, and these markers are copy number markers that can be assessed by varietal counting. In principle, sample coding, pooling, and parallel capture and sequencing can be used to assay a large number of samples in a single sequencing lane. The major difference is that we are not restricted to minute quantities of DNA, so WGA is not needed, and assays will be more robust. We have shown in our autism work that excellent copy number data can be obtained from read depth after capture. Thus we can use direct library preparation with just sample tags. The target area for capture is likely to be different and smaller than the exome. Thus we can pool more samples per sequencing lane, and the sequencing costs for assaying hundreds of markers in hundreds of samples can be reduced. Of course, there are costs associated with the development and

testing of the capture reagents, and the cost of the capture itself, and these are not negligible, but if a test were to come into widespread clinical use, it would be inexpensive, and this is one of our major objectives. There is a clear alternative, namely creating a bank of region specific PCR primers with tags and sample tags. The tags would be necessary because random unique ends created by shearing would not be available for counting based methods. This alternative plan could be less expensive and more robust than capture (see Figure 10).

*Single cell RNA measurements*

Several methods have now been published whereby researchers have profiled the RNA content of single cells by sequencing (Tang et al., 2009). This is potentially a very powerful method for analysis of tumor cells. A varietal counting method could be very useful for RNA profiling from single cells, and for the same reasons that it is useful with DNA. When variety is not introduced by nucleic acid processing varietal counting provides a way to avoid the distortions created by PCR. RNA and DNA can be analyzed from the same cells. The DNA analysis enables us to cluster cells into stromal and tumor subpopulations, and this in turn facilitates the interpretation of the RNA profiling. As an example of utility, suppose one is interested in searching thousands of circulating cells, or cells from a biopsy, or from the margins of a tumor, for rare tumor cells. DNA copy number can facilitate its identification, and the RNA from that cell can then be reliably used to infer additional properties, such as the tissue of origin. This combination is especially powerful for the early detection of the new incidence or recurrence of cancer.

**Examples**

*Methylation sequencing*

Currently, high throughput sequencing requires physical purification of sub-genomic regions of interest in order to obtain high coverage of those regions. The method of sequence capture has been applied to examine methylation in large sets of CpG islands at the sequence level. Methylation sequencing is based on the conversion of C residues to T using bisulfite reagents, with methylated C's detected as those protected from C to T conversion. This means either that sequence capture must take place before bisulfite treatment, which requires more starting DNA than is available for patient specimens, or that sequence capture must be designed to accommodate the C to T conversion. The bisulfite based sequencing method used herein adopts the latter and so is suitable for high throughput single molecule


IPR2025-01353; IPR2025-01355
CSHL EX2002

sequencing by either the Illumina or Roche 454 methods, using as little as 100 ng of input DNA. The sequencing method depends on hybrid capture of bisulfite treated DNA followed by sequencing through the Illumina Solexa sequencer and mapping back to the genome using an algorithm developed Dr. Andrew Smith (19). It was observed that CpG islands can exhibit multiple different levels of methylation, from imprinted (allelic) to randomly partially methylated, to completely methylated or unmethylated. In particular, one observation that is important for this invention is that many islands switch from completely unmethylated in normal tissues to virtually completely methylated in matched tumor tissue ("DNA-methylation switch domains"). Figure 3 shows three examples of the simplest cases in normal cells versus the breast cancer cell line MDA-MB-231: stably unmethylated (A and B); the ALX gene, switched from unmethylated to completely methylated (C and D); and a case, SSTR4, where only a segment of the island is switched (E and F). This complete switch makes it possible to detect rare tumor cell DNA in bodily fluids or other cell populations using nested PCR of bisulfite treated DNA and opens up the possibility of a creating a method for the early detection of recurrence in the form of circulating metastatic cells.

*Comparative genomic hybridization (CGH) by high throughput sequencing*

Until now the most economical method for high resolution genomic analysis has employed some form of microarray. The rapid rise of 'next generation' DNA sequencing is changing that landscape. Both the Illumina (Solexa) and Roche 454 instruments can provide millions of individual sequence reads randomly distributed across any input DNA (Craig et al., 2008). The depth of coverage of these technologies depends on the total length of the input DNA, from whole genome (large breadth, lower depth) to narrowly focused regions achieved by hybrid capture of specific sequences (narrow breadth, higher depth) (Albert et al., 2007; Hodges et al., 2007; Okou et al., 2007). Converting sequencing reads to copy number analysis is a matter of mapping reads and then put them into 'bins' of a certain size, depending on the purpose, counting up the numbers in the bins and performing appropriate statistical analysis to define copy number change breakpoints.

Sequence counting was applied to the degraded formalin-fixed paraffin-embedded (FFPE) DNA that was available from retrospective clinical trials with very encouraging results. Figure 4 displays a comparison of microarray and sequence counting results in a region around Her2 on chromosome 17 for a 10 year old paraffin-embedded clinical trial sample JZ33. The gray and orange points are (respectively) the raw ratio data for 244K Agilent

microarray and one lane of Solexa sequencing reads divided into 30K bins. Note that the sequencing (orange) track shows significantly less noise. The green and blue lines are the result of segmenting each data type. The segments are generally concordant, but the sequence counting method reveals a somewhat more detailed pattern. Another feature of note from Figure 4 is the robustness of our segmenting program, irrespective of the variation of noise in the two sets of raw data.

The efficacy of sequence counting on FFPE DNA and the rapidly decreasing cost of DNA sequencing in general provides an opportunity to convert our previous microarray methods to a highly flexible and scalable platform for breast cancer genome diagnostics. Both the breadth of DNA to be sequenced and the size and location (and therefore resolution) of individual bins can be easily adjusted to fit the particular loci to be assayed.

*Single cell microarray and sequencing*

A representation of DNA amplified from single cells was used for genomic profiling by 'sequence counting'. As a preliminary step in developing that technology, DNA was amplified from batches of 1, 10 and 100 cells of the cell line SKBR3 using the Whole Genome Amplification (WGA) Kit from Sigma-Aldrich and compared these amplified representations to equivalent microarrays on the same format done with a standard SKBR3 DNA prep. Individual cells were picked from adherent culture dishes using a cell transfer micropipette. After random WGA1 amplification the DNA preps were digested with DpnII and AluI, adaptors added and only DpnII fragments without AluI sites were amplified by PCR. These representations were labeled and applied to ROMA custom microarrays with 390,000 probes designed to detect DpnII fragments of appropriate size across the genome. The idea was to test how reproducible this two step process would be when only 2-6 chromosomes are present in each single cell preparation (SKBR3 is roughly triploid).

The results are shown in Figure 5, presenting a comparison of the segmented profiles of a batch preparation of SKBR3 DNA (Figure 5A) against our normal male reference cell line (SKN-1) and three independent single cell DNAs (Figure 5 B,C,D). A visual inspection of these profiles makes it clear that after segmentation, the results for single cells are virtually identical to a standard DNA, leading us to have confidence that sequencing DNA from a representation will yield useful results. Although the raw ratio data is not presented here, we note that the standard deviation of the ratio data is much higher for the three single cell experiments (0.706, 0.725, 0.676) than for a batch DNA preparation (0.313). The extra noise

does not adversely affect segmentation, although in certain regions very fine detail may be lost (red arrow in Figure 5A). This is comparable to other WGA1 experiments such as the one depicted in gray in Figure 4. In this light it is also worth noting that our particular use of the WGA1 method does not depend on absolutely uniform amplification across the genome. We benefit from the redundancy of data that we gain from segmentation to identify breakpoints. Even if only half of the potential fragments from a single cell were amplified in a given reaction, we would still obtain useful data.

*Sequencing across chromosome breakpoints*

One of the methods for the detection of rare tumor cells, in lymph nodes, blood or other bodily fluids is the amplification or sequencing of unique breakpoints resulting from deletions and/or translocations that occur during breast tumor development. Although events such as deletions and duplications are apparent from array CGH, the actual structure underlying the copy number changes is not.

The DNA sequences surrounding 22 breakpoints from two breast cancer cell lines were enriched by hybrid capture on custom microarrays as described by Hodges et al (20). Following hybrid capture, the enriched DNA was amplified and prepared for 'paired end' sequencing on the Illumina GA2 instrument. Paired end sequencing yields approximately 36 bp of sequence from each end of a single DNA molecule (after preparation for sequencing, each molecule is approximately 200 bp in length), and those sequences are then mapped back to the consensus genome. For a molecule from a normal, non-rearranged region the two matched reads should map approximately 130 bp apart. In contrast, at the edges of a deletion, the two reads from a single molecule would be separated on the map by the width of the deletion, or in the case of a translocation, would map to another chromosome. These are referred to as 'bridging fragments' because they connect regions not normally adjacent to one another. Examples of two such bridging fragments are shown at the top of Figure 6, along with the original CGH pattern, and graphs of the distance between paired ends (blue) and the total number of reads (red) across the capture region (X axis). By mapping the exact location of a breakpoint, highly sensitive and specific probes that will distinguish tumor cells from large numbers of normal cell can be prepared. Of 22 regions displaying deletions captured from cell lines T47D and MDA-MB-436, 15 behaved as simple deletions from which unique 'bridging fragments' could be identified as in Figure 6. Two additional sites 'jumped' to different chromosomes indicating translocations. The other 5 sites involved repeated

sequences on one of the paired for which the identification of the exact breakpoint was not possible and which would not be suitable for creating unique probes.

*Copy number profiling using varietal tags on restriction enzyme fragments.*

As one illustration of varietal counting, copy number information was obtained using a "total different tagged nucleic acid molecules method.", DNA fragments were first associated with tags by using terminal transferase and nick-ligation (Figure 9). Genome DNA was obtained from the cell line SKBR3. DNA (2μg) was digested by NlaIII, resulting DNA fragments with 4bp, "CATG", 3' overhangs. This 3'OH end was extended with terminal transferase by adding a poly-A tail. One of primer adaptors was designed to have tags, 12 Ts and CATG at the 3' end, so it could be annealed to the 3' ends of the DNA fragments with a nick site (Figure 9, Product 3). This nick site was ligated by DNA ligase, forming duplex DNA fragments with tags. Another primer adaptor was then added to these duplex DNA fragments by one cycle of PCR (Figure 9, Step 5), allowing the resulting products of this reaction to be further amplified by PCR, followed by sequencing (Illumina/Solexa). Then, by mapping the subsequence of each tag associated sequence read to a location on the genome, the number of DNA fragments (assigned to the same genomic location) with different tags was counted.

We have previously obtained copy number profiles of SKBR3 genomic DNA from a million cell library, by sequencing and sequence counting. Figure 11 compares the use of varietal counting (a "total different tagged nucleic acid molecules method") with 'sequence counting' for analyzing copy number data, genome-wide (Figure 11A) and partial-genome (Figure 11B). The essentially equivalent result obtained using varietal counting illustrates that by adding varietal tags to DNA fragments before PCR one can use varietal counting as described herein to reliably obtain copy number information unaffected by amplification distortion.

Data analysis for the procedure described above employed a modification of the 'bin-counting' method because the NlaIII restriction sites are not distributed evenly in the genome. Therefore, in order to ensure equal weight for each of the 50K bins, the bin boundaries for this experiment were designed to encompass and equal number of NlaIII fragments sized between 100bp and 250bp, the size range from which the Illumina libraries were made.

**Discussion**



IPR2025-01353; IPR2025-01355
CSHL EX2002

Single cell data is much more informative than data obtained from multiple cells, for example from array hybridization of multiple sections of a primary tumor. First, one can afford to analyze many more genomes by multiplex single cell sequencing than by array hybridization (on the order of a hundred single cells versus a dozen sections). Second, one is able to directly assess the exact proportion of cells in a region of a tumor that are malignant. Third, one is able to make a far more detailed analysis of the temporal order of genetic lesions, since one will not be looking at the average state of a population. Fourth, one is able to assess whether tumor heterogeneity is of the anatomically segregated or the co-habitation type, or both.

The invention described herein may be used to determine the original genome copy number (or the gene expression profile) of a sample without amplification distortion. All measurements of copy number or gene expression, especially involving multiple samples, are potential applications. Samples can be nucleic acids derived from single cells or populations of cells. The invention can be used to study cellular heterogeneity in tumors or normal populations of cells.

The invention may be used to generate a genomic profile of a cell, e.g., a cancer cell, or population of cells. From the genomic profile it can be determined if specific genes are amplified or deleted. When the sample is derived from expressed RNA, and the regions are expressed genes, the relative counts comprise an expression profile. From the expression profile the tissue of origin of a cancer cell can be identified. When the sample is a mixture of genomes from different species, and the genomic regions distinguish the different species, the relative counts comprise a census of those different species. For example, from a census we can determine the population of microbial flora and relate the same to disease.

One of the advantages of the use of tags is that it allows the use of restriction endonuclease fragments in genome analysis. Counting methods for genome copy number determination use counts of independent reads. Two reads are not considered independent if they are the result of DNA amplification occurring during sample processing. Typically, this entails shearing sample DNA fragments, and the position of ends of the fragments are used to disambiguate PCR duplicates. If two DNA reads have the same ends, they could be PCR duplicates, and not independent, so they are counted as only one true read. One cannot disambiguate restriction fragments this way, because two reads of the same fragment will have the same ends. However, by the addition of tags, independent reads can be distinguished, if they are

added before any amplification step, as they will have different tags. By contrast, if two reads of the same genome region have the same tag, they are highly likely to be PCR duplicates, and not independent.

The ability to use sequence of restriction fragments in genome analysis can be exploited to reduce the cost of determining copy number profile, much as performing ROMA facilitates genome analysis. By size selection of fragments, one can sample the genome at greater depth at fewer loci, and still obtain a high resolution copy number profile but with fewer reads and therefore at lower cost. This is the essence of representation, and results in complexity reduction, the central element of ROMA. In ROMA, the analysis of the representation is done by hybridization to arrays, and here the representation is analyzed by sequencing.

Further complexity reduction can be achieved by sequence capture. The sequence capture can be designed specifically for the fragments that can be predicted by in silico methods to map to the regions where one wants copy number mutation. In this way one can target special regions of the genome, such as the exome, or known cancer genes. Very little sequence of the sample is then needed for copy number determination, and samples can be multiplexed at a great cost saving.

The invention disclosed herein may also be used to determine the proportion of components in a mixture. For example, suppose one possesses a set of reagents, such as antibody molecules, that bind to a set of substrates, such as different proteins, and one desires to detect the relative proportion of those substrates in a sample, such as blood. First, one covalently combines a tag having a sample tag with each reagent, where the sample tag portion identifies the reagent, and the varietal tag will be used in counting. The tagged reagents are then combined with the sample, and reagents that combine with substrates are physically separated. For example, the substrates may be biotinylated before being mixed with reagents, then reagents combining with substrates separated by avidin affinity chromatography. The tags are amplified by PCR and sequenced, effectively yielding relative counts for each reagent species and therefore each substrate. In this case, the constant portions of the tags include the sample identifier and nucleotides used for amplification.

The invention provides the following advantages: (1) the method is simple; (2) with statistical analysis the confidence of a measurement can be determined with great accuracy; (3) this is a sequence based method, which in the long run will be cheaper than arrays; (4) long read lengths are not necessary, providing savings; (5) because tagged nucleic acid molecules are

counted, many processing methods that create distortion in yields can be applied without major loss in the tagged nucleic acid molecule count; (6) among these processing methods are hybrid selection (aka capture microarrays), which restrict the sequence information to regions of importance, and thus reduce cost by reducing the total number of reads per sample needed for determination of the desired information; (7) the molecules from one sample can be combined with tags and a sample tag, creating sample tagged nucleic acid molecules which can then be pooled, and after processing and sequencing, the sample tags decoded, giving a profile for each sample; (8) pooling adds efficiency and uniformity; (9) the samples can be individual cells, yielding large numbers of single cell measurements cheaply.

The invention described herein has clinical applications as described above, including: (1) determination of the genomic heterogeneity of tumors; (2) detection of tumor cells in needle biopsies, in particular an assessment of the heterogeneity of the cells in a biopsy; (3) assessment of the spread of a cancer to regional lymph nodes; and (4) detection of malignant cells in sites such as blood and bone marrow. The presence of cells with the molecular signature of the primary in the blood or bone marrow opens up the possibility of using more general single cell/molecule assays as a first line of defense in the early detection of cancer, i.e. breast cancer in women at risk.

When there is radiological evidence of a breast mass, the patient most commonly undergoes fine needle biopsy, which is examined by the pathologist who makes an assessment of the malignancy of the mass. Because needle aspiration destroys the histology of the biopsy, one of the pathologist's canonical tools for assessing malignancy is not available. Single cell sequencing would not only more accurately predict the presence of malignant cells, but also give what we might call "genomic staging." Is the tumor a simplex type of complex type? Is it mongenomic or polygenomic? Are there critical loci that are amplified or deleted that might immediately suggest a preoperative therapy?

The single cell/molecule approach allows the assessment of the spread of cancer to regional lymph nodes and to establish the DNA-methylation states at switch regions.

Successful application of single cell/molecule detection in the above contexts will lead to application to the detection of cancer recurrence, and the measurement of cancer load. Detection of recurrence at the earliest possible time may prolong survival for a large class of patients by earlier intervention than they would otherwise receive. Moreover, if cancer load were monitored in the patient, even before the site of the cancer growth is discernible by


IPR2025-01353; IPR2025-01355
CSHL EX2002

imaging, then one could measure the response of the cancer to a variety of therapeutic measurements. That is, the most effective choice of therapy could be determined empirically in the patient.

Additional applications exist in the diagnosis and selection of treatment for patients with auto-immune diseases, and in vitro fertilization (IVF) screening of pre-implant embryos for genetic defects such as trisomy-21.



IPR2025-01353; IPR2025-01355
CSHL EX2002

**References**

The disclosures of the following references in their entireties are hereby incorporated by reference into this application.

1. Pinard R, de Winter A, Sarkis GJ, Gerstein MB, Tartaro KR, Plant RN, et al. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics. 2006; 7:216. PMCID: PMC1560136.

2. Klein, C. et al. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. PNAS. 1999; 96, pp. 4494–4499.

3. Stoecklein, N. et al. SCOMP Is Superior to Degenerated Oligonucleotide Primed-Polymerase Chain Reaction for Global Amplification of Minute Amounts of DNA from Microdissected Archival Tissue Samples. American Journal of Pathology. 2002; 161(1):43-51.

4. Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130.

5. U.S. Patent No. 7,622,281, Ronaghi et al.

6. Eid et al. Real-Time DNA Sequencing from Single Polymerase Molecules. Science. 2009; 323:133-138.

7. U.S. Patent No. 5,639,603, Dower et al.

8. U.S. Patent Application Publication No. 2006/0073506, Christians et al.

9. Navin N, Krasnitz A, Rodgers L, Cook K, Meth J, Kendall J, et al. Inferring tumor progression from genomic heterogeneity. Genome Res. 2010;20(1):68-80. PMCID: 2798832.

10. Fisher B, Redmond CK, Fisher ER. Evolution of knowledge related to breast cancer heterogeneity: a 25-year retrospective. J Clin Oncol. 2008;26(13):2068-71.



IPR2025-01353; IPR2025-01355
CSHL EX2002

11. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25. PMCID: PMC2690996.

12. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4(4):406-25.

13. De Baetselier P, Roos E, Brys L, Remels L, Gobert M, Dekegel D, et al. Nonmetastatic tumor cells acquire metastatic properties following somatic hybridization with normal cells. Cancer Metastasis Rev. 1984;3(1):5-24.

14. Duelli DM, Padilla-Nash HM, Berman D, Murphy KM, Ried T, Lazebnik Y. A virus causes cancer by inducing massive chromosomal instability through cell fusion. Curr Biol. 2007;17(5):431-7.

15. Jorgensen HF, Adie K, Chaubert P, Bird AP. Engineering a high-affinity methyl-CpG-binding protein. Nucleic Acids Res. 2006;34(13):e96. PMCID: PMC1540740.

16. Meehan RR, Lewis JD, Bird AP. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. Nucleic Acids Res. 1992;20(19):5085-92. PMCID: PMC334288.

17. Huang J., Pang J., Watanabe T., Ng HK, Ohgaki H. Whole genome amplification for array comparative genomic hybridization using DNA extracted from formalin-fixed, paraffin-embedded histological sections. J Mol Diagn. 2009 Mar;11(2):109-16. Epub 2009 Feb 5.

18. Talseth-Palmer BA, Bowden NA, Hill A, Meldrum C, Scott RJ. Whole genome amplification and its impact on CGH array profiles. BMC Res Notes. 2008 Jul 29;1:56.

19. Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L, McCombie WR, Wigler M, Hannon GJ, Hicks JB. High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res. 2009 Sep;19(9):1593-605. Epub 2009 Jul 6.



IPR2025-01353; IPR2025-01355
CSHL EX2002

20. Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ, McCombie WR. Genome-wide in situ exon capture for selective resequencing. Nat Genet. 2007 Dec;39(12):1522-7. Epub 2007 Nov 4.

21. genome.ucsc.edu/index.html?org=Human&db=hg19&hgsid=171216665

22. Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S.. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135.

23. U.S. Patent No. 7,537,897, issued May 26, 2009 (Brenner et al.).

24. McCloskey M.L., Stoger, R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767.



IPR2025-01353; IPR2025-01355
CSHL EX2002

What is claimed is:

1. A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises one segment of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

2. The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome by assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.


IPR2025-01353; IPR2025-01355
CSHL EX2002

3. The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. The method of any one of claims 4-5, wherein the second region of the genome comprises a centromere.

7. The method of claim 1, further comprising summing the counts obtained in step (f) for locations which comprise a region of the genome, and comparing the sum to a sum

obtained from a reference sample for the same region of the genome, thereby estimating a relative genomic copy number of the region of the genome.

8. A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA;

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, such that each tagged nucleic acid molecule comprises

i) one cDNA from step (a) or one segment of the cDNA from step (b), and

ii) a tag;

d) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

e) generating tag associated sequence reads by sequencing the product of step (d);

f) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a cDNA or a segment of cDNA to a location on the cDNA library; and

g) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

9. A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:


IPR2025-01353; IPR2025-01355
CSHL EX2002

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,


IPR2025-01353; IPR2025-01355
CSHL EX2002

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. A method for obtaining from genomic material DNA methylation information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands to a polymerase reaction in the presence essentially unique primers, wherein the essentially unique primers are capable of hybridizing to the polynucleotide tail of the zero-th order strands, under conditions that promote the formation of only one complement, thereby generating tagged nucleic acid molecules;

d) separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules;

e) subjecting each group of step (d) to amplification by polymerase chain reaction (PCR);

f) generating tag associated sequence reads by sequencing the product of step (e);

g) assigning each tagged nucleic acid molecule to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the genomic material to a location on the genome; and

h) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,


IPR2025-01353; IPR2025-01355
CSHL EX2002

thereby obtaining DNA methylation information unaffected by amplification distortion.

11. The method of any one of claims 1-7, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. The method of any one of claims 1-7, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. The method of any one of claims 9-12, wherein adding a polynucleotide tail comprises the use of a terminal transferase.

14. The method of any one of claims 1-7, wherein tagging the segments to generate tagged nucleic acid molecules comprises ligation of adaptors comprising the tags to at least one end of the segments of the genomic material.

15. The method of claim 14, wherein the adaptors comprising the tags are ligated to only one end of the segments of the genomic material.

16. The method of any one of claims 1-15, wherein the tags comprise a sequence that aids PCR amplification.

17. The method of any one of claims 1-16, wherein each tagged nucleic acid molecule comprises one tag.

18. The method of any one of claims 1-17, wherein each tagged nucleic acid molecule comprises more than one tag.

19. The method of any one of claims 1-7 or 10-18, wherein segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

20. The method of claim 8, wherein segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

21. The method of any one of claims 8-9, 13, 17-18 or 20, wherein the maximum copy number of a location on a cDNA library is not less than the number of tagged nucleic

acid molecules having a different tag that have been assigned to the same location on the cDNA library.

22. The method of any one of claims 1-7 or 10-19, further comprising analyzing mRNA copy number.

23. The method of any one of claims 10, 13, 16-19, or 22, wherein separating the tagged nucleic acid molecules into a group consisting of hemi-methylated tagged nucleic acid molecules and a group consisting of unmethylated tagged nucleic acid molecules is by cleavage with methylation sensitive restriction enzymes, partitioning with antibodies, or partitioning with methyl-C binding proteins directed to methylated or hydroxymethylated cytosine.

24. The method of any one of claims 1-23, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or prior to sequencing.

25. The method of any one of claims 1-24, wherein each tagged nucleic acid molecule differs at more than one nucleotide.

26. The method of any one of claims 1-25, wherein the tag sequences further comprise a sample tag.

27. The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to sequencing.

28. The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to sample tag.

29. The method of any one of claims 1-28, wherein the tagged nucleic acid molecules are generated from a single species.

30. The method of claim 29, wherein the tagged nucleic acid molecules are generated from a single organism.

31. The method of any one of claims 1-30, wherein tagged nucleic acid molecules are generated from a single cell.

32. The method of claim 31, wherein the single cell is from a needle aspirate of suspected cancer lesions.

33. The method of claim 31, wherein the single cell is from a core biopsy of suspected cancer lesions.

34. The method of claim 29, wherein the tagged nucleic acid molecules are generated from two or more organisms.

35. The method of claim 34, wherein the tagged nucleic acid molecules are generated from a single cell of each organism.

36. The method of any one of claims 30-35, wherein the single species is human.

37. The method of any one of claims 1-28, wherein the tagged nucleic acid molecules are generated from two or more species.

38. The method of claim 37, wherein the tagged nucleic acid molecules are generated from a population of microbes.

39. The method of claim 38, comprising comparing the genomic copy number information obtained for different species of the population to determine the relative count of those different species in the population.

40. A composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of the genomic material;

b) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

c) subjecting the zero-th order derivative strands of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

d) adding a polynucleotide tail to the first order derivative strands;

e) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (c) and (e) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

41. A composition of matter derived from genomic material comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining segments of genomic material;

b) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

c) subjecting the zero-th order derivative strands of step (b) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (b) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

d) subjecting the product of step (c) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (b) under conditions which promote the formation of only one complement, wherein the primers of step (d) have different nucleotide sequences than the primers of step (c), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (c) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.



IPR2025-01353; IPR2025-01355
CSHL EX2002

42. A composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining cDNA from the mRNA transcripts;

b) optionally obtaining segments of the cDNA; and

c) tagging the cDNA or segments of cDNA with substantially unique tags to generate tagged nucleic acid molecules, wherein each tagged nucleic acid molecule comprises

iii) one cDNA from step (a) or one segment of the cDNA from step (b) and

iv) a tag.

43. A composition of matter derived from mRNA transcripts comprising tagged nucleic acid molecules, said tagged nucleic acid molecules being produced by a process comprising:

a) obtaining mRNA transcripts;

b) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (c) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,


IPR2025-01353; IPR2025-01355
CSHL EX2002

wherein the primers of at least one of steps (b) and d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

44. The composition of matter of any one of claims 40-41 or 43, wherein adding a polynucleotide tail comprises the use of a terminal transferase.

45. The composition of matter of any one of claims 40-44, wherein each tagged nucleic acid molecule comprises one tag.

46. The composition of matter of any one of claims 40-44, wherein each tagged nucleic acid molecule comprises more than one tag.

47. The composition of matter of any one of claims 40-41 or 44-46, wherein segments of the genomic material are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

48. The composition of matter of any one of claims 42-46, wherein segments of the cDNA are produced by restriction endonuclease digestion, mechanical shearing, heating, or sonication.

49. The composition of matter of any one of claims 40-48, wherein the complexity of the tagged nucleic acid molecules is reduced by hybrid capture.

50. The composition of matter of any one of claims 40-49, wherein each tagged nucleic acid molecules differs at more than one nucleotide.

51. The composition of matter of any one of claims 40-50, wherein the tag sequences further comprise a sample tag.

52. A composition of matter comprising at least two pools of tagged nucleic acid molecules according to claim 51.

53. The composition of matter of any one of claims 40-52, wherein the tagged nucleic acid molecules are generated from a single cell.

54. The composition of matter of claim 53, wherein the single cell is from a needle aspirate of suspected cancer lesions.


IPR2025-01353; IPR2025-01355
CSHL EX2002

55. The composition of matter of claim 53, wherein the single cell is from a core biopsy of suspected cancer legions.

56. The composition of matter of any one of claims 40-52, wherein the tagged nucleic acid molecules are generated from two or more organisms.

57. A kit for determining nucleic acid copy number information unaffected by amplification distortion comprising:

a) a terminal transferase, and

b) plurality of substantially unique primers,

wherein the substantially unique primers comprise substantially unique tags, and wherein the substantially unique primers are capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the terminal transferase.

58. The kit of claim 57, further comprising a DNA polymerase having 3'-5' proofreading activity.

59. The kit of any one of claims 57-58, wherein the plurality of substantially unique primers comprises $10^n$ primers, wherein n is an integer from 2 to 9.

60. The kit of any one of claims 57-59, further comprising a second terminal transferase and a primer capable of hybridizing to the polynucleotide tail of a nucleic acid molecule which has a polynucleotide tail added by the second terminal transferase.

61. The kit of any one of claims 57-60, wherein the tags are six nucleotides long.

62. The kit of any one of claims 57-60, wherein the tags are 15 nucleotides long.

63. The kit of any one of claims 57-62, further comprising sample tags.

64. The kit of claim 63, wherein the sample tags are 2 nucleotides long.

65. The kit of claim 63, wherein the sample tags are 4 nucleotides long.

66. The kit of any one of claims 63-65, wherein the sample tags comprise a sample tag set.

67. The kit of any one of claims 63-66, wherein the substantially unique tags comprise the sample tags.



IPR2025-01353; IPR2025-01355
CSHL EX2002

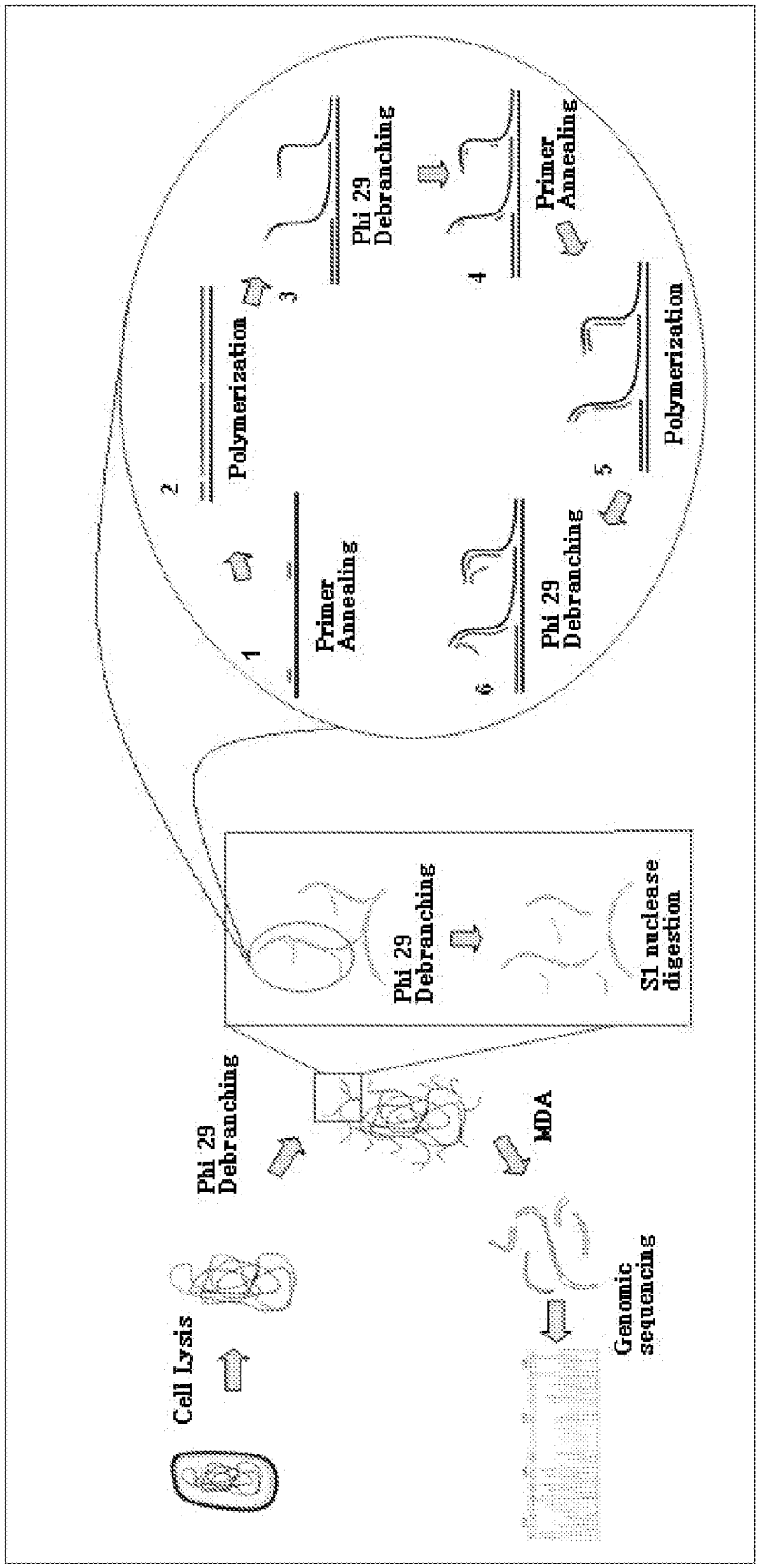



Fig. 1


IPR2025-01353; IPR2025-01355
CSHL EX2002

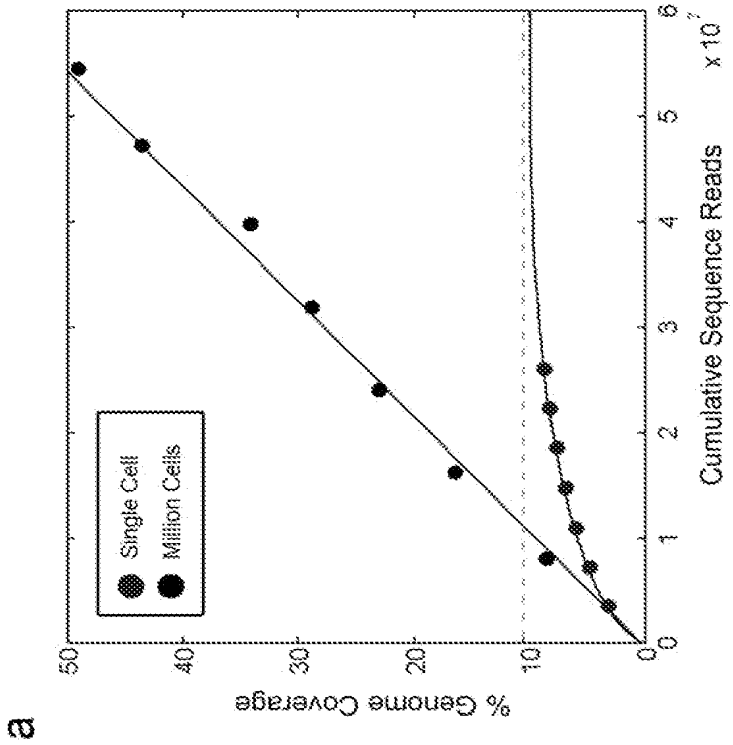



Fig. 2


IPR2025-01353; IPR2025-01355
CSHL EX2002

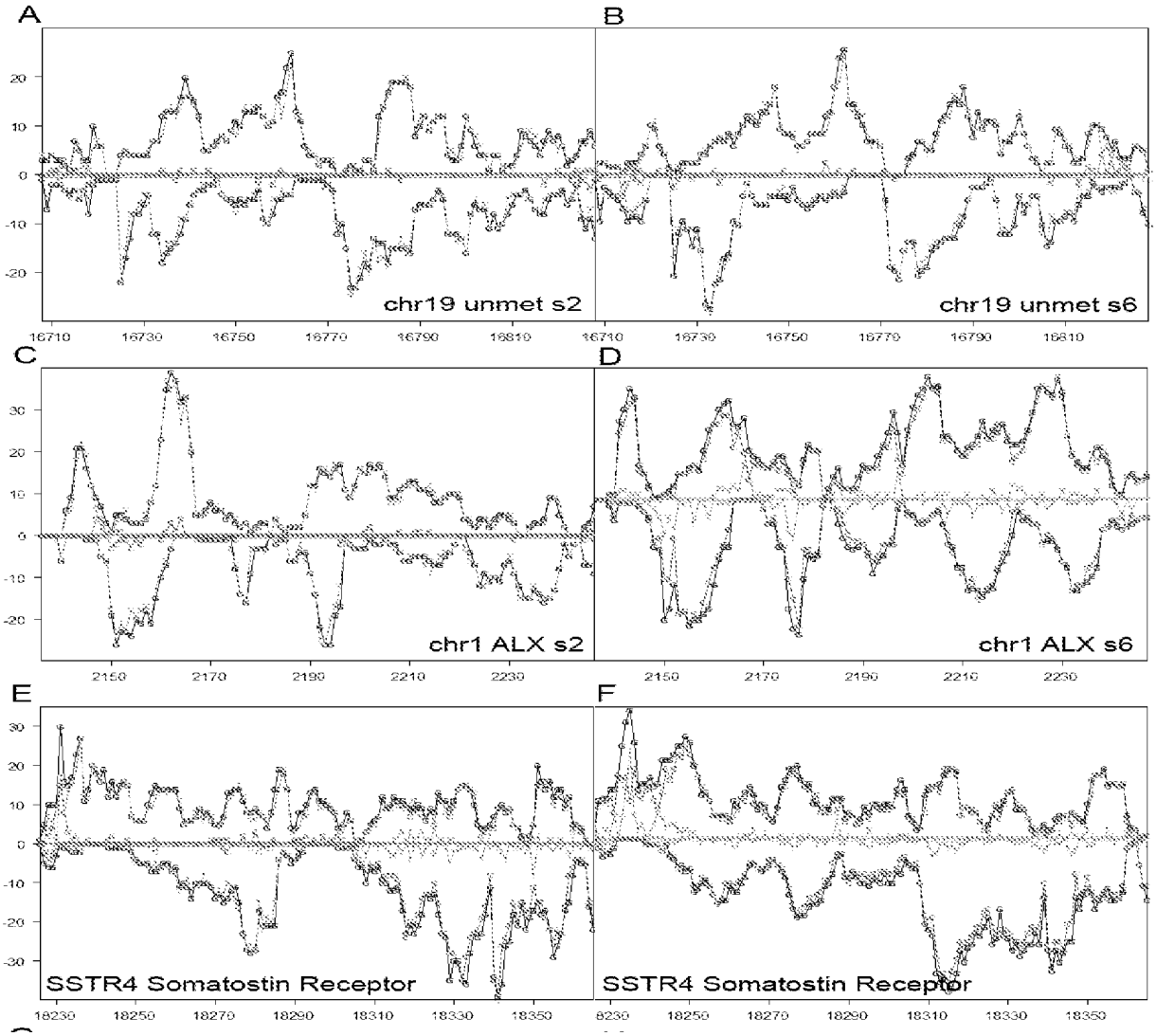

A
B
chr19 unmet s2
chr19 unmet s6
C
D
chr1 ALX s2
chr1 ALX s6
E
F
SSTR4 Somatostin Receptor
SSTR4 Somatostin Receptor

Fig. 3

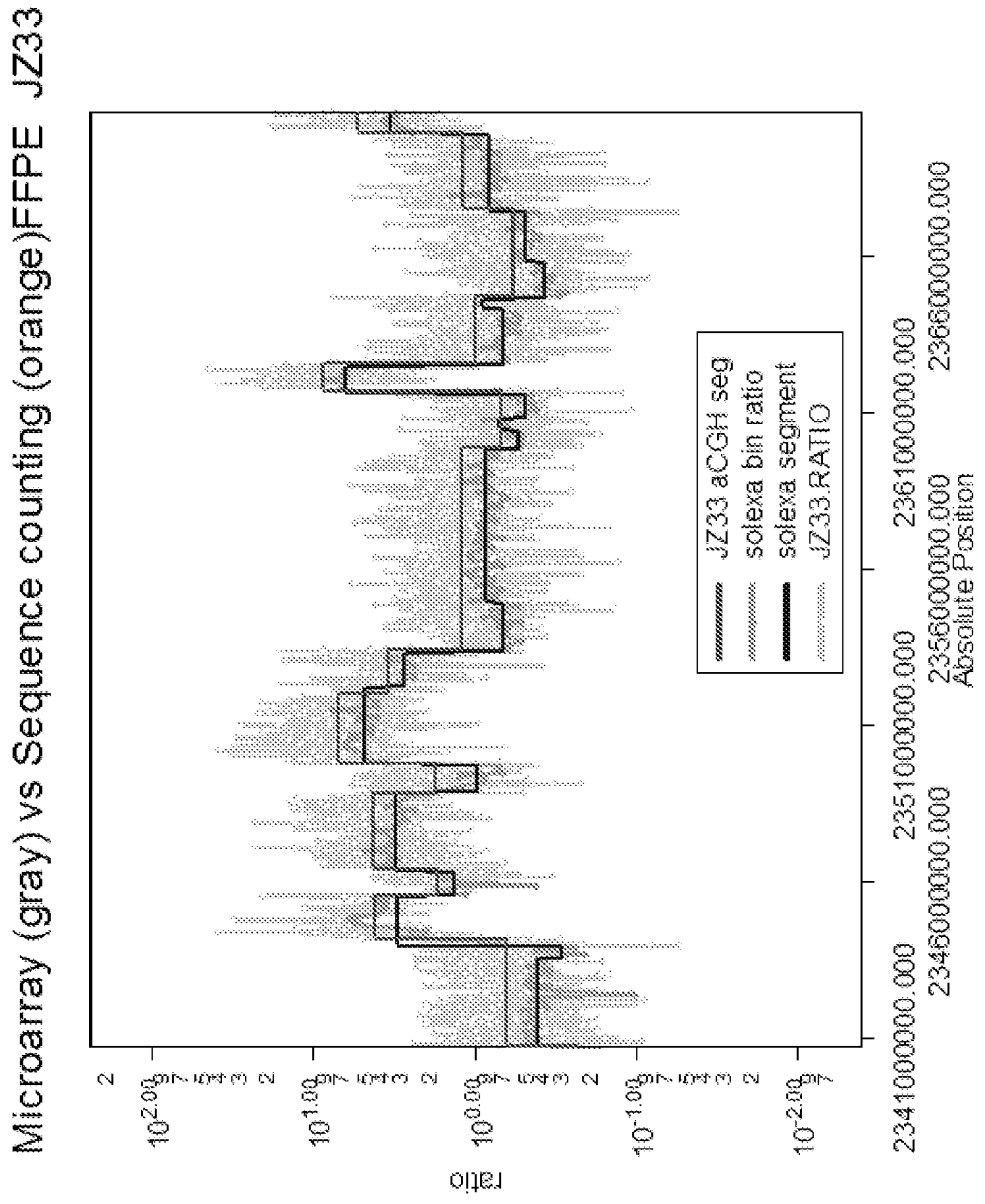



Fig. 4


IPR2025-01353; IPR2025-01355
CSHL EX2002

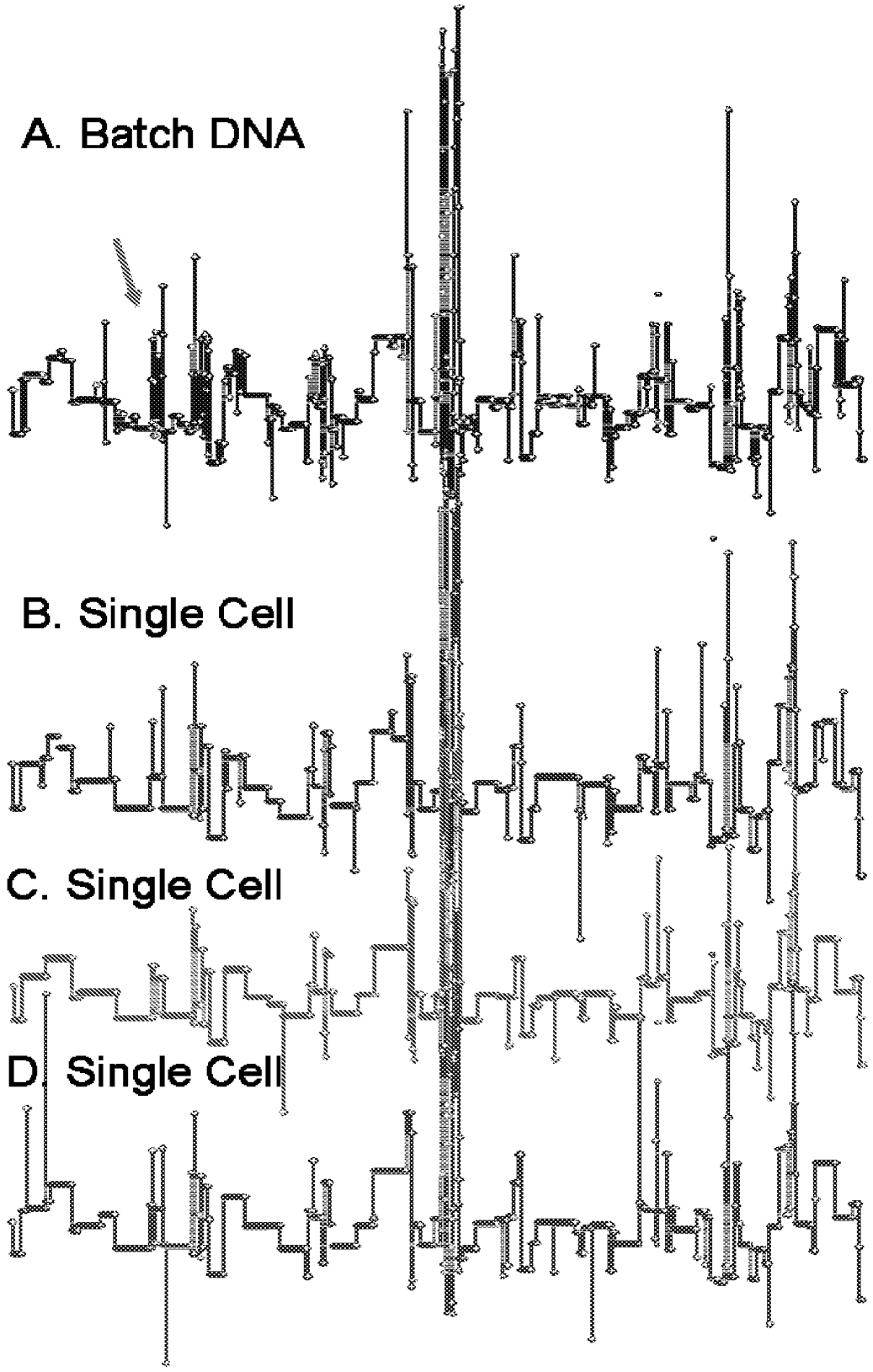



Fig. 5

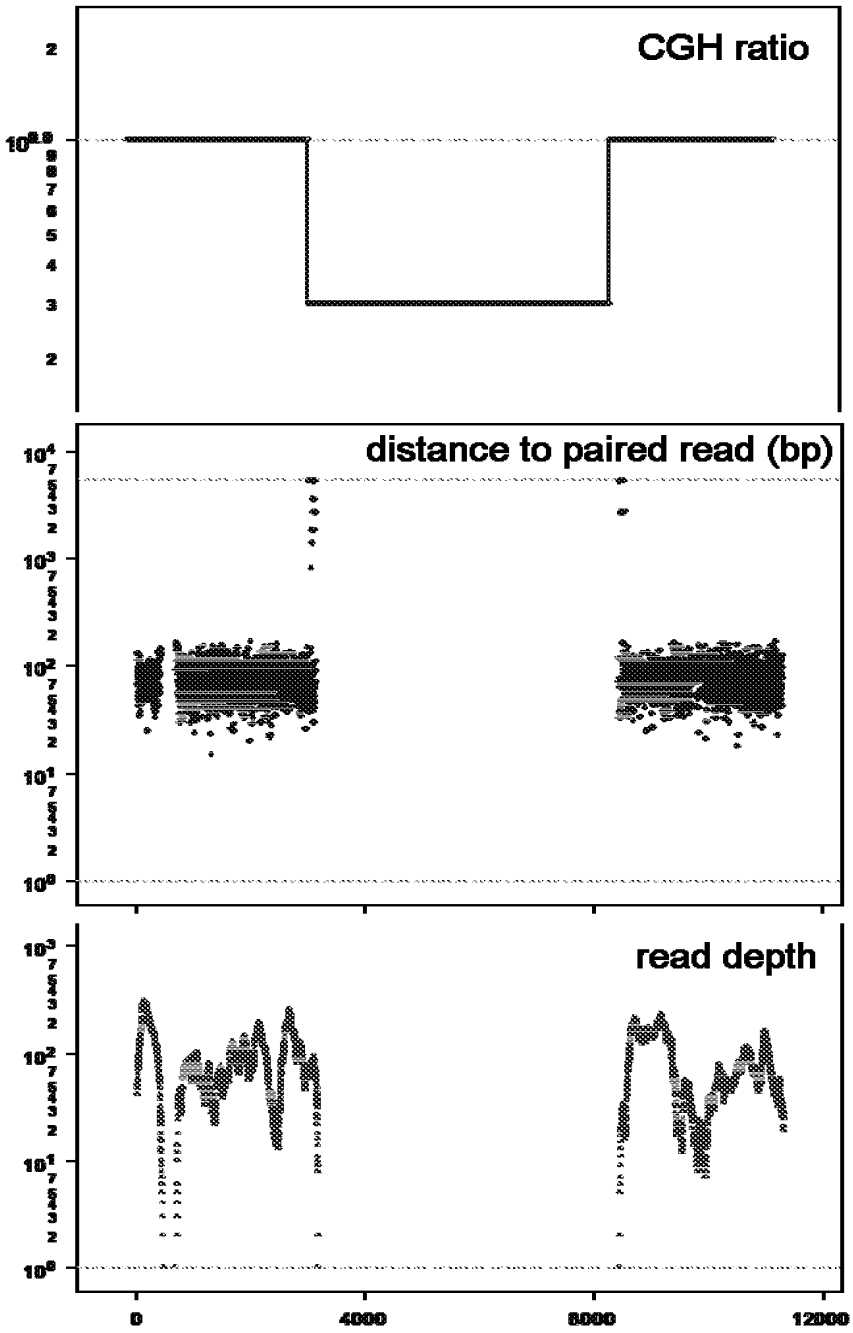



Fig. 6


IPR2025-01353; IPR2025-01355
CSHL EX2002

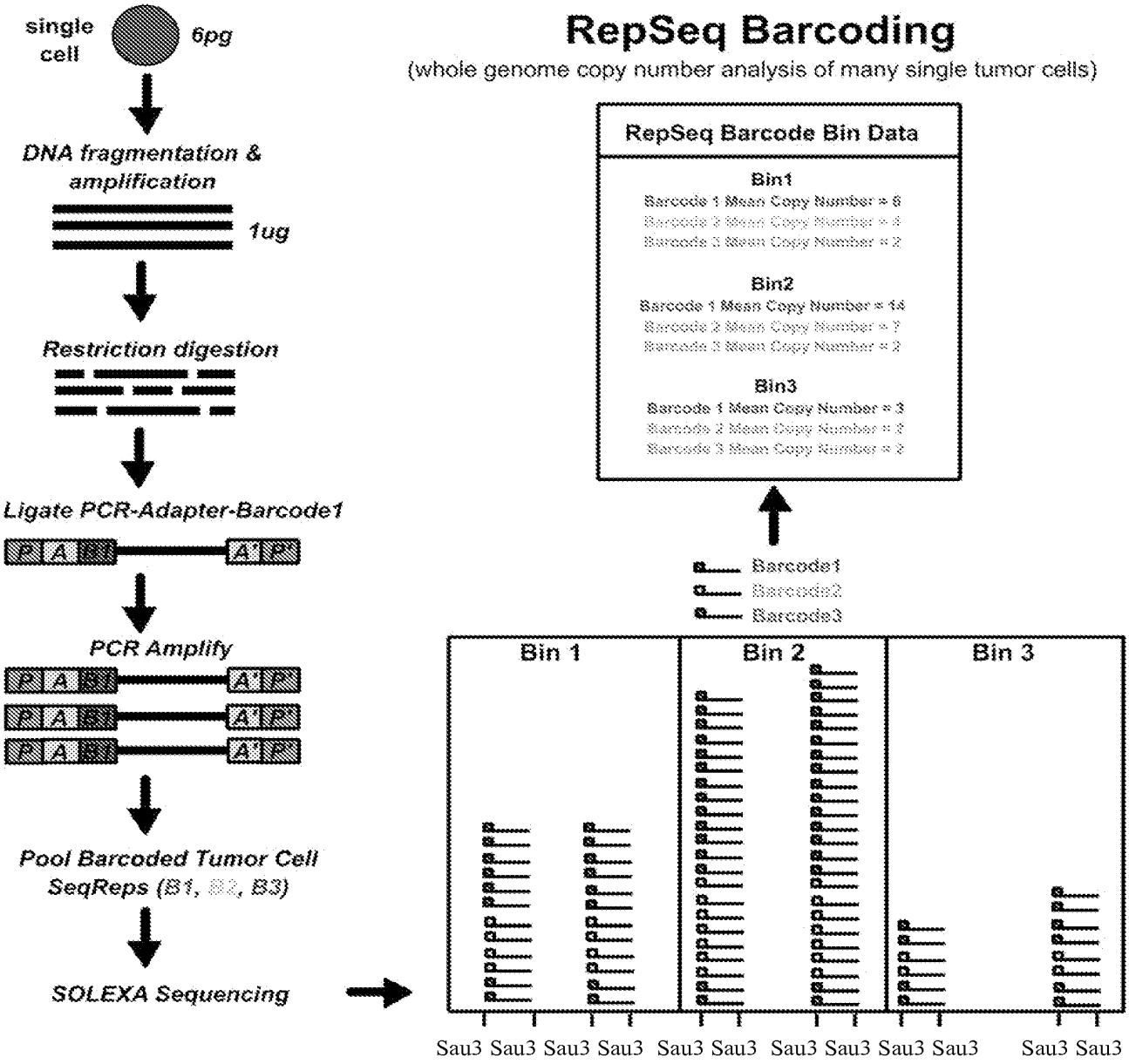



Fig. 7


IPR2025-01353; IPR2025-01355
CSHL EX2002

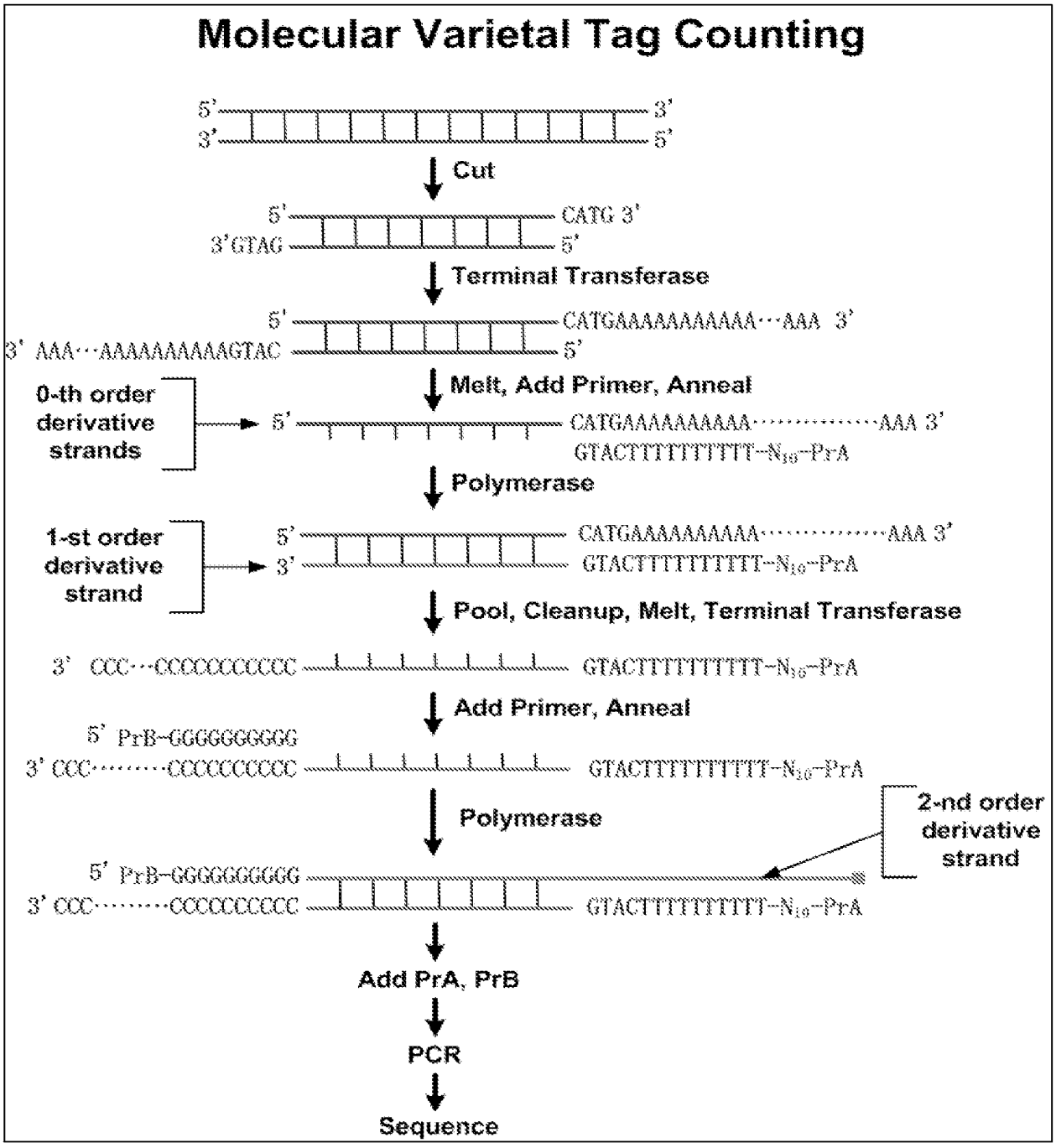



Fig. 8


IPR2025-01353; IPR2025-01355
CSHL EX2002

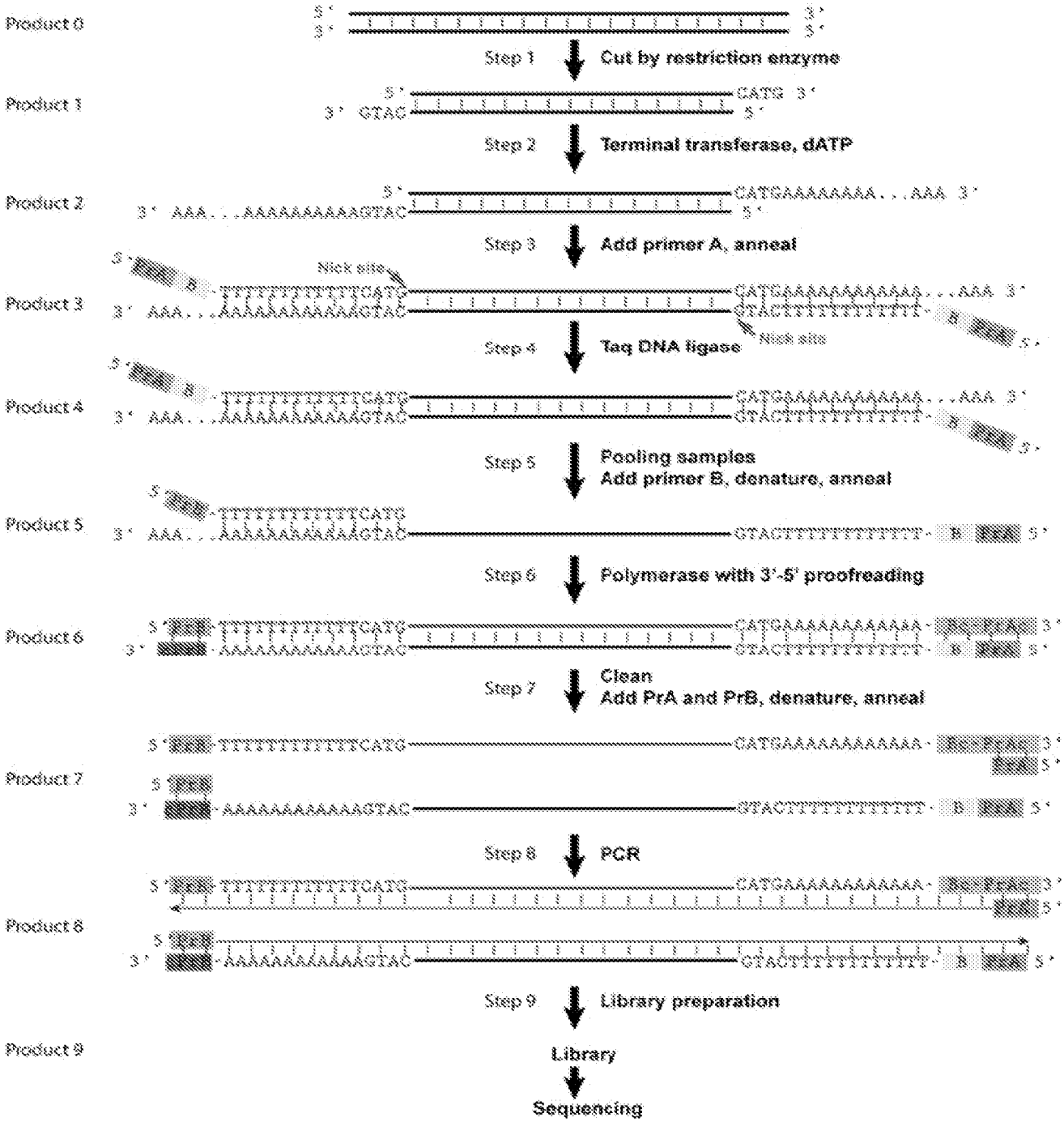

Product 0
Step 1
Cut by restriction enzyme
Product 1
Step 2
Terminal transferase, dATP
Product 2
Step 3
Add primer A, anneal
Nick site
Product 3
Step 4
Taq DNA ligase
Nick site
Product 4
Step 5
Pooling samples
Add primer B, denature, anneal
Product 5
Step 6
Polymerase with 3'-5' proofreading
Product 6
Step 7
Clean
Add PrA and PrB, denature, anneal
Product 7
Step 8
PCR
Product 8
Step 9
Library preparation
Product 9
Library
Sequencing

Fig. 9

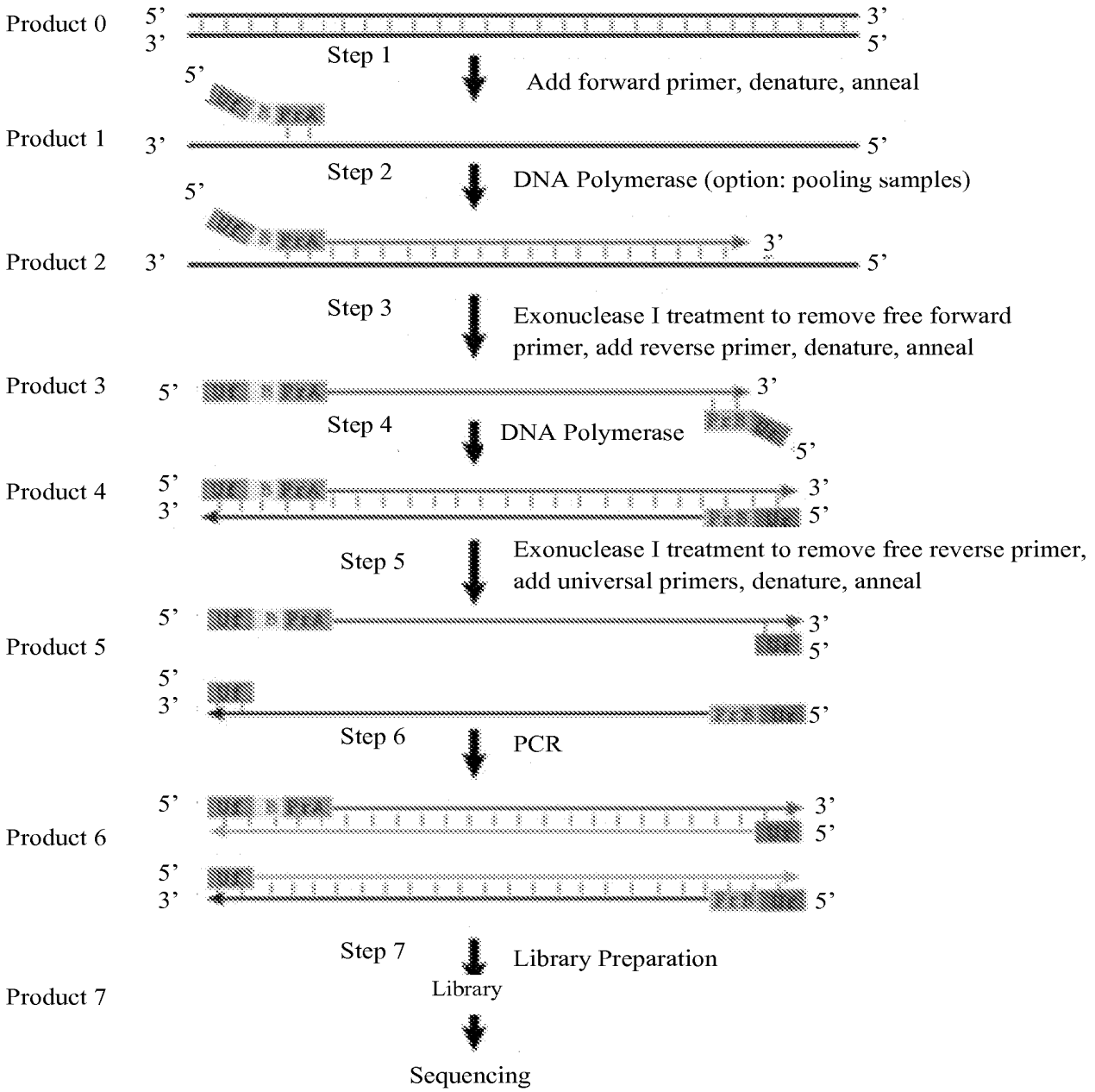



Fig. 10


IPR2025-01353; IPR2025-01355
CSHL EX2002

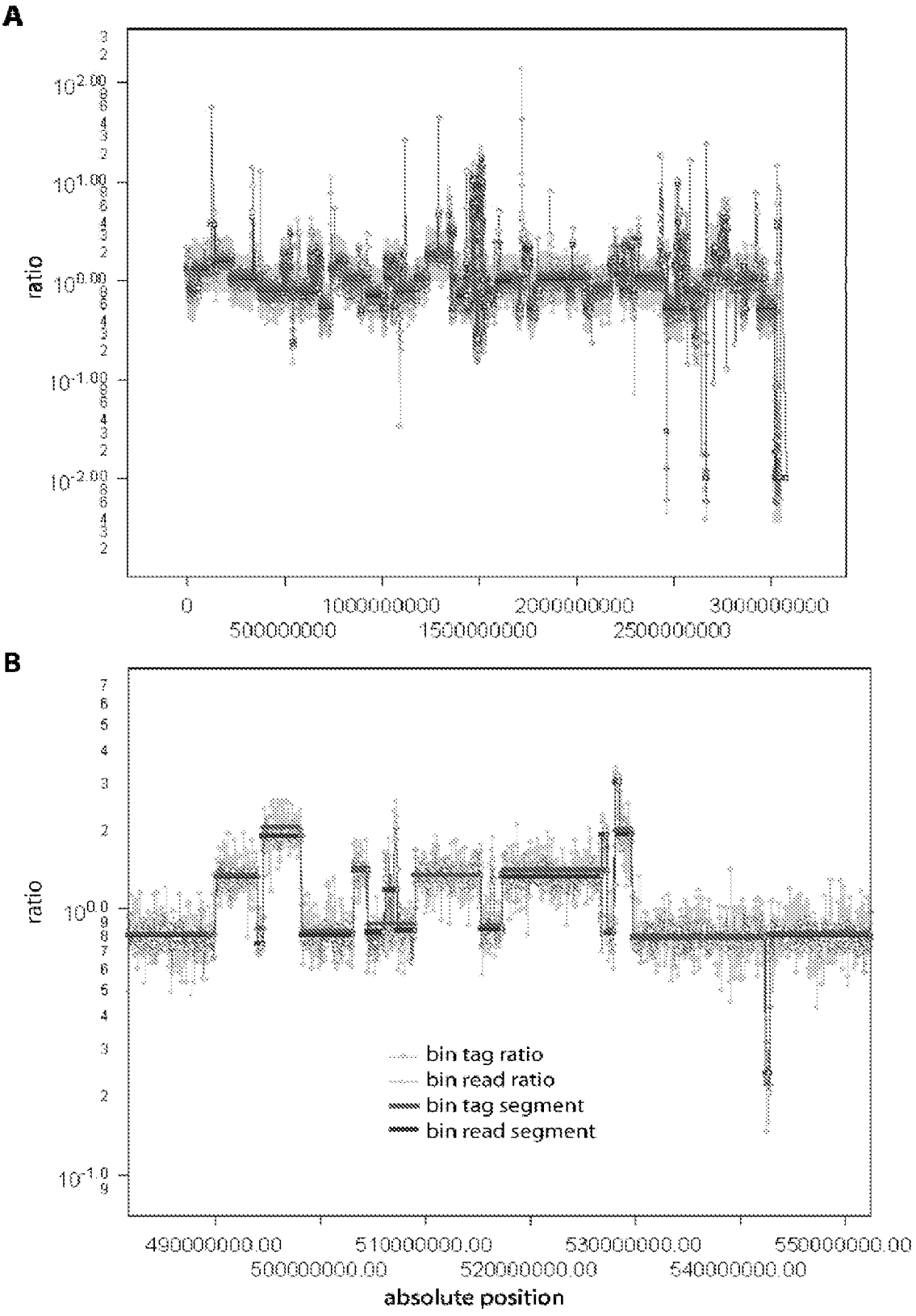

A
ratio
0
500000000
1000000000
1500000000
2000000000
2500000000
3000000000
B
ratio
bin tag ratio
bin read ratio
bin tag segment
bin read segment
490000000.00
500000000.00
510000000.00
520000000.00
530000000.00
540000000.00
550000000.00
absolute position

Fig. 11

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 16405029 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Robert Bolcome |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 24-JUL-2013 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 15:58:45 |
| **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 | | 130724_5981_81503_A_Supplemental_IDS_REB.pdf | 478742<br>3fcd40e7f7266491d026687cc53ec7125e83186e | yes | 4 |

| | Multipart Description/PDF files in .zip description | | |
|---|---|---|---|
| | **Document Description** | **Start** | **End** |
| | Transmittal Letter | 1 | 2 |
| | Information Disclosure Statement (IDS) Form (SB08) | 3 | 4 |

| | | | | | |
|---|---|---|---|---|---|
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Non Patent Literature | 130724_5981_81503_A_Exhibit_1_REB.pdf | 237106<br>1b4a5d9c3ad3bfa390d95513bddf589a8a2d61ad | no | 7 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Foreign Reference | 130724_5981_81503_A_Exhibit_2_REB.pdf | 3572469<br>1169cac7199cca0cb2c6ee8c3307735286ec918e | no | 79 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 4288317 | | |

**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.**

Docket No. 5981/81503-A/GJG/REB

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicant(s) : James Hicks et al.

U.S. Serial No. : 13/278,333

Filing Date : October 21, 2011

Confirmation No. : 4469

For : VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

30 Rockefeller Plaza, 20th Floor
New York, New York 10112
July 24, 2013

**BY EFS**
Commissioner for Patents
Alexandria, VA

**SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT**

In accordance with their duty of disclosure under 37 C.F.R. §1.56, Applicants would like to direct the Examiner's attention to the following documents which are listed on Form PTO-1449 (Substitute), attached hereto as **Exhibit A**, and are also listed below.

This Information Disclosure Statement is being submitted pursuant to 37 C.F.R. §1.97(b)(3) before the mailing of a first Office Action on the merits. Thus, this Supplemental Information Disclosure Statement should be entered and considered.

Reference item **1** is a U.S. Patent. Pursuant to C.F.R. 1.98(a)(2), copies of these reference items are not being submitted.

Copies of reference items **2** and **3** are attached hereto as **Exhibits 1** and **2**, respectively.

1. U.S. Patent No. 7,700,324, issued April 20, 2010 (Issa et al.);


IPR2025-01353; IPR2025-01355
CSHL EX2002

2. International Preliminary Report on Patentability, issued April 23, 2013 in connection with PCT International Application No. PCT/US2011/057350 **(Exhibit 1)**; and

3. PCT International Patent Application Publication No. WO 2012/054873, published April 26, 2012 (Hicks et al.).

The Examiner is respectfully requested to make of record each item listed on the Form PTO-1449 Substitute by initialing and dating the attached Form PTO-1449 Substitute, and returning a copy of the initialed and dated form to applicants' undersigned attorney.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicants' undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this Supplemental Information Disclosure Statement. However, if any additional fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik

Gary J. Gershik
Registration No. 39,992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**
I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

7-24-13

Robert E. Bolcome III Date
Reg. No. 69,032

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 |
|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC |

**CONFIRMATION NO. 4469**

**NEW OR REVISED PPD NOTICE**

23432
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

*OC000000065271431*

# NOTICE OF NEW OR REVISED PROJECTED PUBLICATION DATE

The above-identified application has a new or revised projected publication date. The current projected publication date for this application is 03/06/2014. If this is a new projected publication date (there was no previous projected publication date), the application has been cleared by Licensing & Review or a secrecy order has been rescinded and the application is now in the publication queue.

If this is a revised projected publication date (one that is different from a previously communicated projected publication date), the publication date has been revised due to processing delays in the USPTO or the abandonment and subsequent revival of an application. The application is anticipated to be published on a date that is more than six weeks different from the originally-projected publication date.

More detailed publication information is available through the private side of Patent Application Information Retrieval (PAIR) System. The direct link to access PAIR is currently http://pair.uspto.gov. Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Questions relating to this Notice should be directed to the Office of Data Management, Application Assistance Unit at (571) 272-4000, or (571) 272-4200, or 1-888-786-0101.

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 12/20/2013

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER | |
|---|---|
| LU, FRANK WEI MIN | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 1634 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 12/20/2013 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. 13/278,333 | Applicant(s) HICKS ET AL. | |
|---|---|---|---|
| | Examiner FRANK LU | Art Unit 1634 | AIA (First Inventor to File) Status No |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 2 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *10/21/2011*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims***

5)☒ Claim(s) *1-7,9,11-13,24,26-28,31,41,44,51 and 53* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☐ Claim(s) _____ is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☒ Claim(s) *1-7,9,11-13,24,26-28,31,41,44,51 and 53* are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

**Application Papers**

10)☐ The specification is objected to by the Examiner.

11)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

**Certified copies:**

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date _____.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

## DETAILED ACTION

### *Preliminary Amendment*

1. The preliminary amendment filed on October 21, 2011 has been entered. The claims pending in this application are claims 1-7, 9, 11-13, 24, 26-28, 31, 41, 44, 51, and 53.

### *Election/Restrictions*

2. This application contains claims directed to the following patentably distinct species:

(1) tagging the segments to generate tagged nucleic acid molecules by subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands (claim 11)

(2) tagging the segments to generate tagged nucleic acid molecules by subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands (claims 12 and 13)

The species are independent or distinct because these species are directed to different methods for generating tagged nucleic acid molecules. In addition, these species are not obvious variants of each other based on the current record.

Applicant is required under 35 U.S.C. 121 to elect a single disclosed species, or a single grouping of patentably indistinct species, for prosecution on the merits to which the claims shall

be restricted if no generic claim is finally held to be allowable. Currently, generic claims are 1-7, 9, 24, 26-28, 31, 41, 44, 51, and 53.

There is a search and/or examination burden for the patentably distinct species as set forth above because the search for species (1) is not required for species (2) while the search for species (2) is not required for species (1).

**Applicant is advised that the reply to this requirement to be complete must include (i) an election of a species to be examined** even though the requirement may be traversed (37 CFR 1.143) **and (ii) identification of the claims encompassing the elected species or grouping of patentably indistinct species**, including any claims subsequently added. An argument that a claim is allowable or that all claims are generic is considered nonresponsive unless accompanied by an election.

The election may be made with or without traverse. To preserve a right to petition, the election must be made with traverse. If the reply does not distinctly and specifically point out supposed errors in the election of species requirement, the election shall be treated as an election without traverse. Traversal must be presented at the time of election in order to be considered timely. Failure to timely traverse the requirement will result in the loss of right to petition under 37 CFR 1.144. If claims are added after the election, applicant must indicate which of these claims are readable on the elected species or grouping of patentably indistinct species.

Should applicant traverse on the ground that the species, or groupings of patentably indistinct species from which election is required, are not patentably distinct, applicant should submit evidence or identify such evidence now of record showing them to be obvious variants or clearly admit on the record that this is the case. In either instance, if the examiner finds one of

the species unpatentable over the prior art, the evidence or admission may be used in a rejection under 35 U.S.C. 103 or pre-AIA 35 U.S.C. 103(a) of the other species.

Upon the allowance of a generic claim, applicant will be entitled to consideration of claims to additional species which depend from or otherwise require all the limitations of an allowable generic claim as provided by 37 CFR 1.141.

Applicant is reminded that upon the cancellation of claims to a non-elected invention, the inventorship must be corrected in compliance with 37 CFR 1.48(a) if one or more of the currently named inventors is no longer an inventor of at least one claim remaining in the application. A request to correct inventorship under 37 CFR 1.48(a) must be accompanied by an application data sheet in accordance with 37 CFR 1.76 that identifies each inventor by his or her legal name and by the processing fee required under 37 CFR 1.17(i).

3. Papers related to this application may be submitted to Group 1600 by facsimile transmission. Papers should be faxed to Group 1600 via the PTO Fax Center. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993)(See 37 CAR § 1.6(d)). The CM Fax Center number is (571)273-8300.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Frank Lu, Ph.D., whose telephone number is (571)272-0746. The examiner can normally be reached on Monday-Friday from 9 A.M. to 5 P.M.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dave Nguyen, can be reached on (571)272-0731.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Frank W Lu /
Primary Examiner, Art Unit 1634
December 18, 2013


IPR2025-01353; IPR2025-01355
CSHL EX2002

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

# BIB DATA SHEET

**CONFIRMATION NO. 4469**

| SERIAL NUMBER | FILING or 371(c) DATE | CLASS | GROUP ART UNIT | ATTORNEY DOCKET NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011<br>RULE | 435 | 1634 | 5981/81503-A/GJG/GC |

**APPLICANTS**

**INVENTORS**

James Hicks, Locust Valley, NY;
Nicholas Navin, Houston, TX;
Jennifer Troge, Oyster Bay, NY;
Zihua Wang, Huntington, NY;
Michael Wigler, Cold Spring Harbor, NY;

**** CONTINUING DATA ***************************

This appln claims benefit of 61/510,579 07/22/2011
and claims benefit of 61/406,067 10/22/2010

**** FOREIGN APPLICATIONS ***************************

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ** ** SMALL ENTITY ****

| Foreign Priority claimed ☐ Yes ☑ No<br>35 USC 119(a-d) conditions met ☐ Yes ☐ No<br>Verified and Acknowledged /FRANK WEI MIN LU/ Examiner's Signature | ☐ Met after Allowance<br>Initials | STATE OR COUNTRY | SHEETS DRAWINGS | TOTAL CLAIMS | INDEPENDENT CLAIMS |
|---|---|---|---|---|---|
| | | NY | 11 | 20 | 3 |

**ADDRESS**

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112
UNITED STATES

**TITLE**

VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

| FILING FEE RECEIVED | FEES: Authority has been given in Paper | ☐ All Fees |
|---|---|---|
| 595 | No._____ to charge/credit DEPOSIT ACCOUNT<br>No._____ for following: | ☐ 1.16 Fees (Filing)<br>☐ 1.17 Fees (Processing Ext. of time)<br>☐ 1.18 Fees (Issue)<br>☐ Other _____<br>☐ Credit |

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
February 20, 2014

**BY EFS**
Commissioner for Patents
Alexandria, VA

**COMMUNICATION IN RESPONSE TO DECEMBER 20, 2013 OFFICE ACTION AND SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT**

This Communication and Supplemental Information Disclosure Statement is submitted in reply to a December 20, 2013 Office Action issued by the United States Patent and Trademark Office in connection with the above-identified application. A response to the December 20, 2013 Office Action is due February 20, 2014. Accordingly, this response is being timely filed.

**Remarks** begin on page **2** of this paper.

A **Supplemental Information Disclosure Statement** begins on page **4** of this paper.

**REMARKS**

Claims 1-7, 9, 11-13, 24, 26-28, 31, 41, 44, 51 and 53 are pending in the subject application. Applicants have not herein amended, added, or cancelled any claim. Accordingly, claims 1-7, 9, 11-13, 24, 26-28, 31, 41, 44, 51 and 53 are pending in the subject application.

**Response to December 20, 2013 Restriction Requirement**

In the December 20, 2013 Office Action, the Examiner required applicants to elect a single species from among the species listed below for initiating prosecution on the merits and to which the claims shall be restricted if no generic claim is finally held to be allowable:

(1) Tagging the segments to generate tagged nucleic acid molecules by subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands (claim 11); and

(2) Tagging the segments to generate tagged nucleic acid molecules by subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands (claims 12 and 13).

The Examiner acknowledged that claims 1-7, 9, 24, 26-28, 31, 41, 44, 51, and 53 were generic.

*Applicants' Reply*

In response, applicants elect for initiating examination, Species (2), *i.e.* tagging the segments to generate tagged nucleic acid molecules by subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands

primer to the 5' ends of the zero-th order derivative strands (claims 12 and 13).

Claims 1-7, 9, 12, 13, 24, 26-28, 31, 41, 44, 51 and 53 encompass the elected species.

As acknowledged by the Examiner, upon the allowance of a generic claim, applicants will be entitled to consideration of claims to additional species which depend from or otherwise require all the limitations of an allowable generic claim.

**Supplemental Information Disclosure Statement**

In accordance with their duty of disclosure under 37 C.F.R. § 1.56, applicants would like to direct the Examiner's attention to the following references, which are listed on Form 1449/PTO (Substitute) (**Exhibit A**).

This Information Disclosure Statement is being submitted pursuant to 37 C.F.R. § 1.97(b)(3) before the mailing of a first Office Action on the merits. Thus, no fee is required and this Supplemental Information Disclosure Statement should be entered and considered.

Reference items **1-12** are U.S. Patents and U.S. Patent Application Publications. Pursuant to 37 C.F.R. § 1.98(a)(2), copies of these reference items are not being submitted.

1. U.S. Patent No. 6,569,617, issued May 27, 2003 (Wigler et al.);
2. U.S. Patent No. 7,531,307, issued May 12, 2009 (Wigler et al.);
3. U.S. Patent No. 8,273,871, issued September 25, 2012 (Hannon et al.);
4. U.S. Patent No. 8,554,488, issued October 8, 2013 (Wigler et al.);
5. U.S. Patent Application Publication No. 2005-0032095, published February 10, 2005 (Wigler et al.);
6. U.S. Patent Application Publication No. 2005-0266444, published December 1, 2005 (Wigler et al.);
7. U.S. Patent Application Publication No. 2007-0207481, published September 6, 2007 (Wigler et al.);
8. U.S. Patent Application Publication No. 2010-0227768, published September 10, 2010 (Wigler et al.);
9. U.S. Patent Application Publication No. 2012-0149593, published June 14, 2012 (Hicks et al.);
10. U.S. Patent Application Publication No. 2012-0328607, published December 27, 2012 (Hicks et al.);
11. U.S. Patent Application Publication No. 2013-0179999,

published July 11, 2013 (Hannon et al.); and

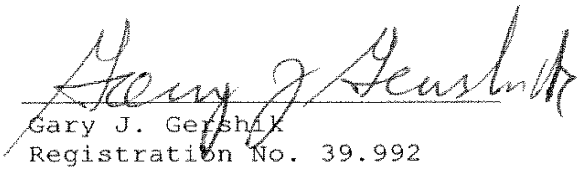
12. U.S. Patent Application Publication No. 2014-0025307, published January 23, 2014 (Wigler et al.).

Applicants request that the Examiner review these references and make them of record in the subject application.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicant's undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this response. However, if any fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

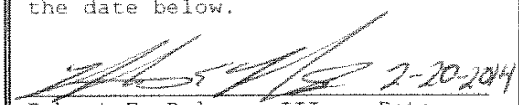
Respectfully submitted,



Gary J. Gershik
Registration No. 39.992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.



2-20-2014

Robert E. Bolcome III Date
Reg. No. 69,032

**Form PTO-1449 (Substitute)** **U.S. Department of Commerce**
**Patent and Trademark Office**

**INFORMATION DISCLOSURE STATEMENT**
**(Use several sheets if necessary)**

| | |
|---|---|
| **Application Number** | 13/278,333 |
| **Filing Date** | October 21, 2011 |
| **First Named Inventor** | James Hicks |
| **Art Unit** | 1634 |
| **Examiner Name** | Frank W. M. Lu |
| **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| | 1 | 6,569,617 | 05-27-2013 | Wigler et al. |
| | 2 | 7,531,307 | 05-12-2009 | Wigler et al. |
| | 3 | 8,273,871 | 09-25-2012 | Hannon et al. |
| | 4 | 8,554,488 | 10-08-2013 | Wigler et al. |
| | 5 | 2005-0032095 | 02-10-2005 | Wigler et al. |
| | 6 | 2005-0266444 | 12-01-2005 | Wigler et al. |
| | 7 | 2007-0207481 | 09-06-2007 | Wigler et al. |
| | 8 | 2010-0227768 | 09-10-2010 | Wigler et al. |
| | 9 | 2012-0149593 | 06-14-2012 | Hicks et al. |
| | 10 | 2012-0328607 | 12-27-2012 | Hicks et al. |
| | 11 | 2013-0179999 | 07-11-2013 | Hannon et al. |
| | 12 | 2014-0025307 | 01-23-2014 | Wigler et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |

**EXAMINER SIGNATURE** **DATE CONSIDERED**

*EXAMINER: Initial if citation 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 of 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.

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 18254628 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Robert Bolcome |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 20-FEB-2014 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 15:19:11 |
| **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 | | 140220_5981_81503_Communication_REB.pdf | 874636<br>915e730938987ac5199dc3843a64b7d154f1034a | yes | 6 |



IPR2025-01353; IPR2025-01355
CSHL EX2002

| | Multipart Description/PDF files in .zip description | | |
|---|---|---|---|
| | **Document Description** | **Start** | **End** |
| | Response to Election / Restriction Filed | 1 | 3 |
| | Transmittal Letter | 4 | 4 |
| | Response to Election / Restriction Filed | 5 | 5 |
| | Information Disclosure Statement (IDS) Form (SB08) | 6 | 6 |
| **Warnings:** | | | |
| **Information:** | | | |
| | **Total Files Size (in bytes):** | 874636 | |

**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.**

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 |
|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC |

**CONFIRMATION NO. 4469**

23432
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

**PUBLICATION NOTICE**




**Title:**VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

**Publication No.**US-2014-0065609-A1
**Publication Date:**03/06/2014

## NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Office of Public Records. The Office of Public Records can be reached by telephone at (703) 308-9726 or (800) 972-6382, by facsimile at (703) 305-8759, by mail addressed to the United States Patent and Trademark Office, Office of Public Records, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently http://pair.uspto.gov/. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

---

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

**Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office**

**INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary)**

| | |
|---|---|
| Application Number | 13/278,333 |
| Filing Date | October 21, 2011 |
| First Named Inventor | James Hicks |
| Art Unit | 1634 |
| Examiner Name | Frank W. M. Lu |
| Attorney Docket No. | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| | 1 | 2004-110153 | 06-10-2004 | Dong et al. |
| | 2 | 2007-172873 | 07-26-2007 | Brenner et al. |
| | 3 | 2010-178655 | 07-15-2010 | Hamilton et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | 4 | WO 2007/062445 | 06-07-2007 | Alexander Seitz | |
| | | | | | |
| | | | | | |

EXAMINER SIGNATURE

DATE CONSIDERED

*EXAMINER: Initial if citation 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 of 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | **Application Number** | 13/278,333 |
| | **Filing Date** | October 21, 2011 |
| | **First Named Inventor** | James Hicks |
| | **Art Unit** | 1634 |
| | **Examiner Name** | Frank W. M. Lu |
| | **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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] |
|---|---|---|---|
| | 5 | February 13, 2014 Extended European Search Report, issued in connection with European Patent Application No. 11835243.4 | |
| | 6 | Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 June 2007 (07.06.2007)

PCT

(10) International Publication Number
WO 2007/062445 A1

(51) International Patent Classification:
*C12Q 1/68* (2006.01)

(21) International Application Number:
PCT/AT2006/000494

(22) International Filing Date:
29 November 2006 (29.11.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
A 1923/2005 29 November 2005 (29.11.2005) AT

(71) Applicant and
(72) Inventor: **SEITZ, Alexander** [AT/AT]; St.-Veit-Gasse 18/2, A-1130 Vienna (AT).

(74) Agent: **SONN & PARTNER PATENTANWÄLTE**; Riemergasse 14, A-1010 Vienna (AT).

(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, 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).

**Declaration under Rule 4.17:**
— *of inventorship (Rule 4.17(iv))*

**Published:**
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 2007/062445 A1

(54) Title: POLYNUCLEOTIDE AMPLIFICATION

(57) Abstract: The present invention provides a method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of a) obtaining a sample or RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining a sample of full length cDNA, b) tailing the 3' end of the transcribed cDNA with a polynucleotide tail after the 3' end, c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (51) primer is specific for the a upstream portion of the polynucleotide tail and the next 1 to 10 nucleotides upstream of the 3' polynucleotide tail of the cDNA.

# Polynucleotide amplification

The present invention relates to the field of specific polynucleotide amplification.

To understand the causalities of the different states life presents itself, one needs to find the factors and processes that lead to it. One central element in this quest is the genome where information is stored in such a fashion that it can be propagated through the generations in a highly conserved manner. In executing the information contained in the genome, DNA is transcribed into RNA and further translated into protein. Studying the genome (DNA), transcriptome (RNA) and proteome (protein) has in recent years greatly contributed to our understanding of the molecular basis of life.

An array of possibilities exists to compare RNA samples in vitro. An early approach is the method of differential hybridization or differential screening (Rebagliati et al., 1985). It involves picking at random clones (or sequences) from cDNA libraries and checking for their presence in different concentration in samples of RNA. An enormous effort has to be put into identifying a single differentially expressed transcript as their population is thought to be below 1%. Subtractive hybridization has been proposed to overcome this problem by enriching for differentially expressed transcripts prior to screening. Another way of addressing the problem of abundance was made by introducing high-density micro arrays. In using micro array technology 100s to 1000s of cDNAs, corresponding to RNAs of genes of interest, are spotted and bound onto a surface. They are normally derived from sequenced EST libraries and - if non redundant - can cover as many transcripts as spots fit onto the membrane. In short, RNA extracted from cells or tissues of interest are labeled, e.g. during reverse transcription, and hybridized to these microarrays. The signal intensity of corresponding spots between two or more RNA preparations can then be compared. Even though this technology has produced massive amounts of data, there are two fundamental drawbacks of all these hybridization techniques. One is cross hybridization. This means that two DNA molecules do not need to be complementary over their entire sequence to hybridize. Therefore this introduces ambiguity towards the identity of the molecule that hybridizes (binds) to



IPR2025-01353; IPR2025-01355
CSHL EX2002

the spot. Secondly one can only compare molecules that have been spotted, limiting the search to sequences that are present in the library used to generate the array.

An alternative approach was introduced in 1992 (Liang et al., 1992; Welsh et al., 1992) that today is generally known as "differential display" (DD). In DD the total mRNAs population of a sample is first divided into defined pools of cDNAs during reverse transcription (RT). This is accomplished by priming the RT with one, two or x base anchored oligo dT primers, subdividing the samples into 3, 12 or $3x4^{x-1}$ fractions. These anchor bases hybridize to the bases immediately upstream of the poly A tail. Therefore selecting for mRNA molecules to be reverse transcribed whose bases immediately upstream of the poly A tail are complementary to the anchor bases of the primer. In a subsequent Polymerase chain reaction (PCR) arbitrary primers and anchored primers are used to amplify defined pools of cDNA. Corresponding pools are run side by side on a gel that separates them by size. Differences of expression levels are reflected in differences in band intensity. Bands of interest are cut out of the gel, reamplified and sequenced. One drawback of the original DD is that it uses arbitrary primers for amplification. Such a shotgun approach can reach full transcriptome coverage only statistically, by covering the transcriptome several times. Therefore some RNA molecules will necessarily be presented several times in the display and others might not be present at all.

Recently methods have been introduced that amplify pools of RNA in a more systematic fashion (Prashar et al., 1996; Prashar et al., 1999). In principle, three additional steps are involved. After first strand cDNA synthesis second strand synthesis is carried out to get double stranded DNA that can be cut with a restriction enzyme. In a third step an adaptor is ligated to the 5′ ends of the fragments. These fragments are then PCR amplified with one primer being complementary to the 5′ adaptor sequence and an anchor primer that was used for RT. It was estimated by the authors that each 6-base cutting restriction enzyme cuts 8% of the cDNAs at a position between 50 and 400 bases from the 3' poly A tail of the mRNA. Therefore more than 12 restriction enzymes will be necessary to approach a complete coverage of the transcriptome. A somewhat more elegant variety of this method was described by Matz et al. (1997)with the intro-

duction of "ordered differential display" that takes advantage of the PCR suppression effect.

However, besides covering the transcriptome only statistically at one level or another, all afore mentioned DD methods suffer from another drawback. They will give the investigator only part of the mRNA sequence that is differentially expressed. Downstream of these DD technologies one will still have to either screen cDNA libraries or where available online sequence libraries for full length clones or employ methodologies that can discover 5′ and 3′ sequences from a partially known mRNA sequence such as RACE (Frohman et al., 1988). Even identifying such full length transcripts will still not yield certainty at least in eukaryotes that the mRNA found is the mRNA that generated the signal as in eukaryotes a major contributor to genome complexity is splicing. Therefore the sequence found on the display can be common to several splice variants. Furthermore different splice variants of a gene may have different functions, making it highly desirable to know the correct splice variant that generated the signal.

A goal of the present invention is to provide a method that can amplify (enrich for) defined subpopulations of transcripts in their entire length with full coverage of all mRNA molecules present in the sample.

The present invention now provides a method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of:

a) obtaining (or providing) a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining (or providing) a sample of full length cDNA,

b) tailing the 3′ end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 10 nucleotides (i.e. the next 1 or 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides), preferably 1 to 5 nucleotides, even more preferred 3 to 5 nucleotides, upstream of the 3' polynucleotide tail of the cDNA. Of course the second primer

can also comprise more specific nucleotides further upstream of the 1 to 10 nucleotide region upstream of the tail specific region (complement).

A full length cDNA can be created by a reverse transcription process described below, or by methods known in the state of the art. Methods to generate full length cDNA are abundantly known to the skilled man in the art, e.g. disclosed in patents US 5,962,271 and US 5,962,272. Such generated cDNA can also be used as starting material to pool full-length cDNA into different fractions. These methods use the template switching ability of reverse transcriptase to add a defined sequence to the 5′ end of the cDNA in a cap dependent manner. The primer that is needed in the subsequent amplification will therefore have a portion that will be complementary to the so generated 3′ end of the cDNA and will define what fraction of cDNA will be amplified through several, e.g. 1 to 10, additional bases defined by the 5′ end of the mRNA molecules amplified.

Therefore it is preferred that a tail is added to the cDNA by the template switching ability of the reverse transcriptase, preferably in a cap dependent manner and the second amplification primer (5′ primer) comprises complementary nucleotides to the so added tail and oligo-dCs and 1 to 10 nucleotides upstream of the oligo-dC sequence.

As used herein, the term "reverse transcriptase" relates to any polymerase that has reverse transcriptase activity and can be used to synthesize cDNA from RNA.

As used herein, the term "full length cDNA" is defined as DNA that includes a sequence complementary to the RNA sequence from the first base to the last base of the RNA. In case of RNA molecules that have a cap and/or a polyA tail as is the case for e.g. most eukaryotic mRNA, "full length cDNA" is defined as DNA that includes a sequence complementary to the RNA sequence from the first base after the cap, e.g. the RNA 7-methylguanosine cap or RNA m7G cap, to the last base before the poly A tail of the RNA used as a template.

As used herein, the term "full length amplification product" is defined as DNA that includes a sequence complementary to the RNA sequence from the first base to the last base of the RNA. In case of RNA molecules that have a cap and/or a polyA tail as is the case for e.g. most eukaryotic mRNA, "full length amplifica-

tion product" is defined as amplification product that includes a sequence complementary to the RNA sequence from the first base after the cap to the last base before the poly (A) tail of the RNA used as a template.

A central element of this invention is the fact that full length RNA molecules can be represented in subpopulations that are defined on their 5′ and/or 3′ end. In principle that means that after the generation of full length cDNA by any means known in the art, the 5' and/or 3' sequence can be used to formulate a set of primer combinations (primer matrix) that has the following quality.

1) Each cDNA will be amplified by only one primer pair in the absence of miss priming. Therefore full transcriptome coverage can be reached without redundancy.
2) Each cDNA sequence within such a defined pool will encode the full amino acid sequence of the protein translated from the corresponding RNA and also
3) Define its transcription start site on its gene (there could be different RNA molecules transcribed from the same gene that can have different transcription start sites).

In comparing quality (1) to original DD one can state that therein the arbitrary primers have the property of either not being complementary to a specific cDNA, being complementary once or being complementary more than once. Therefore each cDNA can be represented either not at all, or once or more than once. When using alternative methodologies that include a restriction enzyme, each restriction enzyme will cut a specific cDNA either not at all or once or more than once. Therefore truly complete transcriptome coverage can never be reached. And increasing the probability for mRNAs to be represented at least one time in the display will mean to increase redundancy.

Qualities (2) and (3) can also not be reached by differential display techniques used to date without employing downstream technologies such as RACE.

The present invention provides methods for amplifying and analyzing the transcriptome through qualitative and quantitative detection of differentially expressed RNA molecules. The technique according to the invention is an organized full length expression display.

In a preferred embodiment the method according to the inven-

tion comprises a primer, which is specific for the (preferably 3' poly A) tail of the RNA or mRNA and the next 1 to 10 nucleotides, preferably 1 to 5 nucleotides, even more preferred 3 to 5 nucleotides, upstream of the (poly A) tail, that is used for the reverse transcription and/or one amplification primer is specific for a corresponding (5' poly-T) stretch of the cDNA, which is complementary to the (3' poly A) tail of the RNA or mRNA, and the next 1 to 10 nucleotides, preferably 1 to 5 nucleotides, even more preferred 3 to 5 nucleotides, downstream of the corresponding (5' poly-T) stretch of the cDNA. Accordingly in addition to the selective 3' primer described above a selective 5' primer can be used to increase the selective power of the primer pair and thus increase the number of cDNA pools defined by the specific nucleotides next to the end sequence anchors.

Preferably the primer for reverse transcription consists of deoxynucleotides.

Generally, the primers throughout this specification can be of any nucleotides, such as deoxyribonucleotides or ribonucleotides. Although deoxynucleotides are preferred in most cases, ribonucleotide primers are also feasible. Therefore a deoxyribonucleotide in a given primer, such as given in examples below, may be substituted by one of its ribonucleotide analogue (e.g. dG by G, dA by A, dC by C, dT by T or U) or other nucleotide analogues.

Preferably the tailing of the 3' end is performed using terminal transferase. - Although other tailing methods are also disclosed, like ligation of a tail sequence, which can be e.g. a defined arbitrary sequence. The terminal transferase can add a certain number of nucleotides preferably uniformly selected from one nucleotide type. Any other means for tailing, adding a tail sequence can also be used, e.g. employing the template switching ability of the reverse transcriptase or by ligating the tail sequence which can be uniformly of one type of nucleotides or of varying nucleotides. Such a tail is preferably a sequence in the range between 5 and 500 nucleotides, more preferred less than 400, less than 300, less than 200, less than 100, less than 50 or less than 30 nucleotides.

In a preferred embodiment the amplification of the cDNA sequences using a selective primer pair is performed by PCR. PCR or polymerase chain reaction is a well established technique

known in the state of the art to amplify polynucleotides using a polymerase, nucleotidetriphosphates, primers, buffer substances and cofactors such as Mg ions and a temperature protocol for hybridization, polymerisation and melting or strand separation.

Preferably the first primer and/or second primer for the amplification consists of deoxynucleotides.

In a preferred embodiment the method according to the invention comprises an additional step of separation of the amplification products according to their length. The separation of the created polynucleotides renders a characterization, isolation and sequence determination possible for each pool after the selective polymerisation which allows a significant advantage for this preferred embodiment in practice.

Preferably the separation is performed by gel electrophoresis, preferably agarose gel electrophoresis, polyacrylamide gel electrophoresis, or capillary electrophoresis.

Even more preferred the method according to the invention comprises an additional step of determining the identity or sequence of the amplification products. Sequence determination can be done for example by gel sequencing, e.g. PCR with dideoxynucleotides, or in an automated sequencer.

Preferably the step of determining the identity or sequence of the amplification products is performed by an automated process on a chip.

The sample used for reverse transcription preferably contains total RNA or mRNA, preferably purified RNA or mRNA, from a specimen. Total RNA includes, but is not limited to, protein coding RNA also called coding RNA such as messenger RNA (mRNA) and non protein coding RNA (non coding RNA or ncRNA), such as ribosomal RNA (rRNA), transfer RNA (tRNA), micro RNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA). Every one of these classes, alone or together with one or more classes, can be analysed according to methods that are within the scope of the invention.

These RNA molecules are preferably purified by affinity chromatography using e.g. an oligo-dT-cellulose column. Thus mRNA with a poly A tail can be preselected.

It is preferred in the inventive method that reverse transcription is carried out at high stringency conditions to in-

crease selectivity of the anchored primers, which can be used as a pool of anchored primers in one polymerisation reaction, as well as separate primers for different polymerisation reactions. In the present invention the phrases "stringent hybridization conditions" or "stringent conditions" refer to conditions under which a primer of the invention will hybridize to its target sequence, but to a minimal number of other sequences. High stringency can be achieved for example by using a high annealing temperature alone or in combination with a chemical additive such as dimethylsulfoxide (DMSO), formamide, tetramethylammonium chloride (TMAC) and Tetramethylammonium oxalate (TMAO) (Kovarova et al. 2000). Examples for high stringency hybridization conditions are given in Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, N.Y. 1989) and Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley and Sons, N.Y. 1994. Under standard PCR conditions often an annealing temperature is chosen that is about 5 degree Celsius below the melting temperature of the primer. Keeping reaction conditions constant the melting temperature in turn is mainly dependent on the length and the GC-content of the oligonucleotides, which can be calculated by common means known in the art. Under more stringent conditions the annealing temperature is raised above the temperature used in standard conditions to decrease the likelihood of mispriming. The annealing temperature can be raised to a temperature that will still allow enough priming events for the amplification to occur. For instance, if a primer has a melting temperature of 50°C, and an annealing temperature of 45°C would be chosen under standard conditions, an annealing temperature of 50°C would thus be more stringent then 45°C. For example primers of the type $dT_{16}VN$ can lead to strong mispriming under normal RT conditions and therefore decrease the specificity of the cDNA pools synthesized (Mizuno et al., 1999). Analogous results have been reported for AMV and HIV-1 reverse transcriptase. Primer promiscuity can be decreased by increasing temperature during annealing and reverse transcription. For this purpose thermostable enzymes can be used that exhibit reverse transcriptase activity or reverse transcriptase enzymes can be thermo protected by adding thermo protectants such as trehalose to the reaction (Carninci et al., 1998). An added benefit of using trehalose is that the

amount of full length cDNA increases (Carninci et al., 1998). This property was also described by Spiess et al. (2002). In addition Spiess showed that betaine can also increase the amount of full length cDNA and that a combination of betaine and trehalose gave the best result (Spiess et al., 2002). Such RT can be derived from moloney murine leukemia virus (M-MLV) (US 4,943,531 A) or in a more preferred embodiment M-MLV lacking RNAse H activity (US 5,405,776 A). Yet another possibility to raise primer specificity is to include competitive oligo nucleotide blockers as described by Mizuno et al. (1999). Therefore the present invention also comprises a method wherein high stringency conditions are used for the processes of reverse transcription and/or the amplification.

Preferably, in the method according to the invention the selectivity of the reverse transcription and/or amplification reaction, in particular the polymerase chain reaction, is increased by the utilization of trehalose, betaine, tetramethylammonium chloride, tetramethylammonium oxalate, formamide and oligo-blockers, or by the utilization of dimethylsulfoxide during the polymerase chain reaction, to reduce the occurrence of mispriming and increase the amount of full length cDNA.

As all reverse transcriptases have a high error rate it is preferred that a proofreading activity is included in the reaction. Enzymes that have such an activity through their 3'-->5' exonuclease function are e.g. polymerases such as PFU, Ultma, Vent, Deep Vent, PWO and Tli, or E.coli. exonuclease III. In addition to reducing errors introduced into a cDNA during the RT the inclusion of a proofreading activity can also enhance full length cDNA synthesis (Hawkins, 2003).

It is further preferred that such a proofreading activity is also included in the PCR step of the method. On the one hand this will reduce errors introduced into the sequence during PCR as well as extend the size of cDNA molecules that can be amplified (Barnes, 1994). As such a 3'-->5' exonuclease activity can degrade primers used in the reaction it is preferred to protect primers from degradation, e.g. by introducing a phosphorothioate bond at the 3' terminal of the primer (Skerra, 1992; Di Giusto, 2004)or by using LNAs (locked nucleic acids).

It is preferred that during reverse transcription and/or PCR, enzymes or enzyme combinations are used that have a high

processivity and fidelity.

The present invention provides a method for full transcriptome coverage and no redundancy at least in principle. However, methodical properties such as mispriming and polymerase stop can influence the outcome nevertheless. Coverage of long mRNAs therefore depends on the used enzymes. Proof reading enzymes, used for long PCR methods, enlarge the scope of transcriptome coverage to at least 99% under ideal conditions for the given enzymes and method parameters (e.g. long extension times). Misannealing, which can be minimized by e.g. stringency and proof reading activities, may lead to a limited redundancy of the covered transcriptome.

The maximal length of an RNA molecule that can be displayed depends in particular on the quality of the reverse transcription and the amplification reaction, e.g. PCR. In principle long PCR can amplify DNA up to at least 35kb (Barnes., 1994) and RT-PCR up to at least 20kb (Thiel et al., 1997), so at least 99.9% of all transcripts can be displayed. However there exist exceptional long RNA molecules such as titin (NCBI accession number X90568) that has about 80kb length and even though in theory displayable, for practical reasons one will probably not tune the display to such transcripts.

Preferably, the reverse transcription is performed by a reverse transcriptase (RT) derived from the moloney murine leukemia virus, AMV, HIV-1, HIV-2, wherein the RNase H activity of these RTs is or is not present or is reduced. Reverse transcription can be performed by any enzyme that has a reverse transcriptase activity, e.g. Tth DNA polymerase that exhibits reverse transcriptase activity in the presence of $Mn^{2+}$. In addition the RNA can be degraded after the reverse transcription step, for example by an RNA degrading enzyme such as RNase to facilitate long PCR.

Preferably for reverse transcription as well as the amplification designed enzymes can be used, which are independent of a thermal cycling procedure, i.e. are capable of isothermal polymerisation.

It is further preferred that primers used for reverse transcription have a wobble in one or more bases, preferably 1, 2, 3, 4 or 5 bases, on their 3' end preferable between position 2 to 4 upstream of the last dT of the terminal dT stretch to fur-

ther increase specificity.

As used herein, the term "primer" refers to a mixture of oligonucleotides that have the same sequence.

As used herein, the term "wobble primer" refers to a mixture of oligonucleotides that have the same sequence, except for the base or bases that are part of the "wobble".

As used herein, the term "wobble" refers to a base within a primer pool that is present as a mixture of two or three or four different nucleotides.

As an example a wobble primer can consist of oligonucleotides of the sequence ACACACN, where N is referred to as the wobble and N is A or C or T or G. This wobble primer would then consist of a mixture of nucleotides of the sequence ACACACA, ACACACC, ACACACT and ACACACG.

When using N nucleotides as a mixture (N is a mixture of dA, dC, dG and dT) in wobble primers, nucleotides with universal bases can be used instead, such as deoxyinosine, 3-nitropyrrole 2'-deoxynucloside and 5-nitroindole 2'-deoxynucleoside. Universal bases will basepair with any nucleotide (dA, dC, dG, dT).

Therefore, according to the invention the primers used for reverse transcription or amplification preferably have a wobble substitution, e.g. dT may be substituted by dG, dA, dC and vice versa, in one or two bases on their 3' end, preferably between position 2 and 4 upstream of the last nucleotide, which is preferably dT for a 5' RT primer (for 5'-->3' strand synthesis) and a 3' amplification primer and dA for a 5' amplification primer. In the PCR step of the method such "wobble" primers can be used to increase specificity for both the 5' and 3' end of the cDNA molecules to be amplified.

Wobbles can also be used to anchor the primer that is complementary to the (e.g. poly A) tail of the original RNAs to the bases immediately upstream of the tail. In this case the selection into different pools can be achieved by primers selective to bases corresponding to bases after the cap of the original RNA. The same rational can be applied to the 3' end of the cDNA where wobbles can anchor a primer to the bases corresponding to bases downstream of the cap of the original RNA. Therefore the selection into different pools would be done through primers that select for bases that correspond to bases immediately upstream of the (e.g. poly A) tail of the original RNAs.

Even further preferred is a method according to the invention, wherein the wobble base represents a substitution by universal bases, preferably deoxyinosine, 3-nitropyrrole 2'-deoxynucloside or 5-nitroindole 2'-deoxynucleoside, which can basepair to any regular nucleobase such as dA, dT, dG or dC.

In a preferred embodiment specificity of the cDNAs amplified can be enhanced by using a combination of a primer that is exonuclease resistant on its 3' end (e.g. by including a phosphorothioate modification or a locked nucleic acid) with a 3'-->5' exonuclease function, such an activity being a property of e.g. Pfu DNA Polymerase. This will provide an "off switch" for amplification if a single mismatch, or multiple mismatches occur in the first 8 bases from the 3' end of the primer. (Yang et al., 2005). As the bases at the 3' prime end of the primers are the ones that discriminate the cDNA pools in methodologies presented in this invention, it is preferred that such an "off switch" is included to stop the extension of primers that mismatched to a cDNA, thus further enhancing specificity. Preferably in the inventive method a polymerase sensitive for a mismatch between nucleotides of the primer and the RNA or cDNA is used (either during reverse transcription or amplification), and the polymerase does not extend a mismatched exonuclease resistant primer, wherein it is preferred that the primers are resistant to exonucleases, especially preferred by a phosphorothioate modification or a locked nucleic acid.

In a preferred embodiment the primer used for reverse transcription consists of an oligo dT stretch or tail complementary stretch without an anchor. Since the selection of specific RNAs occurs on the cDNA level via the amplification step described above, specificity (of the anchor) at the RT level is not a necessity. Of course mispriming should be avoided, especially where the (oligo T) primer aligns with the sequence upstream of the (poly A) tail stretch of the RNA. Mispriming can be avoided using stringent conditions and/or using an enzyme with proof reading properties. Preferably the annealing temperature during the RT process lies in the range of 30 to 50 °C, depending on the melting temperature of the primer used.

Even though utilization of stringent conditions, such as the use of higher annealing temperatures, wobbled primers and oligoblockers, can decrease mispriming events during reverse tran-

scription (RT), a preferred solution is to not separate into different pools during RT. As the reverse transcriptase reaction allows for more misprimed sequences to be extended than the amplification reaction (e.g. PCR reaction), it is most preferred that the separation into pools is done during the amplification reaction. This also decreases the RT reactions necessary to one per sample to be displayed.

In a preferred embodiment the generation of full-length cDNA can be further advanced. Under non optimal conditions reverse transcriptase might stop polymerisation and/or fall off the RNA strand before reaching the 5' end of the RNA molecule. Generally this is attributed to secondary structure that has not melted and that the polymerase cannot traverse. Therefore generating also partial cDNA's. To lower the probability that such transcripts will interfere with subsequent (PCR) amplification one can exploit the intrinsic ability of reverse transcriptase to add a few ($Mn^{2+}$ and $Mg^{2+}$ dependent 1-6 nucleotides, Schmidt et al., 1999) dC once it reaches the 5' cap end of the mRNA molecule (Clark, 1988). The cap (Furuichi et al., 1975) structure greatly increases the nontemplated nucleotide adding capability of RT compared to mRNAs that do not have the cap structure. This can be used to further enrich for full-length cDNAs. In a subsequent PCR reaction cDNAs can be selected that have these Cs present by including dG nucleotides in the 5' primer at the complementary position. In such a case the tailing can be done with any nucleotide except dC, preferably with dT or dA. The general structure of a 5' PCR primer that would select for cDNAs that have three Cs added to their 5' end, then will be $dT_xdGdGdGHN_y$ or $dA_xdGdGdGHN_y$ or $dC_xdGdGdGHN_y$; H can either be dA or dC or dT.

In a preferred embodiment the method according to the invention, includes the generation of cDNA with a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, during reverse transcription, dependent on a 5' cap of the template RNA or mRNA.

Preferably the ability of a reverse transcriptase to generate the 3' sequence of oligo-dC is increased by the addition of $Mn^{2+}$ ions during the process of reverse transcription. The number of nucleotides added can be increased by including Mn ions in the RT reaction as described in Schmidt et al., 1999. Therefore in a preferred embodiment Mn ions are included in the RT reac-

tion to increase quantitatively the nontemplated addition of dC on the 3' end of cDNA in a cap dependent manner. Subsequent tailing is carried out with any nucleotide but dC in order to allow the selection of the cap dependent dC. The selection of added nucleotides during tailing can be done by providing just the wanted nucleotide during this process. One primer that is used for (PCR) amplification can therefore have the general formula $dP_xdG_y\,HN_z$; $P_x$ representing nucleotides complementary to the tail; dG being deoxyguanylate and y an integer between 0 and 10, preferably between 1 and 5 more preferable between 3 and 5; H can either be dA or dC or dT; and N is a sequence of nucleotides, its members independently selected from either dA, dC, dG, dT; z is an integer between 0 and 10 preferably between 0 and 5.

In a preferred embodiment the cap dependent 3' sequence of oligo-dC is used to isolate fully transcribed cDNA by using a second primer during the amplification which is specific for the oligo-dC sequence. This can be done by using a primer as described above.

Preferably the tailing of the 3' end of the cDNA is performed with a polynucleotide sequence, which is characterized by the absence of dC.

Even more preferred the DNA polymerase used for the amplification reaction is Taq DNA polymerase, Tfl DNA polymerase, aTaq DNA polymerase, Sequenase or Klentaq.

In special embodiments an enzyme with proof reading activity is used, preferably selected from PFU, Ultma, Vent, Deep Vent, PWO and Tli polymerases and E.coli. exonuclease III, is used during the amplification reaction. Such an enzyme can be used in addition to a polymerase, preferably in lower concentrations than the polymerase, or, if it is a polymerase itself, instead. Polymerases with proof reading activity prevent stops during DNA polymerisation and ensure long transcripts with significantly reduced errors. Preferably a combination of such an enzyme with proof reading capability is used with a polymerase above.

In a more preferred embodiment a primer used for the amplification reaction, which is preferably a polymerase chain reaction, is characterized by the general formula $dP_xdG_yHN_z$, wherein $P_x$ represents nucleotides complementary to the tail added to the 3' end of the cDNA, e.g. by the terminal transferase or by the

template switching property of the reverse transcriptase or by any other means, dG being deoxyguanylate and y an integer, including 0, preferably between 2 and 5, H being either dT or dA or dC, N being a sequence of nucleotides with a length z wherein its member nucleotides are independently selected from dA, dC, dG and dT, and z being an integer between 0 and 10, preferably between 0 and 5.

The 5′ end of primers used for RT and PCR can also be anchored to any sequence that will allow RT and PCR to be carried out. 5' and 3' PCR primers may have different melting temperatures and PCR conditions may be optimized by equalizing melting temperatures of primer pairs through the addition of an anchor to one or both of the primers. For certain downstream application it might also be of benefit to add a specific sequence such as a promoter that can start RNA Polymerase transcription (e.g. T7 or T3), or the inclusion of a restriction site.

In a most preferred embodiment the amplification is performed by PCR and the PCR primers contain anchoring sequences, which allow proper placement at the end of the 3' or 5' oligonucleotide end stretch of the cDNA.

Preferably an amplification primer or a reverse transcription primer contains a promoter for RNA polymerases. This optional feature allows the transcription of the cDNA and its PCR products.

Preferably the promoter enhances transcription by the T7, T3 or SP6 RNA polymerase.

Primers as well as PCR products can be labeled by any reporter group known in the art that enables their detection. Examples of reporter groups include fluorescent, chemiluminescent or radioisotopes and others known in the art. Any reporter group can be used that can be detected by such techniques as fluorescent measurement, light emission measurement, scintillation counting and other means known in the art.

Accordingly, the amplification primers and/or the amplification products are labeled by a reporter group, preferably fluorescence markers, chemiluminescence markers or radioisotopes in a preferred embodiment of the present invention. Preferably, these reporter groups are used to detect the DNA.

Recently, it has been shown that during transcription at least some eukaryotes such as Arabidopsis (Jin et al., 2004) can

add arbitrary nucleotides to RNA molecules on their 3' end immediately before the poly A tail, in most cases one or two. And even within a certain gene such nucleotide additions can be heterogenic. Therefore using a set of primers that discriminate for the first and/or second base upstream of the poly A during reverse transcription the intended pooling of a mRNA species into certain pools of cDNA will fail for certain transcripts. However a higher degree of pooling can be achieved using nucleotides upstream of these arbitrarily added nucleotides for the primers that define the fractions. A preferred primer for reverse transcription and PCR amplification that will be complementary to the poly A tail and select for bases above these additions be of following general structure $dT_xVN_yM_z$, where dT depicts deoxythymidylate with x repetitions; V and N are wobble bases where V is either dA, dC or dG ; N is a sequence of dA, dC, dG, dT, independently selected for each member; y is an interger equal to or greater than 0, preferable between 1 and 5, and M is either dA, dC, dG, dT; z is an integer above 1, preferably between 1 and 10. In such a primer $M_z$ will specify the cDNA pool to be amplified. As the number of nucleotides added to RNA molecules before polyadenylation might differ even for a specific transcript, a primer can be a mixture of oligonucleotides where the amount of y in N is different. An additional benefit is that in order for the Polymerase to be able to extend the chain the pressure for correct hybridization of bases after the wobble will strongly increase, therefore decreasing mispriming of the bases that define the fractions that will be amplified, hence specificity will be increased.

In a preferred method according to the invention arbitrary nucleotides, preferably one to three nucleotides, are added to the 3' end of the RNA or mRNA posttranscriptionally but before the addition of a (e.g. poly A) tail, and the primers for reverse transcription and/or PCR contain a sequence with a wobble corresponding to the arbitrary sequence.

A simplified version of the invention that will not specifically select for full length cDNA can be formulated by not using the 3' (dC) nontemplated adding function of the RT enzyme. There will still be a portion of full length cDNAs that have no dC added under normal RT condition. Tailing and PCR can be used as described to selectively amplify pools of cDNA. Each cDNA

will still be present in exactly one pool. However the extent of PCR products present that will correspond to full length RNA molecules will be determined by the quality of the RT reaction alone. As one does not select in the PCR reaction that follows for full length cDNA copies, more than one cDNA molecule might give rise to PCR products that represent a single RNA molecule, therefore introducing redundancy. Still each RNA molecule will be represented at least once. Redundancy can be minimized by employing methods that favour full length reverse transcription, such as addition of trehalose and/or betaine and/or increasing temperature as disclosed above. Even though redundancy might not be abolished in this embodiment, still each RNA molecule will be present at least once. This is an improvement compared to other DD techniques that can reach a representation for a certain mRNA molecule only statistically.

Also there exist species of RNA molecules that do not have a poly A tail and a cap structure. However such RNA molecules will also have 5' and 3' prime end that can be used during subsequent amplification to enrich for defined fractions as is intended by this invention.

Examples of RNA that may or may not have a cap and may have or may not have a poly A tail are non protein coding RNAs (ncRNA) such as ribosomal RNA (rRNA), transfer RNA (tRNA), micro RNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA) and small nuclear RNA (snRNA).

In RNA molecules that do not have a poly A tail, a tail such as a poly A tail, is added synthetically (e.g. enzymatically) prior to reverse transcription. Preferably a polynucleotide tail is added where the nucleotides are selected independently from A or C or G or U.

There exist species of RNA that have a certain size range, such as miRNA that are about 19 to 23 bases long. Therefore size separation of pools of amplification products derived from miRNAs is limited. When displaying miRNAs with a size range of 19 to 23 bases, a maximum of 5 molecules per pool can be resolved according to size. Therefore a quite large primer set (primer matrix) has to be used to resolve all miRNA molecules.

In an alternative approach the nucleotides of one end of the miRNA can be used to artificially enlarge the amplification products for a specific number of nucleotides according to the

terminal sequence of the miRNA. A general formula of such a primer is $P_xN_y$; with P representing nucleotides complementary to a tail; N represents nucleotides independently selected from either A, C, G, T; y is an integer between 0 and 10; x defines the number of nucleotides and is different for all $N_y$. If the 5′ end of the miRNA is used for such a size enlargement, a primer mix containing certain or all possible 5′ primers is used in combination with a 3′ primer that is selective for nucleotides at the end of the miRNA molecule, to define a certain pool. One possible primer matrix according to the principles of this invention is shown in figure 6.

Therefore in a preferred embodiment of the invention, amplification products - in particular full length amplification products - are generated that are differentially enlarged by certain numbers of nucleotides depending on the 5′ or 3′ sequences of the RNA or the full length cDNA.

Preferably, amplification products are generated that are differentially enlarged by defined numbers of nucleotides at their 5' and/or 3' end as compared to the RNA, by using different primers which have different lengths and recognize different 5' or 3' endsequences of the RNA or the full length cDNA. Also, pools of amplification products can be generated that are enlarged by defined numbers of nucleotides at their 5' and or 3' end as compared to the RNA, preferably by using as a 5' and/or as a 3' primer a mixture of primers, where each primer recognizes a different endsequence and has a different length. Preferably the primers have a defined sequence that is not complementary to the cDNA (during the amplification step) or RNA (during the reverse transcription step).

In prokaryotes mRNA has no poly A tail. Methods have been devised (Amara et al., 1997) that can make full-length cDNA from such molecules, such as adding a synthetic poly A tail. It is also within the scope of this invention to apply said methods to such molecules.

Preferably in a method according to the invention RNA or mRNA samples are used, which do not have a poly A tail or a tail sequence and, preferably a polynucleotide tail, especially preferred a poly A tail, is added synthetically (enzymatically) prior to reverse transcription.

To decrease the necessary amount of RNA to perform the dis-

play a preamplification step can be introduced before the amplification reaction, e.g. the PCR. Whether or not separation into pools was done in the reverse transcription (RT) step, the preamplification step has to be designed in a way that will amplify all the cDNA that is intended to be separated into defined pools in the amplification (PCR) step. The preamplification step can be introduced before the tailing or after the tailing, before the main amplification with the specific primers.

For example a promoter such as T7 RNA Polymerase can be added to the 5'- end of the primer that is used for priming the RT reaction. After the RT and before the amplification (PCR) RNA can be synthesized by an RNA polymerase specific to the promoter, for example T7 RNA polymerase. Such obtained amplified RNA has to be reverse transcribed again and depending whether or not a sequence is present downstream of the cDNA base that is complementary to the first base of the RNA and that can be used to prime the amplification reaction (PCR) on this end, such a sequence has to be added, for example by the template switching ability of the reverse transcriptase or by tailing with terminal deoxynucleotidyltransferase.

Another example for a preamplification step is the use of a preamplification PCR. In case of a terminal deoxynucleotidyltransferase tailed cDNA one primer has to be specific to the so added tail. For example a poly T tailed cDNA will need a primer such as $dA_x$ and a primer complementary to the poly A tail of the original RNA for example $dT_x$. Preferably one can preamplify only full length cDNA by including variable amounts of G, preferably between 1 and 5, to enable the primer to hybridize only to cDNA that was fully reverse transcribed and thus where the reverse transcriptase added Cs through its non templated nucleotide adding capacity in a cap dependent manner. Additionally the primer that can hybridize to the sequence complementary to the original RNA poly A tail can be anchored to the 5'-end of the poly A tail by including one or more wobbles. This will help to increase the length of the amplification (PCR) products obtained, by not needlessly amplifying long stretches of poly A tail. Such a primer has the general formula of $dT_xVN_y$ (x is the amount of dT, V either G, C or A, and N any nucleotide).

The discrimination of the cDNA into different pools can also be accomplished by using more than one amplification (PCR) step.

Each amplification (PCR) step would be more specific to the pool to be amplified. For instance in a first amplification (PCR) a primer such as $dT_x$ could be used for hybridizing to cDNA on the side that is complementary to the poly A tail side of the RNA. This primer would be not discriminating. In a next step, 3 reactions could be used with one being primed with $dT_xG$, one with $dT_xA$ and one with $dT_xC$. Thus creating 3 pools. In another PCR step the pools can be further divided. For example the $dT_xG$ pool can be divided in to 4 pools by priming with $dT_xGA$, $dT_xGC$, $dT_xGG$ and $dT_xGT$. Thus creating a total of 12 pools. The same rational can be applied to the other end of the cDNA and the pooling of both sides can be combined if desired. Such a sequential pooling can on the one hand further decrease the amount of starting material (RNA) needed or help to achieve a pooling tuned to the complexity of the sample to be investigated or the depth the investigator wants to cover.

It is within the scope of the invention that two or more samples of RNA that are combined by methods that yield normalization or subtraction between such RNA or their derived cDNA or PCR products can also be displayed according to disclosed methods. For example RNA can be (pre)treated by reducing RNA species in high concentration.

Therefore the present invention also relates to a method wherein the RNA in the sample or the cDNA is preamplified, preferably by PCR, or wherein RNA samples are normalized or subtracted. These (pre)treatment steps are preferably introduced before the main (discriminating) amplification c).

In a further embodiment of the method according to the invention the step of fractionation of the amplification product is performed by gel electrophoresis, pulsed-field electrophoresis, agarose gel electrophoresis, PAGE, capillary electrophoresis or pulsed-field capillary electrophoresis. Fractionation and comparison of samples generated during PCR can be done by any methods that allow distinguishing DNA molecules according to specific properties, such as sequence or size. Such methodologies include e.g. gel electrophoresis, such as polyacrylamide gel electrophoresis (PAGE). Especially suited is capillary gel electrophoresis because of its high resolving power and because many samples can easily be analyzed in parallel (Behr et al., 1999). Standard capillary electrophoresis machines today can

process 384 samples in one run. And through the use of appropriate protocols one can achieve the necessary resolving power over a wide range of sizes. Especially long fragments can be resolved sufficiently with pulsed-field capillary electrophoresis (Heller et al., 2001; Morris et al., 2001). Also methods have been devised to collect fractions of the molecules separated in a capillary (Irie et al., 2000; Magnusdottir et al., 2001). Such fractions that represent transcripts of interest (e.g. differentially expressed genes) can then be subjected to further analysis, such as sequencing, to determine the identity of the molecule.

In a preferred embodiment the detection of differences between two cDNA pools can be included into the PCR step of the method. Such methodologies, that can measure DNA levels and/or the change of such levels during PCR, are generally referred to as "real time PCR"(e.g. Wittwer et al. 1997a, Wittwer et al. 1997b). One procedure applicable measures fluorescence from SYBR Green I dye, that has been added to the PCR reaction. This dye will show a strong increase in fluorescence when intercalating with double stranded DNA. The more DNA present in the reaction the stronger the fluorescence. Therefore providing the means to calculate the amount of DNA present in the reaction. The measurement is usually done at the end of the polymerization step as the products will be double stranded. As the amount of DNA increases during subsequent cycles, so will the fluorescence. This data can be compared e.g by blotting on a curve and curves can be compared between different samples. If there is only one cDNA amplified the curve will have an exponential phase. By comparing the start of the exponential phase between samples one can calculate relative amounts of DNA. If two or more products will be present in a given sample, curves will add onto each other and will result in a change of steepness. Still such data can be compared between different samples, if the cDNA pool amplified is not too complex. It is preferred that the amount of cDNAs amplified in one PCR reaction is reduced to such a point that differences in amplification of each product can still be detected. Another way of analysing a sample, when using real time PCR, is the possibility of calculating a melting curve of the DNA present at the end of each cycle. Each PCR product present in the reaction will melt at a specific temperature,

that is defined by the sequence. As a product is melting into its single strands the dye is not intercalating with DNA anymore, whereby a drop of fluorescence will result. Products that will melt at different temperatures will be represented as an increase and then a decrease of the steepness at different areas of the melting curve. Comparing curves between cycles and different samples of the same cDNA pool to be amplified, one can identify samples with a different profile. A melting curve analysis can also be done only at the end of the PCR, however differences in products that have plateaued might not be detected.

The use of such analysis methods preferably in an automated procedure can rapidly screen for samples that show differences during or at the end of PCR and eliminate the need to analyse all of them through a fractionation procedures such as gel electrophoresis.

Therefore in a preferred embodiment the detection of differences between two cDNA pools, preferably corresponding pools between two or more samples, is done during or at the end of the amplification step, preferably being a PCR step, of the method, preferably by real time PCR.

Clearly the methods described in this invention can be integrated into a partially or fully automated process. For example the size fractionation, comparison of different samples, collection of fractions, and sequencing can be all done on a chip. The technologies have already been demonstrated in principle (Xie et al., 2001).

Such an integration of the method onto a chip can also be done for the PCR step of the method. It is preferred that the primers, or primer pairs are spotted and/or bound onto a solid support to enable solid phase DNA amplification (e.g. by PCR). Through microarraying or nanoarraying or single molecule arraying the whole set of primers onto a solid support, a single reaction vessel can be created that contains the whole primer matrix or parts thereof. Therefore instead of using an own reaction vessel for each 5′-3′ primer pair, a single reaction vessel contains all 5′-3′ primer pairs. It is preferred that each primer pair is located in a defined area.

A primer matrix can also be created where each 5′-3′ primer pair is bound onto its own solid support (e.g. a bead) and the solid support itself is then arranged in such a manner that each

primer combination or primer pair can be addressed independently. For instance a matrix of beads can be arranged randomly in microwells and the location of each bead decoded by methods known to the art (e.g. Gunderson et al., 2004). The method as described by Gunderson can thus be adapted and incorporated by the present invention. This method requires only a few labels and several sequential hybridizations to identify thousands of different DNA sequences with great accuracy. Regarding the primer matrix, the primers on the areas may comprise the information for decoding themselves (e.g. as sequence which can be decoded by other labeled oligo nucleotide by several steps of hybridisation, detection and washing) or together with the primers of the invention decoding oligonucleotides can be bound onto the carrier.

Therefore in a preferred embodiment a 5'primer or a 3'primer or a 5'-3' primer pair is spotted and/or bound onto a solid support, preferably the spot or the solid support is located in a defined 2-dimensional or 3-dimensional area and it is preferred that such an area contains information to decode the identity of the primer or primer pair located in that area.

In a preferred embodiment the primer matrix is generated by direct on chip synthesis of the 5' primers or 3' primers or 5'-3'primer pairs.

In a preferred method according to the invention the steps of reverse transcription and amplification are performed in one step.

It is obvious to the one skilled in the art that other possibilities exist to enrich for full-length cDNA. Such methods can easily be adapted to amplify subsequently pools of full length cDNA that are defined through nucleotides that are complementary to the 5' end and/or nucleotides on the 3' end upstream of the poly A tail.

For example the CapSelect method (Schmidt et al., 1999) can be used to substitute the steps of reverse transcription and tailing the 3' end of the cDNA. This method is also drawn on the nontemplated addition of C nucleotides of reverse transcriptase, and addition of a poly N tail, preferably a poly A tail. However the poly A tail consists of ribonucleotides instead of deoxyribonucleotids as it is claimed that terminal transferase will preferably add only 3 to 4 ribonucleotides, compared to an

unspecified amount of deoxyribonucleotids. This is important for the method as in a next step an adaptor will be specifically linked to the 3' end of the cDNA selecting for 3'-AAACCC-5'. Using a primer that is complementary to this adaptor and a primer complementary to the poly A tail it is intended to amplify full length cDNA. Instead of using universal primers to amplify all cDNAs one can use sets of primers that select for the 5' and/or 3' nucleotides and will amplify pools of cDNAs. The primer structure for hybridizing to the poly A tail and selecting for x amount of bases before the poly A tail will be the same as described. For amplifying the 5′ end of mRNAs one can use the adaptor sequence followed by 3 to 4 Ts and 3 to 4 Gs and an x amount of bases that correspond to the 5′ sequence of the RNA transcripts. Increasing x will decrease the number of transcripts amplified through such a primer.

Preferably the polynucleotide tail added by the terminal transferase consists of ribonucleotides, preferably 2 to 6 ribonucleotides, preferably adenosine or thymidine, after a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, which is generated during reverse transcription, dependent on a 5' cap of a template RNA or mRNA, and the second amplification primer (5' primer) comprises complementary nucleotides to the ribonucleotides and oligo-dCs and 1 to 10 nucleotides upstream of the olio-dC sequence.

The method for amplification of certain cDNAs, derived from mRNA, can also be used to determine the expression level of individual mRNAs. Preferably for measuring the expression level of messenger RNA a method according to the invention is employed, including the separation and/or identification of amplification results. Like in common DD results of certain cell lines can be compared for better visualisation of different expressions.

Furthermore the present invention provides a kit for a method according to the invention comprising a DNA polymerase, preferably Taq polymerase, a reverse transcriptase or mixtures of different polymerases, cofactors or salts of metal ions, preferably $Mg^{2+}$ and $Mn^{2+}$, requested by a polymerase, primers and optionally a terminal transferase and additional buffer substances. The primers are preferably selective for the 5' and 3' end of a RNA or cDNA as described above. Even more preferred the primers of the primerset are selective for the first, second,



IPR2025-01353; IPR2025-01355
CSHL EX2002

third, fourth and/or fifth terminal nucleotide of the RNA or cDNA (of either end, independently from each other). Preferably the kit comprises 1, 2, 3, 4, 5, 10, more than 10, more than 20 or more than 30 primer pairs. With this kit a display method according to the invention can easily and quickly be performed. Preferably instructions for the execution of the method are also included. Preferably the full primerset (primer matrix) or parts thereof - preferably one of each primer pair - are bound to a solid support such as a chip, preferably a glass chip.

In another aspect of the present invention there is provided a solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a pair of primers, wherein different primers are selective for at least two, preferably at least three, even more preferred at least four, most preferred at least 10, different 3' or 5' end sequences of target polynucleotides, preferably selected from RNA or cDNA, and each area comprises primers for at least two different end sequences (e.g. a pair of a 5' and a 3' primer), preferably one primer is selective for a (e.g. poly A) tail or a complement thereof and even more preferred selective for the next 1 to 10 nucleotides, preferably the next 1, 2, 3 or 4 nucleotides (of either the 5' and/or the 3' primer. The primers can be specific as described above. These primers may also, e.g. in these combinations, be used in the kit or the method. A primer pair in one area selective for mRNA is specific for the poly A tail and the 5' (cap) end, as well as the next 1 nucleotide (different in each area). Preferably the solid surface comprises at least two, at least 4, at least 10, at least 20 or at least 40 different primer pairs.

Provided is further a solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a 5′ primer and a 3′ primer or a pair of primers, wherein the primers are as defined above being specific for both ends of a cDNA or a RNA, especially comprising formula $dP_xdG_y\,HN_z$; in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or primer pairs.

Also provided is a solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas,



IPR2025-01353; IPR2025-01355
CSHL EX2002

preferably in spots, a 5′ primer or a 3′ primer comprising the general formula as defined above ($dP_xdG_y Hn_z$) or a complement thereof, preferably selected from a primer comprising one of the following sequences: AGA GAT TTT TTT TTT TTT TT GA, AGA GAT TTT TTT TTT TTT T GG, AGA GAT TTT TTT TTT TTT T GC, AGA GAT TTT TTT TTT TTT TT GT, AGA GAT TTT TTT TTT TTT TT CA, AGA GAT TTT TTT TTT TTT TT AG, AGA GAT TTT TTT TTT TTT TTT AA, AAA AAA AAA AAA AAA GGG ATA, AAA AAA AAA AAA AAA GGG ACA, AAA AAA AAA AAA AAA GGG AGA, AAA AAA AAA AAA AAA GGG AAA, AAA AAA AAA AAA AAA GGG GTA, AAA AAA AAA AAA AAA GGG CTA, AAA AAA AAA AAA AAA GGG TTA, in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or primer pairs.

In a further aspect a method for amplifying a pool of polynucleotide molecules in a sample is provided, comprising the steps of:

a) obtaining (or providing) a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining (or providing) a sample of full length cDNA,

b) tailing the 3′ end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides (i.e. nucleotide 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10), upstream of the 3' polynucleotide tail of the cDNA, preferably also comprising unspecific wobble nucleotides, especially preferred in the upstream portion within the 1 to 100 nucleotide range - wherein at least 1, 2 ,3, 4, 5, 6, 7, 8, 9 or 10 nucleotide (s) is/are specific for a nucleotide of the cDNA; the method further preferably defined as given above.

The method preferably includes the use of a primer, which is specific for a 3' tail of the RNA (optionally an added 3' tail) or mRNA (preferably the endogenous 3' poly A tail) and the next

1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides (i.e. nucleotide 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10), upstream of the 3' tail is used for the reverse transcription and/or a first amplification primer is specific for a 5' tail of the cDNA, which is complementary to the 3' tail of the RNA or mRNA, and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides (i.e. nucleotide 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10), downstream of the 5' tail. Of course, also both or independently from one another, in the 5' and the 3' primers, the nucleotide stretch (1 to 100 nucleotides) next to the tail (or tail complement) may comprise unspecific nucleotides (such as wobble nucleotides) in addition to at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more specific nucleotides selecting for different pools (with or without pool separation) of the amplification product. Such primers can also of course be used for the reverse transcription as well. Products of such wobble primers (where any nucleotide may be present at a corresponding position to the wobble position) still fulfil the requirement of full-length cDNA or full-length amplification product although the sequence may not be preserved at positions that correspond to the unspecific nucleotide positions. Using longer unspecific nucleotide primer stretches primer-loops may occur thus resulting in a product with an altered length which is however still considered to be equivalents to full-length cDNA or amplification products for the present invention.

The present invention is further illustrated by the following example and figure without being limited thereto.

Fig. 1 shows a graphical representation of a procedure using the principles of the present invention to amplify full length cDNA into defined subpopulations. The oligo-T and oligo-G stretches indicated in the primers are terminal anchors of the primers and the N, NN and $NN_y$ indicate nucleotides selected from A, T, G, C for defining the subpopulations.

Fig. 2: Gel elecrophoresis product of a display of a pool of artificial mRNA molecules according to example 2: In lane 1-7 primer nr. 8 was used together with primers nr. 1-7. In lanes 8-13 primer nr. 1 was used together with primers 9-14.

Fig. 3: Gel elecrophoresis product of an organised full length expression display of mouse liver RNA according to example 3.

Fig. 4: Gel elecrophoresis product of an organised full length expression display of mouse liver RNA according to example 4 with PCR preamplification.

Fig. 5: NCBI nucleotide database screening for non redundant full length mouse cDNA clones. 99% of the 16854 clones present in the database are between 200 and 6000 bases in length and 99.9% below 7500 bases and no clone was above 9500 bases.

Fig. 6 shows how RNA without a cap and a tail, such as miRNA, is divided into subpopulations. a) Steps of the procedure. Primer nomenclature as in fig.1 b) One possible Primer matrix that increases the resolving power by increasing the size range of the amplification products. The B in the 5′ primers indicate a mixed base (wobble) that is selected from T, G, C.

E x a m p l e s

Expample 1: Principle of procedure

An overview of an example of a polynucleotide amplification procedure according to the present invention is given in fig. 1. Anchored primers for reverse transcription can follow the formula $dT_xV$ or $dT_xVN_y$ (with x defining a repetition of deoxythymidylate ("dT") that has in combination with the V or $VN_y$ part a high enough specificity and melting temperature to successfully prime the RT reaction; V is either deoxyguanylate ("dG"), deoxyadenylate ("dA") or deoxycytidylate ("dC"); N is a sequence of y nucleotides selected independently from dA, dC, dG or dT; if more than one base is used for anchored primers y defines their number; increasing y will increase the subpopulations of reverse transcripts. The tailing of the 3′ end of such synthesized cDNA with a poly N tail can be done using terminal transferase – N stands for a sequence of uniformly nucleotides selected from dT, dA, dC or dG, independently form the "N" indicated in the reverse transcription primers. The amplification of the cDNA pools, optionally preselected by the $VN_y$ stretches defined by polymerase chain reaction (PCR) is performed using primers that hybridize to the 5′ and 3′ end of the cDNA by selecting to terminal sequences like the nucleotides appended by the terminal

transferase (3' of the cDNA; corresponding to the 5' end of the original mRNA) and the poly T stretch (5' of the cDNA; complementary to the 3' poly A tail of the original mRNA) and selecting for an oligo $dN_y$ sequence, wherein N is a stretch of several nucleotides, again independently selected from dT, dA, dC or dG and define different mRNA populations through y bases immediately adjacent to the tails.

Example 2: Amplification of all the RNA molecules present in a sample.

As all the mRNA molecules present in a sample are not known, even in well characterized tissues such as liver from human or mouse, and their number maybe in the thousands, an approach was chosen to use a pool of 14 artificially synthesized mRNA molecules to show that all RNA molecules present in that sample are detected. For this purpose stretches with different lengths of genomic DNA from the mouse GAPDH gene (GenBank Accession: NC_000072) were amplified by standard PCR. 5' primers were used that had an anchored T7 polymerase promoter. Following the promoter a combination of 4 bases were included that would define each mRNA on its 5' end. Primers chosen for the 3' end included, after the GAPDH specific sequence, a combination of 4 bases, that would define these molecules upstream of the poly A tail. Following these 4 bases 21 dT or in one of the RNAs 22 dT were included for the poly A tail. The PCR products obtained were gel purified and subjected to standard T7 RNA transcription, that included 7-methyl-guanosine (the cap). Such T7 reaction products were again gel purified. A mixture of such obtained synthetic mRNA molecules was used as starting material for the reverse transcription (RT). To show that the methods disclosed are specific for the 5' as well as 3' end of the mRNA molecules, 7 of the 14 molecules had a fixed sequence on the 5' end and varied in the 3' end and 7 molecules had a fixed sequence on the 3' end and varied at the 5' end. For sequence of the resulting artificial mRNA molecules see table 1.

Reverse transcription was carried out in a 25µl reaction consisting of 50mM Tris-HCl (ph 8.3), 75 mM KCl, 3mM $MgCl_2$, 10mM DTT, 0.6M trehalose, 2M betaine, 0.5mM dATP, 0.5mM dTTP, 0.5mM dCTP, 0.5mM dGTP, 40nM primer (oligo $dT_{(21)}$). The artificial mRNA mix was included at a level of about 4 ng/µl reaction volume.

The reaction (without the RT) was heated to 70 °C for 20 sec, 32°C for 2 min. Then the Molony Murine Leukemia Virus Reverse Trancriptase lacking RNase H activity (M-MLV RT (-H)) was added a concentration of 4-8 units/µl reaction volume and the reaction continued for 2 min at 32°C. At this point 0.5 U of Pfu polymerase was added. Four cycles at 50°C for 10min and 60°C for 10 sec were used to fully reverse transcribe the RNA, after that the reaction was cooled to 4°C.

Table 1: Sequences of artificial RNAs to be displayed

| RNA Nr. | 5' sequence ..... 3 sequence | Length of RNA | Predicted Length of PCR product |
|---|---|---|---|
| 1 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CTG TGG GCA AGG TC **TC** $A_{(21)}$ | 1515 | 1531 |
| 2 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....TGT CCC TTT TGG TC **CC** $A_{(21)}$ | 1333 | 1348 |
| 3 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....TTG TGG AAG GGC TC **GC** $A_{(21)}$ | 1243 | 1258 |
| 4 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CGA TGC CCC CAT TC **AC** $A_{(21)}$ | 1112 | 1128 |
| 5 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CAC CAC CAT GGA TC **TG** $A_{(21)}$ | 942 | 958 |
| 6 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CAA CGG GAA GCC **TC TA** $A_{(21)}$ | 843 | 859 |
| 7 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....ACG ACC CCT TCA TC **TT** $A_{(21)}$ | 736 | 753 |
| 8 | $^{m7}$G **AT** AT GAT GCG CAA AGG.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 687 | 703 |
| 9 | $^{m7}$G **AC** AT GGC CTA AGC AAG.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 786 | 802 |
| 10 | $^{m7}$G **AG** AT GAG CAC GGA AGC.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 880 | 896 |
| 11 | $^{m7}$G **AA** AT CCG CTT TGA TTT.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 994 | 1010 |
| 12 | $^{m7}$G **GT** AT CAG CTC CCC ATC.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 1070 | 1086 |
| 13 | $^{m7}$G **CT** AT GGG CCT GGG TCA.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 1178 | 1194 |
| 14 | $^{m7}$G **TT** AT GGG CGG AGT GGA.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 1287 | 1303 |

The cDNA was phenol/chloroform extracted, ethanol precipitated and dissolved in 10 µl of $H_2O$. Poly T tailing was carried in a 20 µl reaction containing 0.2mM dTTP, 100mM cacodylate buffer (pH 6.8), 1mM $CoCl_2$, 0.1mM DTT and 15 units of terminal deoxynucleotidyl transferase for 15 min at 37 °C. The reaction was Phenol/Chloroform extracted, ethanol precipitated and dissolved in 100 µl $H_2O$.

PCR was carried in a 10 µl reaction containing 2.5µl tailed cDNA, 50mM Tris (pH 8.8), 14mM $(NH_4)_2SO_4$, 4mM $MgCl_2$, 0.2mM of each dNTP, 0.2 µM of each primer (table 2, fig. 2), 0.5 units Taq

polymerase and 0.03 units Pfu polymerase. Samples were denatured at 93°C for 30 sec, and cycled 21 times at 93°C for 10 sec, 55°C for 30 sec, 72°C for 10 min. The initial annealing temperature of 55°C was decreased each cycle by 0.5°C min. Thereafter followed 19 cycles at 93°C for 10 sec, 49°C for 30 sec, 72°C for 10 min with a final extension step at 72°C for 2 min.

Table 2: PCR primers

| Primer nr. | Sequence |
|---|---|
| 1 | AGA GAT TTT TTT TTT TTT TT **GA** |
| 2 | AGA GAT TTT TTT TTT TTT T **GG** |
| 3 | AGA GAT TTT TTT TTT TTT T **GC** |
| 4 | AGA GAT TTT TTT TTT TTT TT **GT** |
| 5 | AGA GAT TTT TTT TTT TTT TT **CA** |
| 6 | AGA GAT TTT TTT TTT TTT TT **AG** |
| 7 | AGA GAT TTT TTT TTT TTT TTT **AA** |
| | |
| 8 | AAA AAA AAA AAA AAA GGG **ATA** |
| 9 | AAA AAA AAA AAA AAA GGG **ACA** |
| 10 | AAA AAA AAA AAA AAA GGG **AGA** |
| 11 | AAA AAA AAA AAA AAA GGG **AAA** |
| 12 | AAA AAA AAA AAA AAA GGG **GTA** |
| 13 | AAA AAA AAA AAA AAA GGG **CTA** |
| 14 | AAA AAA AAA AAA AAA GGG **TTA** |

2.5 µl of the reaction product was added to 2.5 µl 100% formamide loading buffer and denatured at 96 °C for 2 min and cooled on ice, loaded on a 3.5 % acrylamide, 7M urea gel and run at 180V for 2h and subsequently silver stained.

The results can be seen in figure 2. In lane 1-7 primer nr. 8 was used together with primers nr. 1-7. In lanes 8-13 primer nr. 1 was used together with primers 9-14. In each lane the PCR amplified the specific product, as the predicted length of the PCR products shown in table 1 perfectly matches the lengths of the products on the gel.

The primer combination in lane 1 amplifies the 2 specific products corresponding to the RNA to be displayed. The reaction carried out in lane 1 is the most demanding reaction of the set

as two products have to be amplified, the ones representing RNA 1 and 8. In addition as can be seen in table 1 the RNA molecules were designed in a way to enhance miss priming of the PCR reaction represented in lane 1. All RNA molecules (except RNA Nr.6)have on position 3 and 4 before the T tail the same bases as on position 1 and 2 of RNA number 1 and 8. As primer nr. 1 uses its last two bases to select the cDNA to be amplified, it potentially can miss prime with any of the other cDNA present in the sample. And as cDNA moleclules 1-8 have the same bases on its other end it is only the two last bases of primer 1 that selects between these 8 molecules in reaction 1. However only the two RNAs to be amplified are displayed, showing the great specificity that can be reached in such a display.

Lane 2 shows the limits of this reaction. In addition to the band to be amplified a faint band that corresponds to RNA nr. 4 was amplified. However the band to be amplified is several times more intense than the band that resulted through the miss priming event. Additionally even though a touch down protocol was used for the thermal cycling, the same profile was used for all the different primer combinations. To show that in principle the reaction can be optimized for a specific primer set another reaction was carried that included 1.302M betaine and 1.3% DMSO. The product is shown in lane 14 and only the correct band is displayed.

Example 3: Amplification of total RNA from mouse liver.

Total RNA (13µg) from mouse liver was reverse transcribed, tailed and PCR amplified under the same conditions as for the whole artificial RNA set in example 2.

The same conditions for RT, tailing and PCR were used to show that it is possible to display RNA molecules from a complex mixture under the same stringent conditions as it is possible with artificial RNA. The PCR products were run on a 0.7% agarose gel to resolve longer molecules. As can be seen in figure 3 each primer combination amplifies a unique profile.

The length distribution of the RNA molecules displayed in figure 3, matches very well the length distribution of cDNA in the NCBI nucleotide database screening for non redundant full length mouse cDNA clones (fig. 5).

As on averaged about 10-15 bands can be seen on each lane

the full set of 768 primer combinations ($3x4^4$) displays about 7680-11520 PCR products representing full length RNA molecules expressed in liver. Being able to resolve such profiles on a simple 0.7% agarose gel shows the power of the method and the ease, with what it can be performed in the laboratory.

Example 4: Utilization of PCR Preamplification step.

Preamplification was carried out in a 10µl reaction containing 2.5µl of a purified tailed cDNA generated as in example 3, 50mM Tris (pH 8.8), 14mM $(NH_4)_2$ $SO_4$, 4mM $MgCl_2$, 0.2mM of each dNTP, 0.2µM of Primer $dT_{(21)}$, 0.2µM of Primer $dA_{(21)}$, 0.5 units Taq polymerase and 0.03 units Pfu polymerase. Samples were denatured at 93°C 30 sec, and cycled 11 times at 93°C 10 sec, 42°C 30 sec, 72°C 11 min with a final extension step at 72°C for 2 min. This reaction was diluted to 1ml by adding 990µl $H_2O$. 2.5µl of this dilute was used in the following PCR using the same PCR conditions as in example 3 with the difference that the final 2 cycles where omitted. As can be seen in figure 4 a distinctive banding can be seen for each primer combination. The preamplification step as used in the example shows that 500ng of total RNA are enough to provide material for even a large display with 768 reactions per RNA sample, as might be needed to resolve highly complex RNA mixtures from tissues like brain. The amount of RNA used can easily be reduced by increasing the cycles during preamplification. Therefore the sensitivity of the display should meet even the most demanding applications, where only limited amounts of RNA are available. Furthermore the amount of bands per primer pair has increased compared to example 3. As the high annealing temperatures used in the discriminating PCR step, in both example 3 and 4, exert an extremely high stringency, some low level transcripts might be just below the detection limit when no preamplification step is used.

Expample 5: miRNA Display

An overview of an example of a polynucleotide amplification procedure according to the present invention, where the RNA has no cap and no poly A tail, in this case microRNA (miRNA), is given in fig. 6.

To be able to reverse transcribe full length cDNA from miRNA that has no Poly A tail, a synthetic tail is added by using poly

A polymerase. In a second step the Poly A tailed RNA is reverse transcribed into cDNA using an oligo dT primer. After polynucleotide tailing of the cDNA by terminal transferase, the cDNA samples have sufficient long nucleotide terminals that allow for primers to anneal and amplify the cDNA sequences into different pools according to the 5′ and 3′ sequences of the original miRNA molecules (fig. 6a).

As microRNA (miRNA) molecules have a size of about 19-23 bases, a sufficiently large primer matrix (all possible combinations of 5′ and 3′ primers) has to be used to separate all miRNAs into different pools where differences in expression between samples can be measured.

Alternatively one end of the miRNA sequence can be used to artificially enlarge the amplification products for a specific number of nucleotides according to the terminal sequence of the miRNA. Figure 6b shows an example of such a matrix. In this case the 5′ primer is used to split miRNA amplification products into different lengths according to the second and third nucleotide of the miRNA. In cases where the comparison of different pools is done by an additional size fractionation (such as capillary gel electrophoresis) all 16 5′ primers of fig. 6b can be used as a primer mix. This 5′ primer mix together with one 3′ primer defines a pool. Therefore 192 pools have a similar resolution when compared to 192 x 16 = 3072 pools, where each 5′ primer would be used together with each 3′ primer.

R e f e r e n c e s

Amara et al., 1997, Nucleic Acids Res. 25, 3465-3470.
Barnes, 1994, PNAS 91, 2216-2220.
Behr et al., 1999, Electrophoresis 20, 1492-1507.
Carninci et al., 1998, Proc. Natl. Acad. Sci. U. S. A 95, 520-524.
Clark, 1988, Nucleic Acids Res. 16, 9677-9686.
Di Giusto et al., 2004 Nucleic Acids Res. 32, e32 1-8.
Frohman et al., 1988, Proc. Natl. Acad. Sci U. S. A 85, 8998-9002.
Furuichi et al., 1975, Nature 253, 374-375.
Gunderson et al., 2004 Genome Res. 14, 870-877.
Hawkins et al., 2003, Biotechniques 34, 768-773.

Heller et al., 2001, Methods Mol. Biol. 162, 293-305.
Irie et al., 2000, Electrophoresis 21, 367-374.
Jin et al., 2004, RNA. 10, 1695-1697.
Kovarova et al., 2000, Nucleic Acids Res. 28, e70i-e70iii.
Liang et al., 1992, Science 257, 967-971.
Magnusdottir et al., 2001, Methods Mol. Biol. 162, 323-331.
Matz et al., 1997, Nucleic Acids Res. 25, 2541-2542.
Mizuno et al., 1999, Nucleic Acids Res. 27, 1345-1349.
Morris et al., 2001, Methods Mol. Biol. 162, 307-321.
Prashar et al., 1996, Proc. Natl. Acad. Sci. U. S. A 93, 659-663.
Prashar et al., 1999, Methods Enzymol. 303, 258-272.
Rebagliati et al., 1985, Cell 42, 769-777.
Schmidt et al., 1999, Nucleic Acids Res. 27, e31.
Spiess et al., 2002, Anal.Biochem. 301, 168-174.
Skerra, 1992, Nucleic Acids Res. 20, 3551-54.
Thiel et al., 1997, Anal.Biochem. 252, 62-70.
Welsh et al., 1992, Nucleic Acids Res. 20, 4965-4970.
Wittwer et al., 1997a, Biotechniques 22, 130-138.
Wittwer et al., 1997b, Biotechniques 22, 176-181.
Xie et al., 2001, Methods Mol. Biol. 162, 67-83.
Yang et al., 2005, Biochem. Biophys. Res. Commun. 328, 265-272.

C l a i m s :

1. Method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of:

a) obtaining a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining a sample of full length cDNA,

b) tailing the 3′ end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 10 nucleotides upstream of the 3' polynucleotide tail of the cDNA.

2. Method according to claim 1 characterized in that a primer, which is specific for a 3' poly A tail of the RNA or mRNA and the next 1 to 10 nucleotides upstream of the poly A tail is used for the reverse transcription and/or a first amplification primer is specific for a 5' poly-T stretch of the cDNA, which is complementary to the 3' poly A tail of the RNA or mRNA, and the next 1 to 10 nucleotides downstream of the 5' poly-T stretch.

3. Method according to any one of claims 1 or 2 characterized in that the tailing of the 3' end is performed using terminal transferase.

4. Method according to any one of claims 1 to 3 characterized in that the amplification is performed by PCR.

5. Method according to any one of claims 1 to 4 comprising an additional step of separation of the amplification products according to their length, preferably by gel electrophoresis, preferably polyacrylamide gel electrophoresis, capillary electrophoresis, pulsed-field electrophoresis, agarose gel electrophoresis, PAGE or pulsed-field capillary electrophoresis.

6. Method according to any one of claims 1 to 5 comprising an additional step of determining the identity or sequence of the

amplification products, preferably by an automated process on a chip.

7. Method according to any one of claims 1 to 6, wherein the sample contains total RNA or mRNA, preferably purified RNA or mRNA, from a specimen.

8. Method according to any one of claims 1 to 7, wherein high stringency conditions are used for the processes of reverse transcription and/or the amplification.

9. Method according to any one of claims 1 to 8, wherein the selectivity of the reverse transcription and/or amplification, preferably a polymerase chain reaction, is increased by the utilization of trehalose, betaine, tetramethylammonium chloride, tetramethylammonium oxalate, formamide and oligo-blockers, or by the utilization of dimethylsulfoxide during the polymerase chain reaction, to reduce the occurrence of mispriming.

10. Method according to any one of claims 1 to 9, wherein the reverse transcription is performed by a reverse transcriptase (RT) derived from the moloney murine leukemia virus, AMV, HIV-1, HIV-2, wherein the RNase H activity of these RTs is present, reduced or not present.

11. Method according to any one of claims 1 to 10, wherein primers used for reverse transcription or amplification have a wobble substitution, preferably dT is substituted by dG, dA or dC or vice versa, in one or two bases on its 3′ end, preferably between position 2 and 4 upstream of the last nucleotide, wherein the wobble base preferably represents a substitution by universal bases, preferably deoxyinosine, 3-nitropyrrole 2'-deoxynucloside or 5-nitroindole 2'-deoxynucleoside, which can basepair to any regular nucleobase such as dA, dT, dG or dC.

12. Method according to any on of the claims 1 to 11, using a polymerase sensitive for a mismatch between nucleotides of the primer and the RNA or cDNA, and the polymerase does not extend a mismatched exonuclease resistant primer, wherein it is preferred that the primers are resistant to exonucleases, especially pre-

ferred by a phosphorothioate modification or a locked nucleic acid.

13. Method according to any one of claims 1 to 12, whereby cDNA with a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, is generated during reverse transcription, dependent on a 5' cap of the template RNA or mRNA, and preferably wherein the cap dependent 3' sequence of oligo-dC is used to isolate fully transcribed cDNA by using a second primer during the amplification which is specific for the oligo-dC sequence.

14. Method according to any one of claims 1 to 13, wherein the tailing of the 3' end of the cDNA is performed with a polynucleotide sequence, which is characterized by the absence of dC.

15. Method according to any one of claims 1 to 14, wherein the DNA polymerase used for the amplification reaction is Taq DNA polymerase, Tfl DNA polymerase, aTaq DNA polymerase, Sequenase or Klentaq and/or an enzyme with proof reading activity, preferably selected from the PFU, Ultma, Vent, Deep Vent, PWO and Tli polymerases and E.coli. exonuclease III.

16. Method according to any one of claims 1 to 15, wherein the RNA in the sample or the cDNA is preamplified, preferably by PCR, or wherein RNA samples are normalized or subtracted.

17. Method according to any one of claims 1 to 16, wherein a primer used for the amplification reaction is characterized by the general formula $dP_xdG_yHN_z$, wherein $P_x$ represents nucleotides complementary to the sequence added to the 3' end of the cDNA, dG being deoxyguanylate and y an integer, preferably between 2 and 5, H being either dT or dA or dC, N being a sequence of nucleotides with a length z wherein its member nucleotides are independently selected from dA, dC, dG and dT, and z being an integer between 0 and 10, preferably between 0 and 5.

18. Method according to any one of claims 1 to 17, wherein the amplification is performed by PCR and the PCR primers contain anchoring sequences, which allow proper placement at the end of the 3' or 5' oligonucleotide end stretch of the cDNA.

19. Method according to any one of claims 1 to 18, wherein an amplification primer or a reverse transcription primer contains a promoter for RNA polymerases, preferably wherein the promoter enhances transcription by the T7 or T3 or SP6 RNA polymerases.

20. Method according to any one of claims 1 to 19, wherein the amplification primers and/or the amplification products are labeled by a reporter group, preferably fluorescence markers, chemiluminescence markers or radioisotopes, preferably wherein these reporter groups are used to detect the DNA.

21. Method according to any one of claims 1 to 20, wherein arbitrary nucleotides, preferably one to three nucleotides, are added to the 3' end of the RNA or mRNA posttranscriptionally but before the addition of a (e.g. poly A) tail, and the primers for reverse transcription and/or PCR contain a sequence with a wobble corresponding to the arbitrary sequence.

22. Method according to any one of the claims 1 to 21, wherein amplification products are generated that are differentially enlarged by defined numbers of nucleotides at their 5' and/or 3' end as compared to the RNA, by using different primers which have different lengths and recognize different 5' or 3' endsequences of the RNA or the full length cDNA.

23. Method according to any one of claims 1 to 22, wherein RNA or mRNA samples are used, which do not have a poly A tail or a tail sequence, and preferably a polynucleotide tail, especially preferred a poly A tail, is added synthetically prior to reverse transcription.

24. Method according to any one of claims 1 to 23, wherein the detection of differences between two cDNA pools, preferably corresponding pools between two samples, is done during or at the end of the amplification step, preferably by real time PCR.

25. Method according to any one of claims 1 to 24, wherein a 5' primer or a 3' primer or a 5'-3' primer pair, is spotted and/or bound onto a solid support, preferably the spot or the solid



IPR2025-01353; IPR2025-01355
CSHL EX2002

support is located in a defined 2-dimensional or 3-dimensional area and preferably such an area contains information to decode the identity of the primer or primer pair located in that area.

26. Method according to any one of claims 1 to 25, wherein the steps of reverse transcription and amplification are performed in one step.

27. Method according to any one of claims 1 to 26, wherein the polynucleotide tail added by the terminal transferase consists of ribonucleotides, preferably 2 to 6 ribonucleotides, preferably adenosine or thymidine, after a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, which is generated during reverse transcription, dependent on a 5' cap of a template RNA or mRNA, and the second amplification primer (5' primer) comprises complementary nucleotides to the ribonucleotides and oligo-dCs and 1 to 10 nucleotides upstream of the olio-dC sequence.

28. Method according to any one of claims 1 to 27, wherein a tail is added to the cDNA by the template switching ability of the reverse transcriptase, preferably in a cap dependent manner, and the second amplification primer (5' primer) comprises complementary nucleotides to the so added tail and oligo-dCs and 1 to 10 nucleotides upstream of the oligo-dC sequence.

29. Method for measuring the expression level of messenger RNA characterized by employing a method according to any one of claims 1 to 28, including the separation and/or identification of amplification results.

30. Kit for a method according to any one of claims 1 to 29 comprising a DNA polymerase, preferably Taq polymerase, a reverse transcriptase or mixtures of different polymerases, cofactors or salts of metal ions, preferably $Mg^{2+}$ and $Mn^{2+}$, requested by a polymerase, primers, preferably being selective for the 5' and/or 3' end of a RNA or cDNA, even more preferred being selective for the first, second, third, fourth and/or fifth terminal nucleotide of the RNA or cDNA, and optionally a terminal transferase and additional buffer substances.

31. A solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a 5′ primer and a 3′ primer or a pair of primers, wherein the primers are defined as in claim 1 point c), claim 2, claim 11, claim 12 or claim 17, in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or different primer pairs.

32. A solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a 5′ primer or a 3′ primer comprising the general formula as defined in claim 17 or a complement thereof, preferably selected from a primer comprising one of the following sequences: AGA GAT TTT TTT TTT TTT TT GA, AGA GAT TTT TTT TTT TTT T GG, AGA GAT TTT TTT TTT TTT T GC, AGA GAT TTT TTT TTT TTT TT GT, AGA GAT TTT TTT TTT TTT TT CA, AGA GAT TTT TTT TTT TTT TT AG, AGA GAT TTT TTT TTT TTT TTT AA, AAA AAA AAA AAA AAA GGG ATA, AAA AAA AAA AAA AAA GGG ACA, AAA AAA AAA AAA AAA GGG AGA, AAA AAA AAA AAA AAA GGG AAA, AAA AAA AAA AAA AAA GGG GTA, AAA AAA AAA AAA AAA GGG CTA, AAA AAA AAA AAA AAA GGG TTA, in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or different primer pairs.

33. Method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of:

a) obtaining a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining a sample of full length cDNA,

b) tailing the 3′ end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides, upstream of

the 3' polynucleotide tail of the cDNA, preferably also comprising unspecific wobble nucleotides, especially preferred in the upstream portion within the 1 to 100 nucleotide range - wherein at least 1 nucleotide is specific for a nucleotide of the cDNA; the method further preferably defined as in any one of claims 1 to 29.

34. Method according to claim 1 or 33, characterized in that a primer, which is specific for a 3' tail of the RNA (optionally an added 3' tail) or mRNA (preferably the endogenous 3' poly A tail) and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides, upstream of the 3' tail is used for the reverse transcription and/or a first amplification primer is specific for a 5' tail of the cDNA, which is complementary to the 3' tail of the RNA or mRNA, and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides, downstream of the 5' tail.



IPR2025-01353; IPR2025-01355
CSHL EX2002

SEQUENCE LISTING

<110> Seitz, Alexander

<120> Polynucleotide Amplifikation

<130> r48881

<150> AT A 1923/2005
<151> 2005-11-29

<160> 34

<170> PatentIn version 3.3

<210> 1
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<220>
<221> misc_feature
<222> (7)..(7)
<223> n is a, c, g, or t

<400> 1
acacacn 7

<210> 2
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 2
acacaca 7

<210> 3
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 3
acacacc 7

<210> 4
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 4
acacact 7

<210> 5



IPR2025-01353; IPR2025-01355
CSHL EX2002

<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 5
acacacg 7

<210> 6
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 6
gatatccgcc ctgtgct 17

<210> 7
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 7
ctgtgggcaa ggtctca 17

<210> 8
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 8
tgtccctttt ggtccca 17

<210> 9
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 9
ttgtggaagg gctcgca 17

<210> 10
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 10
cgatgccccc attcaca 17

<210> 11
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 11
caccaccatg gatctga 17

<210> 12
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 12
caacgggaag cctctaa 17

<210> 13
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 13
acgacccctt catctta 17

<210> 14
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 14
gatatgatgc gcaaagg 17

<210> 15
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 15
gacatggcct aagcaag 17

<210> 16
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 16


IPR2025-01353; IPR2025-01355
CSHL EX2002

gagatgagca cggaagc 17

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

<220>
<223> RNA-terminus

<400> 17
gaaatccgct ttgattt 17

<210> 18
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 18
ggtatcagct ccccatc 17

<210> 19
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 19
gctatgggcc tgggtca 17

<210> 20
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 20
gttatgggcg gagtgga 17

<210> 21
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 21
agagattttt ttttttttt ga 22

<210> 22
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer



IPR2025-01353; IPR2025-01355
CSHL EX2002

<400> 22
agagattttt ttttttttg g 21

<210> 23
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 23
agagattttt ttttttttg c 21

<210> 24
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 24
agagattttt tttttttttt gt 22

<210> 25
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 25
agagattttt tttttttttt ca 22

<210> 26
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 26
agagattttt tttttttttt ag 22

<210> 27
<211> 23
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 27
agagattttt tttttttttt taa 23

<210> 28
<211> 21
<212> DNA
<213> Artificial



IPR2025-01353; IPR2025-01355
CSHL EX2002

<220>
<223> primer

<400> 28
aaaaaaaaaa aaaaagggat a 21

<210> 29
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 29
aaaaaaaaaa aaaaagggac a 21

<210> 30
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 30
aaaaaaaaaa aaaaagggag a 21

<210> 31
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 31
aaaaaaaaaa aaaaagggaa a 21

<210> 32
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 32
aaaaaaaaaa aaaaaggggt a 21

<210> 33
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 33
aaaaaaaaaa aaaaagggct a 21

<210> 34
<211> 21
<212> DNA



IPR2025-01353; IPR2025-01355
CSHL EX2002

<213> Artificial

<220>
<223> primer

<400> 34
aaaaaaaaaa aaaaagggtt a 21



IPR2025-01353; IPR2025-01355
CSHL EX2002

# INTERNATIONAL SEARCH REPORT

International application No
PCT/AT2006/000494

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

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, BIOSIS, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 03/035841 A (SENTION INC [US]; SLEPNEV VLADMIR I [US]) 1 May 2003 (2003-05-01) Composition of claims 60-73 | 30 |
| A | US 5 962 272 A1 (CHENCHIK ALEX [US] ET AL) 5 October 1999 (1999-10-05) cited in the application the whole document | |
| A | DE 199 20 611 A1 (ROCHE DIAGNOSTICS GMBH [DE]) 9 November 2000 (2000-11-09) | |
| A | WO 02/065093 A (BAYLOR COLLEGE MEDICINE [US]; RES FOUND MENTAL HYGIENE [US]; GINSBERG) 22 August 2002 (2002-08-22) | |

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document 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 |
|---|---|
| 19 March 2007 | 03/04/2007 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL – 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | Cornelis, Karen |

Form PCT/ISA/210 (second sheet) (April 2005)



IPR2025-01353, IPR2025-01355
CSHL EX2002

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | EP 1 369 480 A (BIONEER CORP [KR]) 10 December 2003 (2003-12-10) figure 5 | |
| X | EP 1 186 673 A (AGILENT TECHNOLOGIES INC [US]) 13 March 2002 (2002-03-13) table 8 | 30-32 |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/AT2006/000494

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 03035841 A | 01-05-2003 | CA 2464691 A1 | 01-05-2003 |
| | | EP 1490511 A2 | 29-12-2004 |
| | | JP 2005507661 T | 24-03-2005 |
| | | US 2003077611 A1 | 24-04-2003 |
| | | US 2006105380 A1 | 18-05-2006 |
| US 5962272 A1 | | NONE | |
| DE 19920611 A1 | 09-11-2000 | EP 1072686 A2 | 31-01-2001 |
| | | JP 2000325079 A | 28-11-2000 |
| | | US 2003203389 A1 | 30-10-2003 |
| | | US 6558927 B1 | 06-05-2003 |
| WO 02065093 A | 22-08-2002 | CA 2437737 A1 | 22-08-2002 |
| | | EP 1366196 A2 | 03-12-2003 |
| | | JP 2004527236 T | 09-09-2004 |
| EP 1369480 A | 10-12-2003 | NONE | |
| EP 1186673 A | 13-03-2002 | US 2003065449 A1 | 03-04-2003 |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 18851836 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Robert Bolcome |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 24-APR-2014 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 15:42:39 |
| **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 | | 140424_5981_81503_A_Supplemental_IDS_REB.pdf | 1335754<br>fafde4d17531f3686dffb4496faaf92cc81f4fbe | yes | 4 |



IPR2025-01353; IPR2025-01355
CSHL EX2002

| | Multipart Description/PDF files in .zip description | | |
|---|---|---|---|
| | **Document Description** | **Start** | **End** |
| | Transmittal Letter | 1 | 2 |
| | Information Disclosure Statement (IDS) Form (SB08) | 3 | 4 |

| | | | | | |
|---|---|---|---|---|---|
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Other Reference-Patent/App/Search documents | 140424_5981_81503_A_Exhibit_2_REB.pdf | 2024213<br>010997bac0ff8dafd021533e23229744f956e34 | no | 12 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Non Patent Literature | 140424_5981_81503_A_Exhibit_3_REB.pdf | 1440727<br>3a471817081e4841ebdae83128787a629a9b49d4 | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 4 | Foreign Reference | 140424_5981_81503_A_Exhibit_1_REB.pdf | 3708666<br>6fda240c75cfb018f10ea96ce1de10b5657eed17 | no | 53 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 8509360 | | |

**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.**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
April 24, 2014

**BY EFS**
Commissioner for Patents
Alexandria, VA

**SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT**

In accordance with their duty of disclosure under 37 C.F.R. § 1.56, applicants would like to direct the Examiner's attention to the following references, which are listed on Form 1449/PTO (Substitute) (**Exhibit A**).

This Information Disclosure Statement is being submitted pursuant to 37 C.F.R. § 1.97(b)(3) before the mailing of a first Office Action on the merits. Thus, no fee is required and this Supplemental Information Disclosure Statement should be entered and considered.

Reference items **1-3** are U.S. Patent Application Publications. Pursuant to 37 C.F.R. § 1.98(a)(2), copies of these reference items are not being submitted.

Copies of reference items **4-6** are attached hereto as **Exhibits 1-3**, respectfully.

1. U.S. Patent Application Publication No. 2004-110153, published June 10, 2004, Dong et al.;
2. U.S. Patent Application Publication No. 2007-172873, published July 26, 2007, Brenner et al.;
3. U.S. Patent Application Publication No. 2010-178655, published July 15, 2010, Hamilton et al.;

4. PCT International Patent Application Publication No. WO 2007/062445, published June 7, 2007, Alexander Seitz (**Exhibit 1**);
5. February 13, 2014 Extended European Search Report, issued in connection with European Patent Application No. 11835243.4 (**Exhibit 2**); and
6. Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 (**Exhibit 3**).

Applicants request that the Examiner review these references and make them of record in the subject application.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicant's undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this response. However, if any fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39.992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

4-24-14

Robert E. Bolcome III Date
Reg. No. 69,032

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 04/30/2014

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER | |
|---|---|
| LU, FRANK WEI MIN | |
| **ART UNIT** | **PAPER NUMBER** |
| 1634 | |
| **MAIL DATE** | **DELIVERY MODE** |
| 04/30/2014 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. 13/278,333 | Applicant(s) HICKS ET AL. | |
|---|---|---|---|
| | Examiner FRANK LU | Art Unit 1634 | AIA (First Inventor to File) Status No |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *2/20/2014*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims***

5)☒ Claim(s) *1-7,9,11-13,24,26-28,31,41,44,51 and 53* is/are pending in the application.

5a) Of the above claim(s) *11* is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *1-7,9,12,13,24,26-28,31,41,44,51 and 53* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

**Application Papers**

10)☒ The specification is objected to by the Examiner.

11)☒ The drawing(s) filed on *10/21/2011* is/are: a)☒ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

**Certified copies:**

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☐ Notice of References Cited (PTO-892)

2) ☒ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date *See Continuation Sheet*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

Continuation of Attachment(s) 2). Information Disclosure Statement(s) (PTO/SB/08), Paper No(s)/Mail Date : 9/13/2012, 7/24/2013, and 2/20/2014.



IPR2025-01353; IPR2025-01355
CSHL EX2002

## DETAILED ACTION

### *Election/Restrictions*

1. Applicant's election of species (2) (tagging the segments to generate tagged nucleic acid molecules by subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands, see claims 12 and 13) in the reply filed on February 20, 2014 is acknowledged. Because applicant did not distinctly and specifically point out the supposed errors in the restriction requirement, the election has been treated as an election without traverse (MPEP § 818.03(a)). Claims 1-7, 9, 12, 13, 24, 26-28, 31, 41, 44, 51, and 53 will be examined.

### *Drawings*

2. Some letters in Figures 7, 9, and 10 are difficult to be recognized. Applicant is required to furnish drawings under 37 CFR 1.81(c). No new matter may be introduced in the required drawings. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d).

### *Specification*

3. The disclosure is objected to because of the following informalities: (1) there are Figures 5A and 5B. However, BRIEF DESCRIPTION OF THE FIGURES only describes Figure 5; and (2) there are some nucleotide sequences having more than 10 nucleotides in Figures 8 and 9.

However, there are no SEQ ID NOs for these nucleotide sequences in either Figures 8 and 9 or BRIEF DESCRIPTION OF THE FIGURES related to Figures 8 and 9.

Appropriate correction is required.

***Claim Objections***

4. Claim 1 is objected to because of the following informalities: (1) "each tagged nucleic acid molecule" in step b) or e) should be "each of the tagged nucleic acid molecules"; and (2) "one segment of the genomic material" in step a) or e) should be "one segment of the segments of the genomic material".

5. Claim 9 is objected to because of the following informality: "each tagged nucleic acid molecule" in step b) should be "each of the tagged nucleic acid molecules".

6. Claim 12 is objected to because of the following informality: "tagged nucleic acid molecules" in line 2 and last line should be "the tagged nucleic acid molecules".

7. Claim 13 or 44 is objected to because of the following informality: "a polynucleotide tail" should be "the polynucleotide tail".

8. Claim 24 or 27 is objected to because of the following informality: "prior to sequencing" should be "prior to said sequencing the product of step (c)".

9. Claim 28 is objected to because of the following informality: "sample tag" should be "the sample tag".

10. Claim 31 is objected to because of the following informality: "tagged nucleic acid molecules" should be "the tagged nucleic acid molecules".

Appropriate correction is required.

***Claim Rejections - 35 USC § 112***

11. The following is a quotation of the first paragraph of 35 U.S.C. 112(a):

> (a) IN GENERAL.—The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor or joint inventor of carrying out the invention.

The following is a quotation of the first paragraph of pre-AIA 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor of carrying out his invention.

12. Scope of Enablement

Claims 1-7, 9, 12, 13, 24, 26-28, and 31 are rejected under 35 U.S.C. 112(a) or 35 U.S.C. 112 (pre-AIA), first paragraph, because the specification, while being enabling for generating tagged nucleic acid molecules, does not reasonably provide enablement for performing the methods recited in claims 1-7, 12, 13, 24, 26-28, and 31. The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to use the invention commensurate in scope with these claims.

Factors to be considered in determining whether a disclosure meets the enablement requirement of 35 USC 112, first paragraph, have been described by the court in *In re Wands*, 8 USPQ2d 1400 (CA FC 1988). *Wands* states at page 1404,

"Factors to be considered in determining whether a disclosure would require undue experimentation have been summarized by the board in Ex parte Forman. They include (1) the quantity of experimentation necessary, (2) the amount of direction or guidance presented, (3) the presence or absence of working examples, (4) the nature of the invention, (5) the state of the prior art, (6) the relative skill of those in the art, (7) the predictability or unpredictability of the art, and (8) the breadth of the claims."

The nature of the invention

The claims are drawn to a method for obtaining from genomic material genomic copy number information unaffected by amplification distortion and a method for obtaining from mRNA transcriptions mRNA copy number information unaffected by amplification distortion. The invention is a class of invention which the CAFC has characterized as "the unpredictable arts such as chemistry and biology." *Mycogen Plant Sci., Inc. v. Monsanto Co.,* 243 F.3d 1316, 1330 (Fed. Cir. 2001).

The Breadth of The Claims

Claims 1-7, 12, 13, 24, 26-28, and 31 encompass a method for obtaining from genomic material genomic copy number information unaffected by amplification distortion comprising a) obtaining segments of the genomic material; b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the genomic material from step (a) and a tag; c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR); d) generating tag associated sequence reads by sequencing the product of step (c); e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to a segment of the segments of the genomic material to a location on the genome; and f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome. Claim 9 encompasses a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion comprising a)

generating tagged nucleic acid molecules, comprising: i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands; ii) adding a polynucleotide tail to the first order derivative strands; and iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands, wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules; b) subjecting the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification; c) generating tag associated sequence reads by sequencing the product of step (b); d) assigning each of the tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each of the tag associated sequence reads corresponding to a mRNA transcript to a location on the cDNA library; and e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the CDNA library.

Working Examples

The specification provides working examples (see pages 56-62) for Methylation sequencing, Comparative genomic hybridization (CGH) by high throughput sequencing, Single cell microarray and sequencing, Sequencing across chromosome breakpoints, and Copy number

profiling using varietal tags on restriction enzyme fragments. The specification provides no working example or the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, and 31.

The Amount of Direction or Guidance Provided and The State of The Prior Art

Although the specification provides working examples (see pages 56-62) for Methylation sequencing, Comparative genomic hybridization (CGH) by high throughput sequencing, Single cell microarray and sequencing, Sequencing across chromosome breakpoints, and Copy number profiling using varietal tags on restriction enzyme fragments, the specification does not provide a guidance to show that the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, and 31 can be performed. Furthermore, there is no experimental condition and/or experimental data in the specification to support the claimed invention. During the process of the prior art search, the examiner has not found any prior art which is related to the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, and 31.

Level of Skill in The Art, The Unpredictability of The Art, and The Quantity of Experimentation Necessary

While the relative skill in the art is very high (the Ph.D. degree with laboratory experience), there is no predictability whether the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, and 31 can be performed.

First, since step c) of claim 1 does not require that the product of step (c) is an amplification product of a specific gene amplified using primers only specific for the specific gene, if segments of the genomic material are a genomic DNA library and the product of step (c)

is an amplification product of the genomic library amplified using a universal primer and one of the substantially unique tags, the product of step (c) comprises a mixture of genomic DNAs, it is unclear how to the product of step (c) such as a mixture of genomic DNAs can be sequenced so that tag associated sequence reads cannot be generated and each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material cannot be assigned by mapping the subsequence of each tag associated sequence read corresponding to a segment of the segments of the genomic material to a location on the genome as recited in steps d) and e) of claim 1. Furthermore, since step d) of claim 1 does not require completely sequencing the product of step c), if step d) of claim 1 only require sequencing the tag of the product of step c), tag associated sequence reads only comprise sequence of the tag of the product of step c). Thus, it is unclear how each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material can be assigned by mapping the subsequence of each tag associated sequence read corresponding to a segment of the segments of the genomic material to a location on the genome so that the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome cannot be counted as recited in steps e) and f) of claim 1. In addition, since claim 1 does not indicate how to differentiate the amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules, it is unclear how to obtain genomic copy number information unaffected by amplification distortion by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome.

Second, in view of step a) of claim 9, tagged nucleic acid molecules are a cDNA library. Since step c) of claim 9 does not require that the product of step (c) is an amplification product of a specific gene amplified using primers only specific for the specific gene, if the product of step (c) is an amplification product of the cDNA library amplified using a universal primer and one of the substantially unique primers, the product of step (c) comprises a mixture of cDNAs, it is unclear how to the product of step (c) such as a mixture of cDNAs can be sequenced so that tag associated sequence reads cannot be generated and each of the tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts cannot be assigned by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript of the mRNA transcripts to a location on the cDNA library as recited in steps c) and d) of claim 9. Furthermore, since step c) of claim 9 does not require completely sequencing the product of step c), if step c) of claim 9 only require sequencing the tag of the product of step c), tag associated sequence reads only comprise sequence of the tag of the product of step c). Thus, it is unclear how each of the tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts can be assigned by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript of the mRNA transcripts to a location on the cDNA library so that the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library cannot be counted as recited in steps d) and e) of claim 9. In addition, since claim 9 does not indicate how to differentiate the amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules, it is unclear how to obtain mRNA copy number information unaffected by amplification distortion by counting the number

of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library.

Third, in view of the specification and claim 2, it is unclear how genomic copy number of a region of the genome comprising more than one location on the genome can be estimated by assigning the highest count obtained in step (f) for the locations within the region as the copy number of the region.

Fourth, in view of the specification and claim 28, it is unclear how the sample tag is correlated with deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads.

In view of above discussions, the skilled artisan will have no way to predict the experimental results. Accordingly, it is concluded that undue experimentation is required to make the invention as it is claimed. The undue experimentation at least includes to test whether the method recited in claims 1-7, 12, 13, 24, 26-28, and 31 can be performed.

Conclusion

In the instant case, as discussed above, the level of unpredictability in the art is high, the specification provides one with no guidance that leads one to claimed methods. One of skill in the art cannot readily anticipate the effect of a change within the subject matter to which the claimed invention pertains. Thus given the broad claims in an art whose nature is identified as unpredictable, the unpredictability of that art, the large quantity of research required to define these unpredictable variables, the lack of guidance provided in the specification, the absence of any working example related to claimed invention and the no teaching in the prior art balanced only against the high skill level in the art, it is the position of the examiner that it would require

undue experimentation for one of skill in the art to perform the method of the claim as broadly written.

13. The following is a quotation of 35 U.S.C. 112(b):

> (b) CONCLUSION.—The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention.

> The following is a quotation of 35 U.S.C. 112 (pre-AIA), second paragraph:
> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

14. Claims 1-7, 9, 12, 13, 24, 26-28, 31, and 51 are rejected under 35 U.S.C. 112(b) or 35 U.S.C. 112 (pre-AIA), second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which the inventor or a joint inventor, or for pre-AIA the applicant regards as the invention.

15. Claim 1 is rejected as vague and indefinite in view of step b) because it is unclear what is the relationship between a tag and substantially unique tags. Please clarify.

16. Claim 1 is rejected as vague and indefinite in view of step e) because it is unclear what is the relationship between tag associated sequence reads in step d) and each tag associated sequence read in step e). Please clarify.

17. Claim 1 is rejected as vague and indefinite in view of step f) because it is unclear what is the relationship between a different tag in step f) and substantially unique tags in step b). Please clarify.

18. Claim 7 is rejected as vague and indefinite because it is unclear what kind of sum means the sum in line 3. Please clarify.

19. Claim 9 is rejected as vague and indefinite in view of step d) because it is unclear what is the relationship between tag associated sequence reads in step c) and each tag associated sequence read in step d). Please clarify.

20. Claim 9 is rejected as vague and indefinite in view of step e). Since step a) does not require that the tagged nucleic acid molecules have different tag, it is unclear where a different tag in step e) comes from. Please clarify.

21. Claim 12 is rejected as vague and indefinite in view of step a) because it is unclear what is the relationship between substantially unique tags in step b) of claim 1 and a polynucleotide tail in step a) of claim 12. Please clarify.

22. Claim 26 recites the limitation "the tag sequences" in the claim. There is insufficient antecedent basis for this limitation in the claim because there is no phrase "tag sequences" in claim 1. Please clarify.

23. Claim 51 recites the limitation "the tag sequences" in the claim. There is insufficient antecedent basis for this limitation in the claim because there is no phrase "tag sequences" in claim 41. Please clarify.

***Claim Rejections - 35 USC § 101***

24. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

25. Claims 41, 44, 51, and 53 are rejected under 35 U.S.C. 101 because the claimed invention is not directed to patent eligible subject matter. Based upon an analysis with respect to the claim as a whole, claims 41, 44, 51, and 53 do not recite something significantly different than a

judicial exception. The rationale for this determination is explained below: a composition of matter derived from genomic material comprising tagged nucleic acid molecules recited in claims 41 and 53 is a natural product that is not markedly different in structure from natural occurring genomic DNAs and claims 44 and 51 are not structural limitations of the composition of matter derived from genomic material comprising tagged nucleic acid molecules.

***Claim Rejections - 35 USC § 102***

26. The following is a quotation of the appropriate paragraphs of pre-AIA 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

27. Claims 41, 44, 51, and 53 are rejected under pre-AIA 35 U.S.C. 102 (b) as being anticipated by Brenner *et al.,* (US 2007/0172873 A1, published on July 26, 2007).

Regarding claim 41, Brenner *et al.,* teach a composition of matter derived from genomic material comprising tagged nucleic acid molecules (ie., $f_4$ fragments with ligated tag adaptors) (see page 5, paragraph [0038] and Fig. 1G). Although said tagged nucleic acid molecules are not produced using the method of Brenner *et al.,* since claim 41 is product-by-process claim and there is no structural difference between the tagged nucleic acid molecules in the claim and the tagged nucleic acid molecules taught by Brenner *et al.,* claim 41 is anticipated by Brenner *et al.*. Note that it is well established that even though product-by- process claims are limited by and defined by the process, the determination of the patentability of the product is based on the product itself. The patentability of a product does not depend on its method of production. If the

product in the product-by-process claim is the same as or obvious from a product of the prior art, the claim is unpatentable even though the prior product was made by a different process. *In re Thorpe*, 227 USPQ 964, 966 (Fed. Cir. 1985).

Regarding claims 44 and 51, since claims 44 and 51 are not structural limitations of the composition of matter derived from genomic material comprising tagged nucleic acid molecules, claims 44 and 51 are anticipated by Brenner *et al.*.

Regarding claim 53, Brenner *et al.,* teach that the tagged nucleic acid molecules are generated from a single cell (see page 10, paragraph [0064]).

Therefore, Brenner *et al.,* teach all imitations recited in claims 41, 44, 51, and 53.

***Conclusion***

28. No claim is allowed.

29. Papers related to this application may be submitted to Group 1600 by facsimile transmission. Papers should be faxed to Group 1600 via the PTO Fax Center. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993)(See 37 CAR § 1.6(d)). The CM Fax Center number is (571)273-8300.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Frank Lu, Ph.D., whose telephone number is (571)272-0746. The examiner can normally be reached on Monday-Friday from 9 A.M. to 5 P.M.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dave Nguyen, can be reached on (571)272-0731.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Frank W Lu /
Primary Examiner, Art Unit 1634
April 18, 2014

| Search Notes | Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|---|
| | 13278333 | HICKS ET AL. |
| | Examiner | Art Unit |
| | FRANK LU | 1634 |

| CPC- SEARCHED | | |
|---|---|---|
| Symbol | Date | Examiner |
| C12Q1/6813, 6816, 682, 6846, 6851,6853, 6855, 686, 6862, 6865 | 4/18/2014 | /FL/ |

| CPC COMBINATION SETS - SEARCHED | | |
|---|---|---|
| Symbol | Date | Examiner |
| | | |

| US CLASSIFICATION SEARCHED | | | |
|---|---|---|---|
| Class | Subclass | Date | Examiner |
| 435 | 6.1, 6.11, 6.12, 91.1, 91.2, 91.51 | 4/18/2014 | /FL/ |
| 436 | 94, 501 | 4/18/2014 | /FL/ |
| 536 | 23.1, 24.3, 24.33, 25.3 | 4/18/2014 | /FL/ |

| SEARCH NOTES | | |
|---|---|---|
| Search Notes | Date | Examiner |
| STN and EAST (see attached file) | 4/18/2014 | /FL/ |
| Inventor searches | 4/18/2014 | /FL/ |

| INTERFERENCE SEARCH | | | |
|---|---|---|---|
| US Class/ CPC Symbol | US Subclass / CPC Group | Date | Examiner |
| | | | |


IPR2025-01353; IPR2025-01355
CSHL EX2002

**Form PTO-1449 (Substitute)** **U.S. Department of Commerce Patent and Trademark Office**

**INFORMATION DISCLOSURE STATEMENT**
**(Use several sheets if necessary)**

| | |
|---|---|
| **Application Number** | 13/278,333 |
| **Filing Date** | October 21, 2011 |
| **First Named Inventor** | James Hicks |
| **Art Unit** | 1634 |
| **Examiner Name** | Frank W. M. Lu |
| **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| /F.L./ | 1 | 6,569,617 | 05-27-2013 | Wigler et al. |
| | 2 | 7,531,307 | 05-12-2009 | Wigler et al. |
| | 3 | 8,273,871 | 09-25-2012 | Hannon et al. |
| | 4 | 8,554,488 | 10-08-2013 | Wigler et al. |
| | 5 | 2005-0032095 | 02-10-2005 | Wigler et al. |
| | 6 | 2005-0266444 | 12-01-2005 | Wigler et al. |
| | 7 | 2007-0207481 | 09-06-2007 | Wigler et al. |
| | 8 | 2010-0227768 | 09-10-2010 | Wigler et al. |
| | 9 | 2012-0149593 | 06-14-2012 | Hicks et al. |
| | 10 | 2012-0328607 | 12-27-2012 | Hicks et al. |
| | 11 | 2013-0179999 | 07-11-2013 | Hannon et al. |
| | 12 | 2014-0025307 | 01-23-2014 | Wigler et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |

| | | | |
|---|---|---|---|
| **EXAMINER SIGNATURE** | /Frank Lu/ | **DATE CONSIDERED** | 04/19/2014 |

***EXAMINER:** Initial if citation 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 of 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | **Application Number** | 13/278,333 |
| | **Filing Date** | October 21, 2011 |
| | **First Named Inventor** | James Hicks |
| | **Art Unit** | |
| | **Examiner Name** | |
| | **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| /F.L./ | 1 | 7,700,324 | 04-20-2010 | Issa et al. |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| /F.L./ | 3 | WO 2012/054873 | 04-26-2012 | Hicks et al. | |
| | | | | | |
| | | | | | |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 04/19/2014 |
|---|---|---|---|

*EXAMINER: Initial if citation 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 of 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | **Application Number** | 13/278,333 |
| | **Filing Date** | October 21, 2011 |
| | **First Named Inventor** | James Hicks |
| | **Art Unit** | |
| | **Examiner Name** | |
| | **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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] |
|---|---|---|---|
| /F.L./ | 2 | International Preliminary Report on Patentability, issued April 23, 2013 in connection with PCT International Application No. PCT/US2011/057350 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 04/19/2014 |
|---|---|---|---|

*EXAMINER: Initial if citation 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.

**Form PTO-1449 (Substitute)** **U.S. Department of Commerce Patent and Trademark Office**

**INFORMATION DISCLOSURE STATEMENT**
**(Use several sheets if necessary)**

| | |
|---|---|
| **Application Number** | 13/278,333 |
| **Filing Date** | October 21, 2011 |
| **First Named Inventor** | James Hicks |
| **Art Unit** | |
| **Examiner Name** | |
| **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| /F.L./ | 1 | 6,172,214 | 01-09-2001 | Brenner et al. |
| | 6 | 2006/0035258 | 02-16-2006 | Tadakamalla et al. |
| | 7 | 2007/0172873 | 07-26-2007 | Brenner et al. |
| | 8 | 2009/0137402 | 05-28-2009 | Wang et al. |
| | 9 | 2010/0203538 | 08-12-2010 | Dube et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |

| | | | |
|---|---|---|---|
| **EXAMINER SIGNATURE** | /Frank Lu/ | **DATE CONSIDERED** | 04/19/2014 |

***EXAMINER:** Initial if citation 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 of 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | **Application Number** | 13/278,333 |
| | **Filing Date** | October 21, 2011 |
| | **First Named Inventor** | James Hicks |
| | **Art Unit** | |
| | **Examiner Name** | |
| | **Attorney Docket No.** | 5981/81503-A/GJG/REB |

**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] |
|---|---|---|---|
| /F.L./ | 2 | Islam et al., "Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq" Genome Research (2001) 21:1160-1167 | |
| | 3 | Kivioja et al., "Counting absolute numbers of molecules using unique molecular identifiers" Nature Methods (2012) 9(1):72-74 and Online Methods (2 pages) | |
| | 4 | Shiroguchi et al., "Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes" PNAS (2012) 109(4):1347-1352 | |
| | 5 | International Search Report, issued in connection with PCT International Application No. PCT/US2011/057350 mailed on 6/21/2012 | |
| | 10 | Castle et al., "DNA copy number, including telomeres and mitochondria, assayed using next-generation sequencing" BMC Genomics (2010) 11(244):1-11 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 04/19/2014 |
|---|---|---|---|

*EXAMINER: Initial if citation 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.

Search Note for Application 13/278,333 (4/18/2013)

Keywords in STN (Biosis, medline, caplus, embase and scisearch) and EAST

1. nucleic acid or DNA or RNA or polynucleotide or oligonucleotide or genomic DNA or cDNA
2. tag$ or adaptor or adapter or linker or label
3. PCR or amplif$ or extens$
4. sequence read$
5. location or position
6. count$ or calculate$
7. number
8. amplification distort$

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
July 30, 2014

**BY EFS**
Commissioner for Patents
Alexandria, VA

**AMENDMENT IN RESPONSE TO APRIL 30, 2014 OFFICE ACTION AND SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT**

This Amendment is submitted in reply to an April 30, 2014 Office Action issued by the United States Patent and Trademark Office in connection with the above-identified application. A response to the April 30, 2014 Office Action is due July 30, 2014. Accordingly, this response is being timely filed.

**Amendments to the Drawings** are on page **2** of this paper.

**Amendments to the Specification** begin on page **3** of this paper.

**Amendments to the Claims** begin on page **4** of this paper.

**Remarks** begin on page **11** of this paper.

A **Supplemental Information Disclosure Statement** begins on page **27** of this paper.

**AMENDMENTS TO THE DRAWINGS**

Applicants attach hereto as **Exhibit A** EIGHT (8) replacement sheets of drawings containing Figures 7-10.

**AMENDMENTS TO THE SPECIFICATION**

Please amend the paragraph that runs from lines 27-29 on page 13 of the subject application as filed, as follows:

--**Figure 5. ROMA CGH of sKBR3 compared with profiles of 3 independent single cell isolates derived from sparse DNA sequencing. (A) Batch DNA. (B, C, D) Single Cell.**--

Please amend the paragraph that runs from lines 3-4 on page 14 of the subject application as filed, as follows:

--**Figure 7(A-C). Flowchart for genomic profiling by representational sequence counting**--

Please amend the paragraph on line 8 of page 14 of the subject application as filed, as follows:

--**Figure 9(A and B). Nick ligation protocol**--

Please amend the paragraph on line 15 of page 14 of the subject application as filed, as follows:

--**Figure 10(A and B). Schema for targeted sequencing of regional markers**--

**AMENDMENTS TO THE CLAIMS**

Pursuant to 37 C.F.R. § 1.121(c), this listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims**

1. (Currently Amended) A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to [[a]]one segment of the segments of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

2. (Currently Amended) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Original) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Original) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Original) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Original) The method of claim 1, further comprising summing the counts obtained in step (f) for locations which comprise a region of the genome, and comparing the sum to a sum obtained from a reference sample for the same region of the genome, thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Currently Amended) A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. (Cancelled)

11. (Previously Presented) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:


IPR2025-01353; IPR2025-01355
CSHL EX2002

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Currently Amended) The method of claim 1, wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands

added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Currently Amended) The method of claim 12, wherein adding [[a]]the polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d)~~prior to sequencing~~.

25. (Cancelled)

26. (Previously Presented) The method of claim 1, wherein the tag sequences further comprise a sample tag.

27. (Currently Amended) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to

PCR amplification or prior to said sequencing of the product of step (c).

28. (Currently Amended) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.

29-30. (Cancelled)

31. (Currently Amended) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (New) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (New) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (New) The method of claim 24, wherein amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (New) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.

**REMARKS**

Claims 1-7, 9, 11-13, 24, 26-28, 31, 41, 44, 51 and 53 are pending in the subject application. By this Amendment, applicants have hereinabove amended claims 1, 2, 9, 12. 13, 24, 27, 28, and 31 and cancelled claims 41, 44, 51 and 53 without prejudice of disclaimer, and added new claims 68-71. Claim 11 has been withdrawn from consideration. Accordingly, claims 1-7, 9, 11-13, 24, 26-28, 31 and 67-71 are pending and under consideration in the subject application.

Support for the amendment to claim 2 may be found, *inter alia*, at [0293] of the subject specification as published as U.S. Patent Application Publication No. 2014/0065609.

Claims 1, 9, 12, 13, 24, 27, 28, and 31 were amended in accordance with the Examiner's suggestions, and/or for clarity.

Support for the new claims 68-71 may be found, *inter alia*, at paragraphs [0172], [0246] and [0339] of the subject specification as published.

**Objections to the Drawings**

The April 30, 3014 Office Action objected to Figures 7, 9 and 10 for allegedly containing letters that were difficult to recognize.

*Applicants' reply*

In response, applicants attach hereto as **Exhibit A** EIGHT (8) replacement sheets of drawings containing Figures 7-10. The letters in Figures 7, 9 and 10 are clear. Additionally, sequences having more than 10 nucleotides are identified by their respective SEQ ID NO in Figures 8 and 9. Additionally, applicants attach as **Exhibit B**, FOUR (4) sheets of drawings containing previous Figures 7-10, marked to show the amendments.

Applicants have also amended the description of Figures 7, 9 and 10 in the subject application to reflect the multiple views of the

amended figures.

The amendments to the drawings and the specification raise no issue of new matter.

Accordingly, applicants respectfully request that the objections to the drawings be withdrawn.

**Objections to the Specification**

The April 30, 2014 Office Action objected to the specification because the BRIEF DESCRIPTION OF THE FIGURES only described Figure 5, but there are Figures 5A and 5B. Additionally, neither the descriptions of Figures 8 and 9 nor the Figures themselves identified the SEQ ID NO for nucleotides in Figures 8 and 9 that had more than 10 nucleotides.

*Applicants' reply*

In response, applicants have amended the description of Figure 5 in the BRIEF DESCRIPTION OF THE FIGURES of the subject application do describe Figures 5A and 5B. As discussed above, Figures 8 and 9 have been amended to identify the SEQ ID NO for oligonucleotides therein.

The amendments to the specification raise no issue of new matter.

Accordingly, applicants respectfully request that the objections to the specification be withdrawn.

**Objections to the Claims**

The April 30, 2014 Office Action objected to claims 1, 9, 12, 13 or 44, 24 or 27, 28, and 31. The Examiner's rationale and suggested claim amendments may be found on page 3 of the April 30, 2014 Office Action.

*Applicants' reply*

In response, without conceding correctness of the Examiner's


IPR2025-01353; IPR2025-01355
CSHL EX2002

position, and solely for the purpose of advancing prosecution, applicants have amended claims 1, 9, 12, 13, 44, 24, 27, 28, and 31 without prejudice or disclaimer. The amendments to these claims are consistent with the Examiner's suggested claim amendments on page 3 of the April 30, 2014 Office Action.

Accordingly, applicants respectfully request that the claim objections be withdrawn.

**Rejection of claims under 35 U.S.C. § 112, First Paragraph (Enablement)**

The April 30, 2014 Office Action rejected claims 1-7, 9, 12, 13, 24, 26-28 and 31 under 35 U.S.C. § 112, first paragraph, because the specification allegedly did not enable any person skilled in the art to which it pertains, or with which it is most clearly connected, to make and use the invention commensurate in scope with the claims. The Examiner's rationale is set forth on pages 4-11 of the April 30, 2014 Office Action.[1]

*Applicants' reply*

In response, applicants respectfully disagree. Additionally, without conceding correctness of the Examiner's position, and solely for the purpose of advancing prosecution, applicants have amended claims 1, 9, 12, 13, 24, 27, 28, and 31 without prejudice or disclaimer. The pending claims comply with 35 U.S.C. § 112, first paragraph (Enablement).

Independent claim 1 relates to a method for "obtaining from genomic material genomic copy number information unaffected by amplification" comprising steps a) through f). Steps a) through d) relate to obtaining, tagging, amplifying, and sequencing segments of the genomic material. Steps e) and f) relate to assigning and

[1] On page 7 of the April 30, 2014 Office Action the Examiner acknowledged that "[d]uring the process of the prior art search, the examiner has not found any prior art which is related to the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, and 31."

counting the segments of the genomic material.

Independent claim 9 relates to a method for "obtaining from mRNA transcripts mRNA copy number information unaffected by amplification" comprising steps a) through c). Steps a) through c) relate to generating, amplifying, and sequencing tagged nucleic acid molecules from mRNA Steps d) and e) relate to assigning and counting the tagged nucleic acid molecules.

The subject specification provides substantial supporting disclosures for each and every step of the claims. *See, e.g.*, paragraphs [0167], [0168], [0308], [0309] and [0331] (describing obtaining segments of genomic material); paragraphs [0150]-[0163], [0241], [0300], [0302]-[0304], [0306], [0307], [0310], [0313], [0314], [0320] and [0331], and Figures 7-10 (describing generating "tagged nucleic acid molecules"); paragraph [0164], [0242], [0314], [0315] and [0331] (describing subjecting the tagged nucleic acid molecules to amplification); paragraphs [0260], [0292], [0311] and [0331] (describing "generating tag associated sequence reads"); paragraphs [0255], [0292]-[0295] and [0331] (describing methods for assigning a tagged nucleic acid molecule to a location on a genome or a cDNA library); and paragraphs [0331]-[0333], [0109]-[0117] (describing methods for counting the number of tagged nucleic acid molecules) of the subject application as published.

Undue experimentation would not be required to perform the claimed method for obtaining genomic or mRNA transcript "copy number information unaffected by amplification distortion", as claimed. The eight factors the Federal Circuit set forth in *In re Wands* show that the subject application enabled a person having ordinary skill in the art ("POSA") to make and use the claimed invention on the effective filing date of the subject application without undue experimentation. *See In re Wands* 858 F.2d 731, 737 (Fed. Cir. 1988). The eight "Wands Factors" are useful to show that "undue" experimentation was not needed; the eight factors are (1) the quantity of experimentation necessary, (2) the amount of direction

and guidance presented, (3) the presence or absence of working examples, (4) the nature of the invention, (5) the state of the prior art, (6) the relative skill of those in the art, (7) the predictability or unpredictability of the art, and (8) the breath of the claims.

Even if, *arguendo*, routine optimization experimentation would have been necessary to perform the claimed invention, such experimentation would not have been undue. Routine experimentation regardless of complexity is not "undue." *See, e.g.*, M.P.E.P. § 2164.01. "'The test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed.'" *In re Wands*, 858 F.2d at 737 (citing *In re Angstadt*, 537 F.2d 489, 502-04 (CCPA 1976)). Additionally, "time and difficulty of experiments are not determinative if they are merely routine." M.P.E.P. § 2164.06.

I The Quantity of experimentation necessary.

On the effective filing date of the subject application, a POSA could have practiced the claimed invention with little or no experimentation. Any experimentation that would have been required would have been routine optimization. The subject application, in light of knowledge in the art, provided all the direction and guidance needed to practice the claimed methods for obtaining cope number information unaffected by amplification distortion, as claimed.

Every step of the claimed invention can be readily performed using disclosures in the subject application, supplemented by knowledge in the art. The result of performing these steps would, without more, provide copy number information, as claimed. No further experimentation would be needed.

II The amount of direction and guidance presented.

The subject application discloses the critical information needed to each and every step of the claims. *See* discussion above.

A POSA was in possession of the skill needed to practice the steps recited in the claims. No other disclosure would have been required to enable a POSA to practice the claimed invention on the effective filing date.

Importantly, the subject application need not disclose subject matter that would have been well-known to a POSA on the effective filing date. M.P.E.P. § 2164.05(a) states:

> "The specification need not disclose what is well-known to those skilled in the art and preferably omits that which is well-known to those skilled and already available to the public. *In re Buchner*, 929 F.2d 660, 661, 18 USPQ2d 1331, 1332 (Fed. Cir. 1991); *Hybritech, Inc.v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384, 231 USPQ 81, 94 (Fed. Cir. 1986), *cert. denied*, 480 U.S. 947 (1987); and *Lindemann Maschinenfabrik GMBH v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1463, 221 USPQ 481, 489 (Fed. Cir. 1984)." (Emphases added).

As discussed below, the subject specification provides explicit teachings that address factors that the April 30, 2014 Office Action alleged to be unpredictable. *See* pages 7-9 of the April 30, 2014 Office Action. Applicants respectfully note that the allegations on pages 7-9 of the April 30, 2014 Office Action were difficult to understand. The discussion below reflects applicants' understanding of the Examiner's allegations.

*II(A) Allegation 1: Applying the claimed steps to mixtures of segments of genomic DNA material or mixtures of cDNAs was not enabled.*

Applicants respectfully disagree with this allegation. The subject application contemplates and describes how to sequence and map tagged nucleic acid molecules derived from mixtures of many segments of genomic material or cDNAs.

Paragraph [0010] of the subject specification states that a method

"that allows for copy number determination free of distortions caused by amplification steps and which allows for accurate and efficient copy number determination of complex samples" using "existing methodologies for high volume, massively parallel sequencing" was needed. The subject specification goes on to explain that it provides such methods, and that steps of the methods provided may use well known and understood high-throughput sequencing and mapping techniques. *See, e.g.*, paragraphs [0285]-[0288], [0321], [0322], [0332] and [0336], and Figure 7. Additionally, paragraph [0332] discloses that applicants successfully obtained "genome-wide" copy number data. *See also*, Figure 11A. Paragraph [0285] shows that applicants successfully sequenced and mapped hundreds of thousands of genomic fragments. *See also*, Figure 7.[2]

Paragraphs [0331] and [0324] of the subject application describe the parallel sequencing of many tagged nucleic acid molecules. For example, paragraph [0324] describes an available sequencing technique which "can provide millions of individual sequence reads randomly distributed across any input DNA" (citation omitted). With respect to mRNA, paragraph [0322] explains that "[s]everal methods have now been published whereby researchers have profiled the RNA content of single cells by sequencing" (citation omitted).

The subject application also describes how to map multiple sequence reads to a genome or cDNA library. *See, e.g.*, paragraph [0255], which states "the term 'mapping' refers to identifying on a genome or cDNA library . . . " (emphasis added)). Paragraph [0282] explains that mapping can readily be achieved by aligning sequence reads to specific locations on a map of the human genome.

Additionally, dependent claims 24 and 68-71 relate to subjecting the tagged nucleic acid molecules to hybrid capture, which allows for the selection of specific tagged nucleic acid molecules of

[2] Methods for reducing sample complexity by representation and bin sequencing are also provided in the subject application. See, e.g., paragraphs [0282], [0283], [0338] and [0339].

interest. *See, e.g.*, paragraphs [0246] and [0339].

*II(B) Allegation 2: Mutant tagged nucleic acid molecules.*

On page 8 of the April 30, 2014 Office Action, the Examiner alleged claim 1 "does not indicate how to differentiate the amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules." On page 9, of the April 30, 2014 Office Action, the Examiner alleged claim 9 "does not indicate how to differentiate the amplification products having original tagged nucleic molecules and the amplification products having mutant tagged nucleic acid molecules."

First, neither claim 1 nor claim 9 recites the term "mutant". Accordingly, applicants are unsure how the Examiner's statements relate to the claims. To the extent the Examiner was referring to mutations introduced during the performance of claim steps, paragraph [0311] explains:

> "Through sequence error during processing, the tag may very occasionally mutate, creating the appearance of a new tagged nucleic acid molecule. But if the set of tags are sufficiently large, the tagged nucleic acid molecules will generally differ at more than one nucleotide, and so tagged nucleic acid molecules that differ by a single nucleotide can (optionally) be ignored when counting."

Accordingly, the subject application provided ample guidance to enable a POSA to account for unwanted mutations which might unexpectedly and rarely occur.

*II(C) Allegation 3: Claim 2 was not enabled by the subject specification.*

On page 10 of the April 30, 2014 Office Action, the Examiner alleged that "in view of the specification and claim 2, it is unclear how genomic copy number of a region of the genome comprising more than one location on the genome can be estimated by assigning the highest count obtained in step (f) for the locations within the region as the copy number of the region."

In response, without conceding correctness of the Examiner's position, and solely for the purpose of advancing prosecution,

applicants have amended claim 2 for clarity. Enabling support for the amendments to claim to may be found in paragraph [0293] of the subject application.

*II(D) Allegation 4: Claim 28 was not enabled by the subject specification.*

On page 10 of the April 30, 2014 Office Action, the Examiner alleged that "in view of the specification and claim 28, it is unclear how the sample tag is correlated with deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads."

In response, applicants respectfully disagree that claim 28 is not enabled. Neither the term "correlated" nor the phrase "correlated with deconvoluting" is recited claims 28. Claim 28 explicitly recites how to deconvolute the tag associated sequence reads; deconvoluting is achieved "by grouping the tag associated sequence reads according to the sample tag".

The sample tag allows tags from different samples within a pool to be grouped so that copy number can be determined in each of the different pooled samples. *See, e.g.*, [0266], [0283] and [0292].

*II(E) Allegation 5: Step d) of claim 1 encompasses generating a sequence read of only the tag of the product of step c).*

On page 8 of the April 30, 2014 Office Action, the Examiner alleged that "step d) of claim 1 does not require completely sequencing the product of step c), if step d) of claim 1 only require sequencing the tag of the product of step c), tag associated sequence reads only comprise sequence of the tag of the product of step c)".

In response, applicants respectfully disagree. Step d) of claim 1 recites "generating tag associated sequence reads by sequencing the product of step (c)". The Examiner's alleged embodiment of sequencing only tags (without also sequencing segments of genetic material) does not fall within the scope of step d). Step d) recites generating "tag associated sequence reads" (emphases added). Step

d) does not recite generating mere "tag sequence reads".

III The presence of working examples.

The subject specification discloses constructive and actual reduction to practice of embodiments of the claimed invention.

Page 6 of the April 30, 2014 Office Action acknowledged that the subject specification provides working examples for copy number profiling using varietal tags on restriction enzyme fragments, but alleges that "[t]he specification provides no working example [f]or the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, 31." These assertions in the April 30, 2014 Office Action are factually inconsistent. Copy number profiling using varietal tags on restriction enzyme fragments falls within the scope of independent claim 1.

The subject application describes how to obtain, and provides experimental examples relating to obtaining, copy number information unaffected by amplification distortion as claimed. However, the claims would be enabled even if absolutely no experimental data was provided; the law does not require applicants to actually reduce the claims to practice. Constructive reduction to practice is sufficient.

Paragraphs [0331]-[0333] of the subject specification provide a working example of obtaining copy number information. Additionally, the subject specification explains that "tagging" "assigning" and "counting" steps may be performed in a variety of ways that would have been well within the skill of a POSA reading the subject application. *See, e.g.*, paragraphs [0150]-[0163], [0241], [0300], [0302]-[0304], [0306], [0307], [0310], [0313], [0314], and [0320] (tagging), [0255], [0292]-[0295] and [0331] (assigning), and [0109]-[0117] (counting) of the subject application.

Additionally, applicants are not required to limit the scope of the claims to the working examples provided in the subject

specification. M.P.E.P. § 2164.02 states:

> "An applicant need **not** have actually reduced the invention to practice prior to filing. In *Gould v. Quigg*, 822 F.2d 1074, 1078, 3 USPQ 2d 1302, 1304 (Fed. Cir. 1987), as of Gould's filing date, no person had built a light amplifier or measured a population inversion in a gas discharge. The Court held that "The mere fact that something has not previously been done clearly is not, in itself, a sufficient basis for rejecting all applications purporting to disclose how to do it." 822 F.2d at 1078, 3 USPQ2d at 1304 (quoting *In re Chilowsky*, 229 F.2d 457, 461, 108 USPQ 321, 325 (CCPA 1956)).
>
> The specification **need not contain an example** if the invention is otherwise disclosed in such manner that one skilled in the art will be able to practice it without an undue amount of experimentation. *In re Borkowski*, 422 F.2d 904, 908, 164 USPQ 642, 645 (CCPA 1970)." (Emphasis added.)

The examples disclosed in the subject application are representative of the subject matter recited in independent claims 1 and 9. The claimed invention is disclosed in such a manner that a POSA at the relevant time could readily have practiced it without an undue amount of experimentation.

IV The nature of the invention.

The claimed invention is within the biological arts.

V The state of the art.

On the effective filing date, the state of the biological arts was such that a POSA could readily have practiced each and every step of the claims. *See, e.g.*, discussion *infra* in Part IX.

VI The relative skill of those in the art.

As acknowledged by the Examiner, "[t]he level of skill in the art is very high." *See* April 30, 2014 Office Action at page 7. Consequently, the POSA was certainly capable of i) generating tagged nucleic acids, ii) amplifying tagged nucleic acids, iii) sequencing, assigning and iv) counting, as claimed. *See* discussion supra in Parts II and III, and *infra* in Part IX.

VII The predictability or unpredictability of the art.

Applicants acknowledge that the biological arts are deemed to be unpredictable. *Mycogen Plant Sci., Inc. v. Monsanto Co.*, 243 F.3d

1316, 1330. However, as discussed above, the state of the biological arts was such that a POSA could readily have performed each and every step of the claims. Thus, while inventing the steps required to arrive at the invention as a whole was unpredictable, once the specific steps were identified, arranged and disclosed by the inventors, applying the steps required the application of predictable techniques to identify the obtain copy number information unaffected by amplification distortion. Nothing unpredictable is required once the claim steps and the proper arrangement thereof are known.

VIII The breadth of the claims.

The pending claims are narrow in that they require a specific combination of steps to be used when obtaining copy number information unaffected by amplification distortion.

A POSA reading the subject application at the relevant time would have been able to readily practice the claimed methods without any undue experimentation using methods described in the specification as supplemented by knowledge in the art.

Even if a considerable amount of routine experimentation were required, the claimed invention would still be enabled. M.P.E.P. § 2164.06 states:

> "'[A]n extended period of experimentation may not be undue if the skilled artisan is given sufficient direction or guidance.' *In re Colianni*, 561 F.2d 220, 224, 195 USPQ 150, 153 (CCPA 1977). ''The test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed.'' *In re Wands*, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988) (citing *In re Angstadt*, 537 F.2d 489, 502-04, 190 USPQ 214, 217-19 (CCPA 1976))."

For the foregoing reasons alone, the claims satisfy the enablement requirement.

IX The preponderance of the evidence of record shows that the claims are enabled.

No evidence of record shows that obtaining copy number information as recited in the claims is not enabled. The disclosure of the subject application must be presumed to be accurate unless Examiner is aware of technical evidence to the contrary. Applicants note that no explicit evidence was provided in the April 30, 2014 Office Action to support Examiner's enablement rejection.[3]

M.P.E.P. § 2164.04 states:

> "In order to make a rejection, the examiner has the initial burden to establish a reasonable basis to question the enablement provided for the claimed invention. *In re Wright*, 999 F.2d 1557, 1562, 27 USPQ2d 1510, 1513 (Fed. Cir. 1993) (examiner must provide a reasonable explanation as to why the scope of protection provided by a claim is not adequately enabled by the disclosure). A specification disclosure which contains a teaching of the manner and process of making and using an invention in terms which correspond in scope to those used in describing and defining the subject matter sought to be patented **must be taken as being in compliance with the enablement requirement** of 35 U.S.C. 112, first paragraph, unless there is a reason to doubt the objective truth of the statements contained therein which must be relied on for enabling support. Assuming that sufficient reason for such doubt exists, a rejection for failure to teach how to make and/or use will be proper on that basis. *In re Marzocchi*, 439 F.2d 220, 224, 169 USPQ 367, 370 (CCPA 1971). As stated by the court, "it is incumbent upon the Patent Office, whenever a rejection on this basis is made, **to explain why it doubts the truth or accuracy of any statement** in a supporting disclosure and to back up assertions of its own with **acceptable evidence or reasoning** which is inconsistent with the contested statement. Otherwise, there would be no need for the applicant to go to the trouble and expense of supporting his presumptively accurate disclosure." 439 F.2d at 224, 169 USPQ at 370." (Emphases added.)

As explained hereinabove, the subject specification teaches that copy number information unaffected by amplification distortion may be obtained using the steps recited in the claims, as supplemented by knowledge in the art on the relevant date. The April 30, 2014 Office Action provides no evidence, other than the Examiner's own

[3] Page 10 of the April 30, 2014 Office Action alleged that undue experimentation was required to practice the claimed invention, and that "[t]he undue experimentation at least includes to test whether the method recited in claims 1-7, 12, 13, 24, 26-28, and 31 can be performed". However, the subject specification discloses that the claimed methods successfully obtain copy number information. The disclosure of the subject application must be presumed accurate.

opinions, to rebut the statements in the subject specification. The dismissal of teachings in the specification in the absence of contradicting evidence is improper. The determination of enablement must be based on the evidence as a whole.

M.P.E.P. § 2164.05 states that in considering applicant's arguments,

> "The examiner must then weigh all the evidence before him or her, including the specification and any new evidence supplied by applicant with the evidence and/or sound scientific reasoning previously presented in the rejection and decide whether the claimed invention is enabled. The examiner should **never** make the determination based on personal opinion. The determination should always be based on the weight of all the evidence." (**Emphasis in original**.)

The evidence in this record consists of applicants' specification and applicants' submissions. There is no evidence of record contradicting applicants' evidence. Therefore, there is no evidence of record that supports an enablement rejection.

Accordingly, applicants respectfully request that this rejection be withdrawn.

**Rejection of claims under 35 U.S.C. § 112, Second Paragraph**

The April 30, 2014 Office Action rejected claims 1-7, 9, 12, 13, 24, 26-28, 31 and 51 under 35 U.S.C. § 112, second paragraph, as allegedly indefinite for failing to particularly point out and distinctly claim the subject matter with applicants regard as the invention. The Examiner's rationale is set forth on pages 11 and 12 of the April 30, 2014 Office Action.

*Applicants' reply*

In response, applicants respectfully disagree. However, without conceding correctness of the Examiner's position, and solely for the purpose of advancing prosecution, applicants have amended claims 1, 9, 12, 13, 24, 27, 28, and 31 and cancelled claim 51 without prejudice or disclaimer. The pending claims comply with 35 U.S.C. § 112, second paragraph.

Accordingly, applicants respectfully request that this rejection be withdrawn.

**Rejection of claims under 35 U.S.C. § 101**

The April 30, 2014 Office Action rejected claims 41, 44, 51 and 53 under 35 U.S.C. § 101, as allegedly not directed to patent eligible subject matter. The Examiner's rationale is set forth on pages 12 and 13 of the April 30, 2014 Office Action.

*Applicants' reply*

In response, applicants respectfully disagree. However, without conceding correctness of the Examiner's position, and solely for the purpose of advancing prosecution, applicants have canceled claims 41, 44, 51 and 53 without prejudice or disclaimer.

Accordingly, this rejection is moot.

**Rejection of claims under 35 U.S.C. § 102**

The April 30, 2014 Office Action rejected claims 41, 44, 51 and 53 under 35 U.S.C. § 102(b) as allegedly anticipated by U.S. Patent Application Publication No. 2007/0172873 ("Brenner"). The Examiner's rationale is set forth on pages 13 and 14 of the April 30, 2014 Office Action.

Applicants' reply

In response, applicants respectfully disagree. However, without conceding correctness of the Examiner's position, and solely for the purpose of advancing prosecution, applicants have canceled claims 41, 44, 51 and 53 without prejudice or disclaimer.

Accordingly, this rejection is moot.

**Concluding Remarks**

Applicants believe the remarks presented hereinabove to be fully responsive to all of the grounds of rejection raised by the Examiner. In view of these remarks, applicants respectfully submit

that all of the claims in the present application are now in order for allowance. Notice to this effect is respectfully requested.

**SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT**

In accordance with their duty of disclosure under 37 C.F.R. § 1.56, applicants would like to direct the Examiner's attention to the following documents which are listed on Form PTO-1449 (**Exhibit C**) and are also listed below.

According to 37 C.F.R. § 1.97(c), a Supplemental Information Disclosure Statement filed after the period specified in 37 C.F.R. § 1.97(b) shall be considered if accompanied by the fee set forth in 37 C.F.R. § 1.17(p) or a statement under 37 C.F.R. § 1.97(e). The fee set forth in 37 C.F.R. § 1.17(p) is NINETY DOLLARS ($90.00) for a small entity, and authorization is hereby given to charge this amount of fee to Deposit Account No. 03-3125.

Reference items **1** and **2** are U.S. Patents. Pursuant to 37 C.F.R. § 1.98(a)(2), copies of these reference items are not being submitted.

1. U.S. Patent No. 8,663,917, issued March 4, 2014, Wigler et al.; and
2. U.S. Patent No. 8,694,263, issued April 8, 2014, Wigler et al.

The Examiner is respectfully requested to make of record each item listed on the substitute Form PTO-1449 by initialing and dating the attached substitute Form PTO-1449, and returning a copy of the initialed and dated form to applicants' undersigned attorney.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicant's undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee, other than $90.00 for an Information Disclosure Statement, is deemed necessary in connection with the filing of this response. However, if any fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39,992
Attorney for Applicants
Cooper & Dunham LLP
30 Rockefeller Plaza
New York, New York 10112
(212) 278-0400

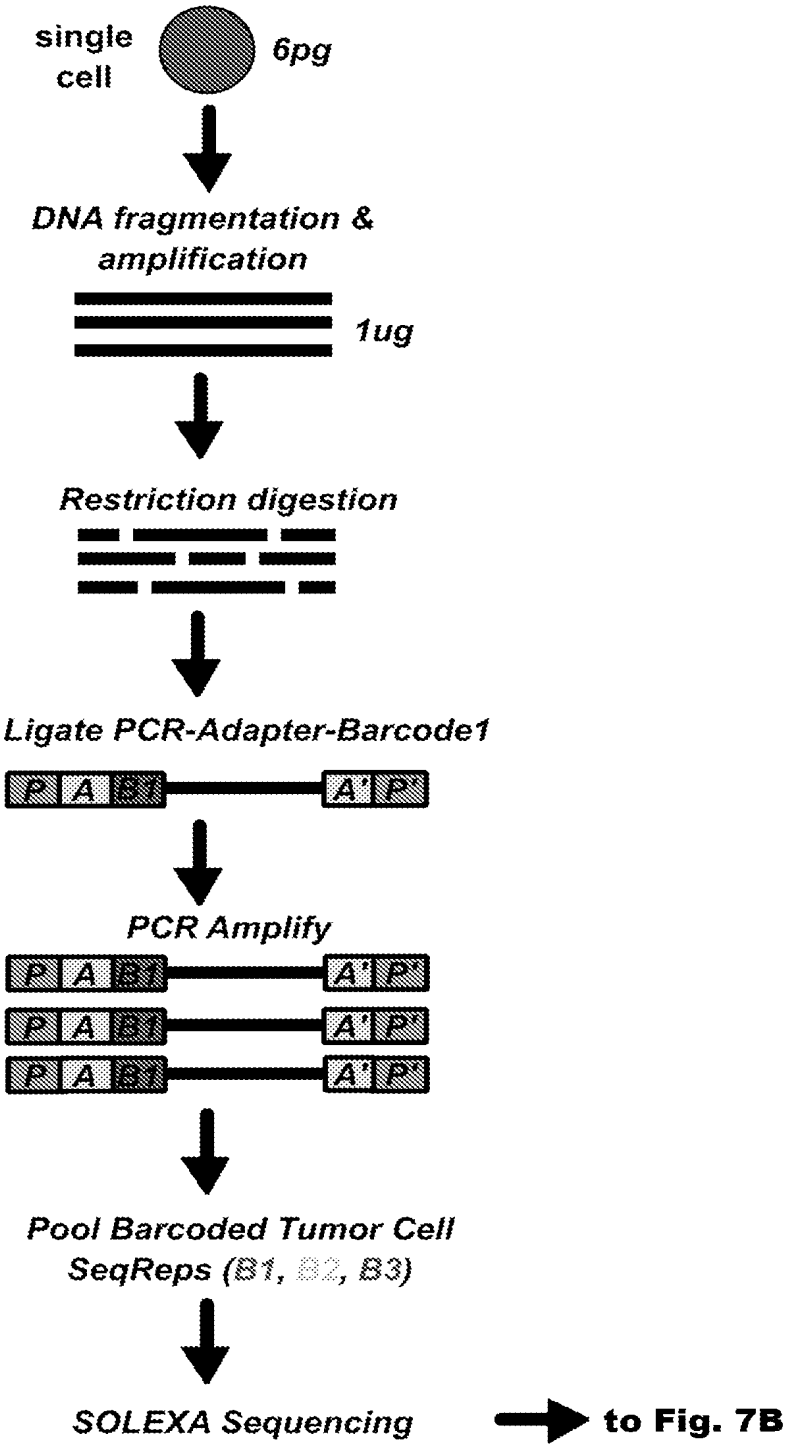
Certificate of Transmittal

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

7-30-14

Robert E. Bolcome III Date
Reg. No. 69,032

Fig. 7A




Applicant(s): James Hicks et al.
U.S. Serial No.: 13/278,333
Filing Date: October 21, 2011
Exhibit A

Fig. 7B (continued from Fig. 7A)

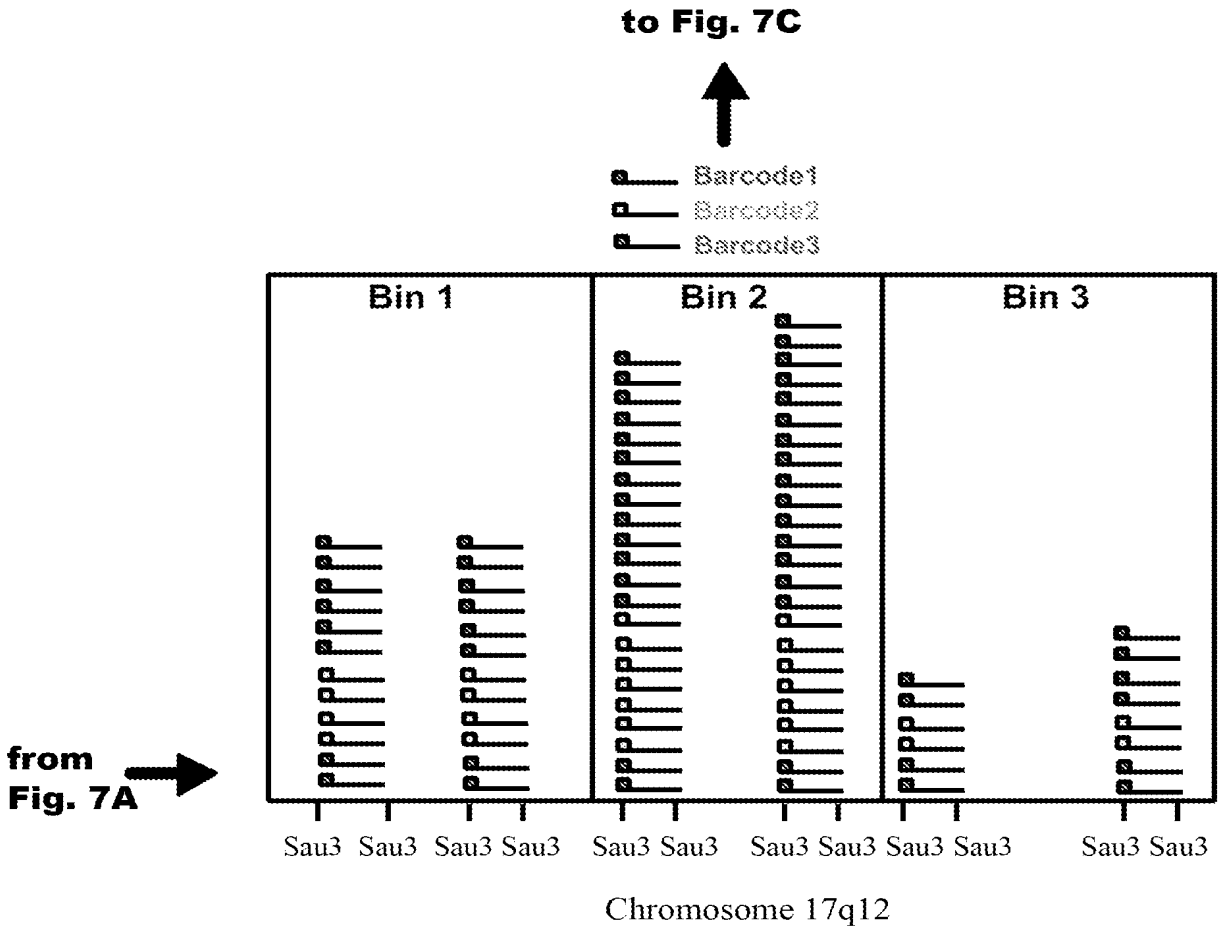
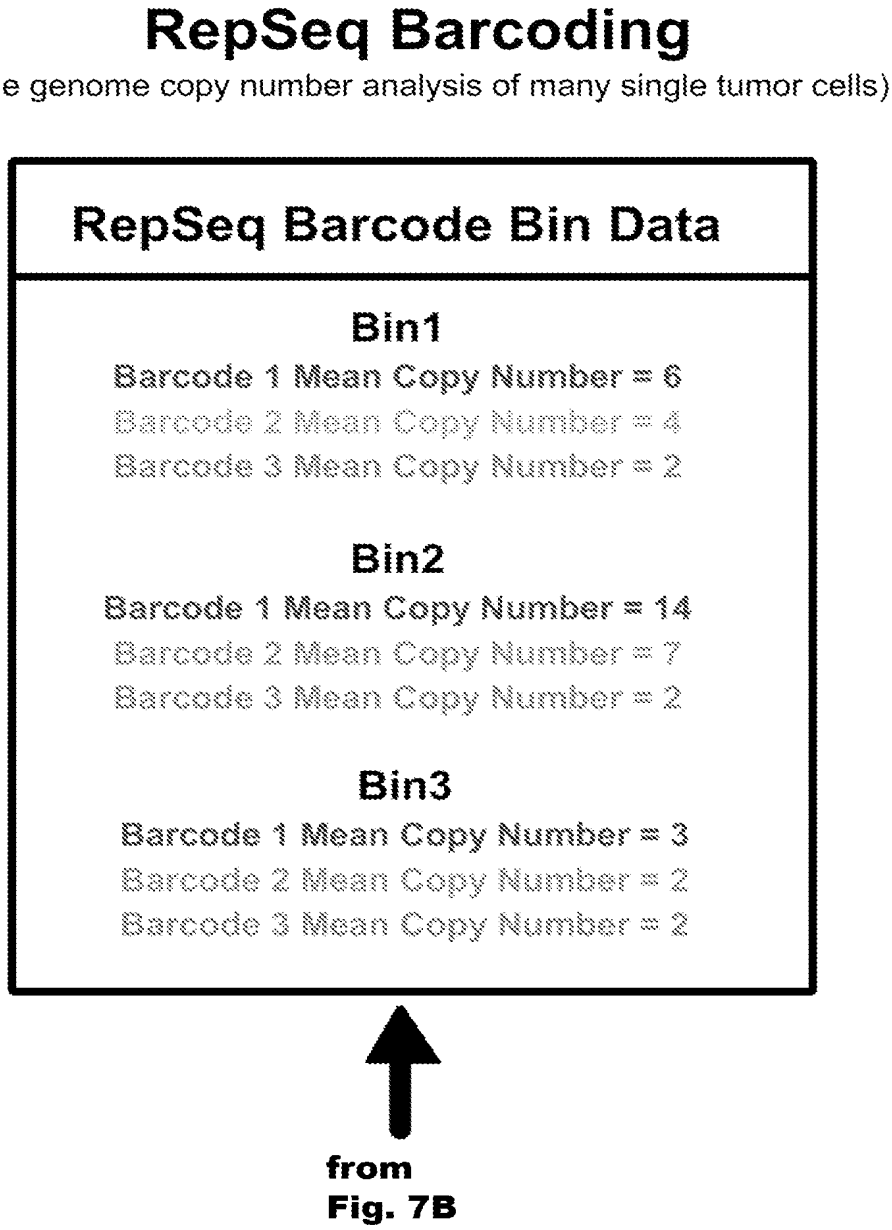


Fig. 7C (continued from Fig. 7B)





IPR2025-01353; IPR2025-01355
CSHL EX2002

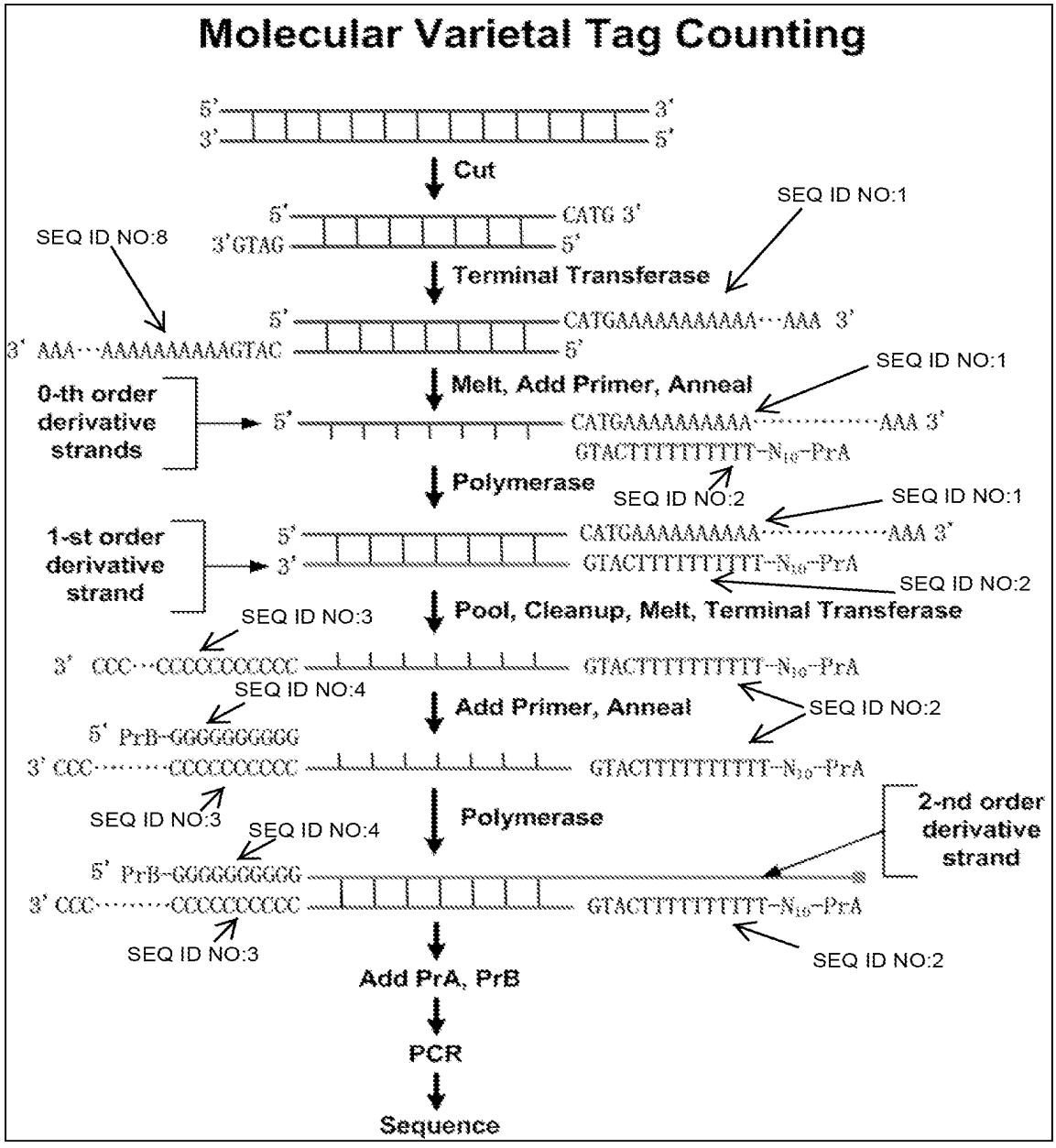
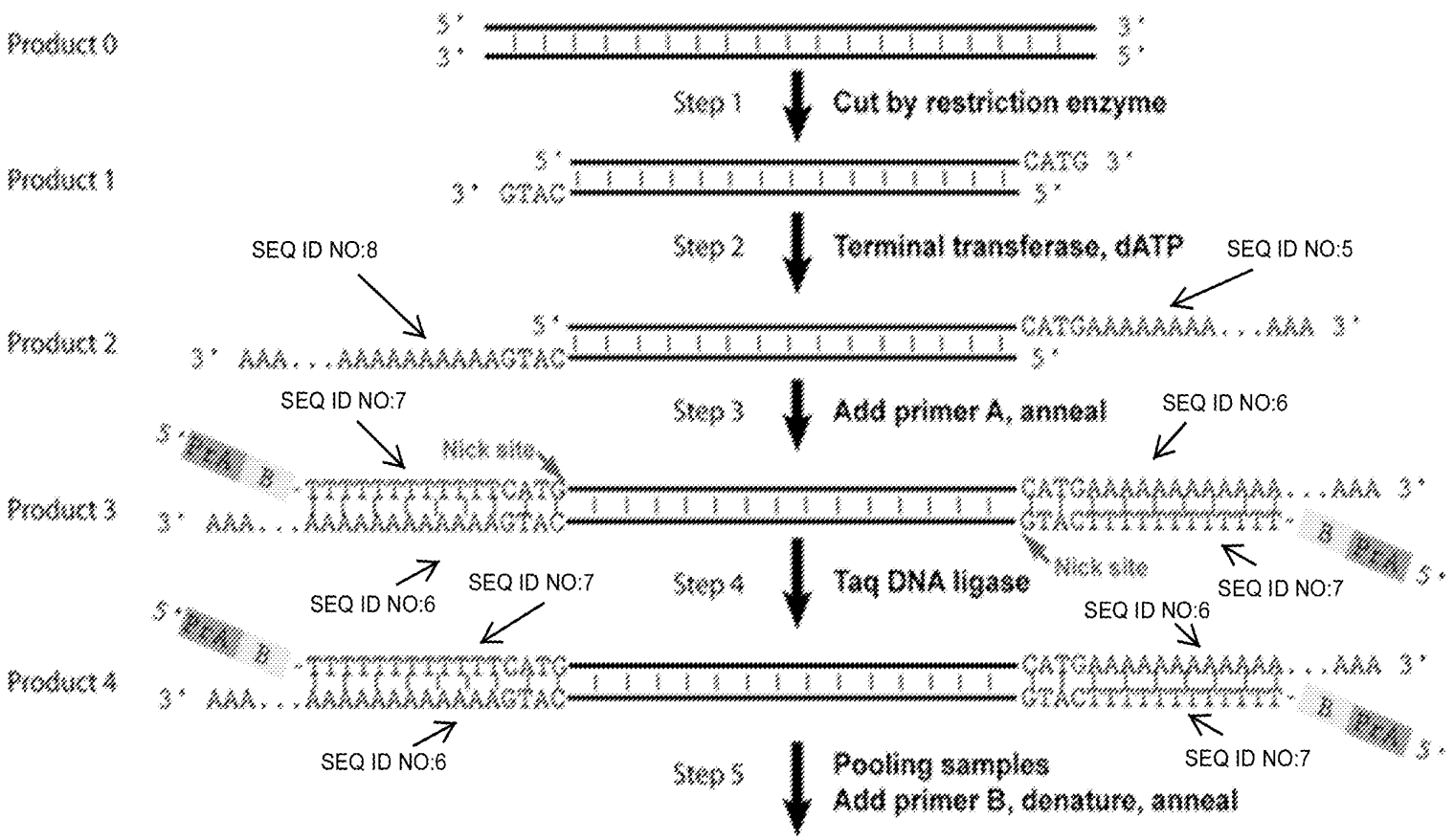

Molecular Varietal Tag Counting
Cut
SEQ ID NO:1
SEQ ID NO:8
Terminal Transferase
Melt, Add Primer, Anneal
0-th order derivative strands
Polymerase
SEQ ID NO:2
1-st order derivative strand
Pool, Cleanup, Melt, Terminal Transferase
SEQ ID NO:3
SEQ ID NO:4
Add Primer, Anneal
Polymerase
2-nd order derivative strand
Add PrA, PrB
PCR
Sequence

Fig. 8

Fig. 9A

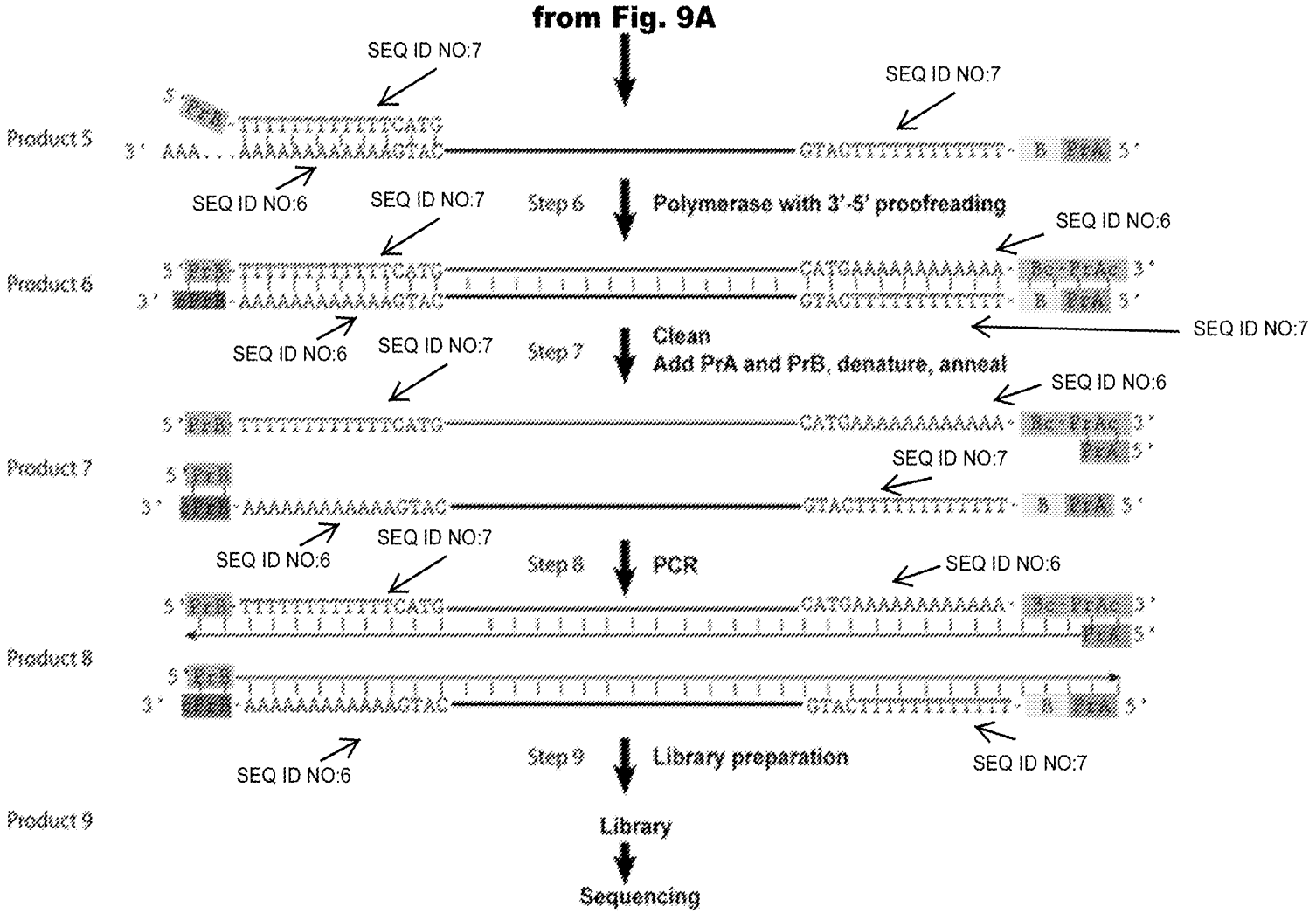



**to Fig. 9B**



IPR2025-01353; IPR2025-01355
CSHL EX2002

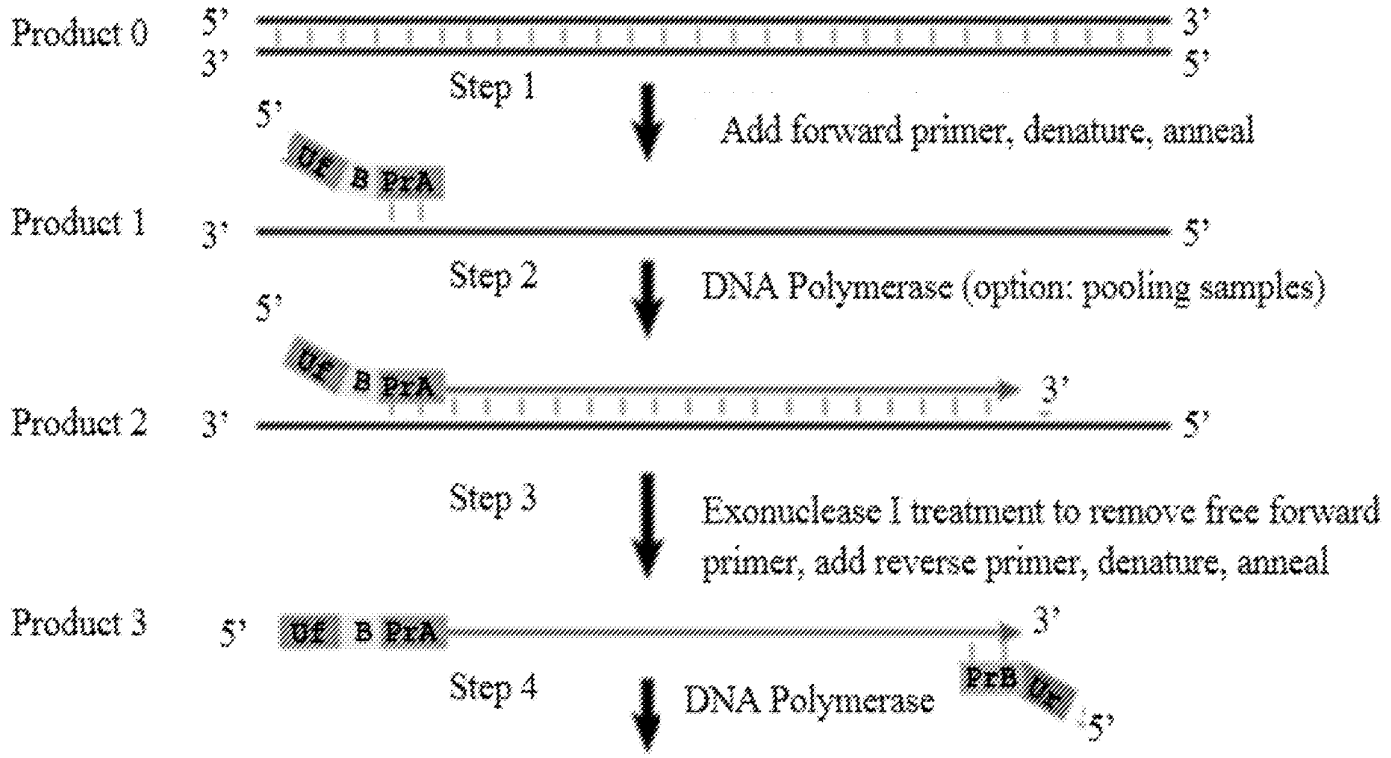
Fig. 9B (continued from Fig. 9A)






IPR2025-01353; IPR2025-01355
CSHL EX2002

Fig. 10A



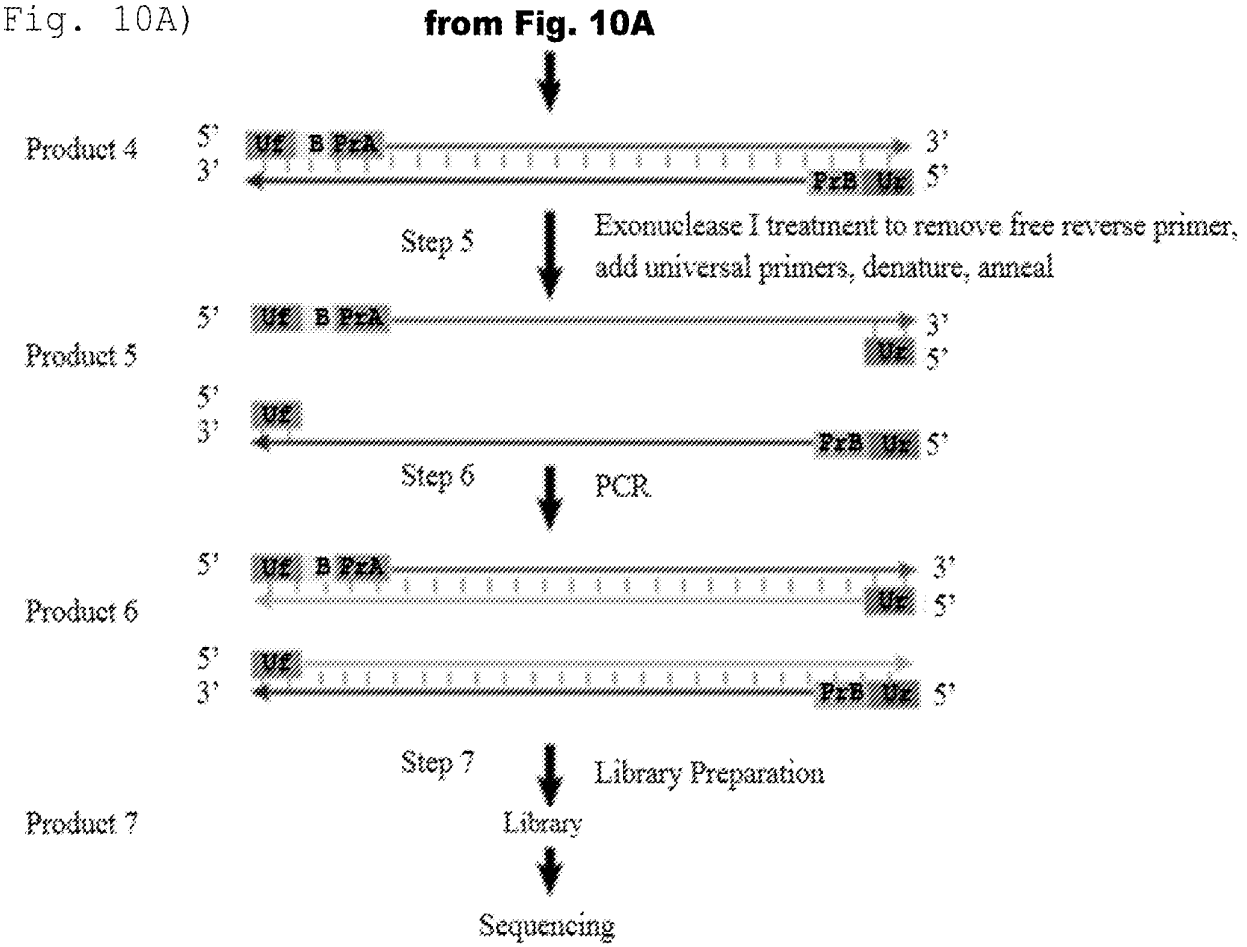

**to Fig. 10B**



IPR2025-01353; IPR2025-01355
CSHL EX2002

Fig. 10B

(continued from Fig. 10A)






IPR2025-01353; IPR2025-01355
CSHL EX2002

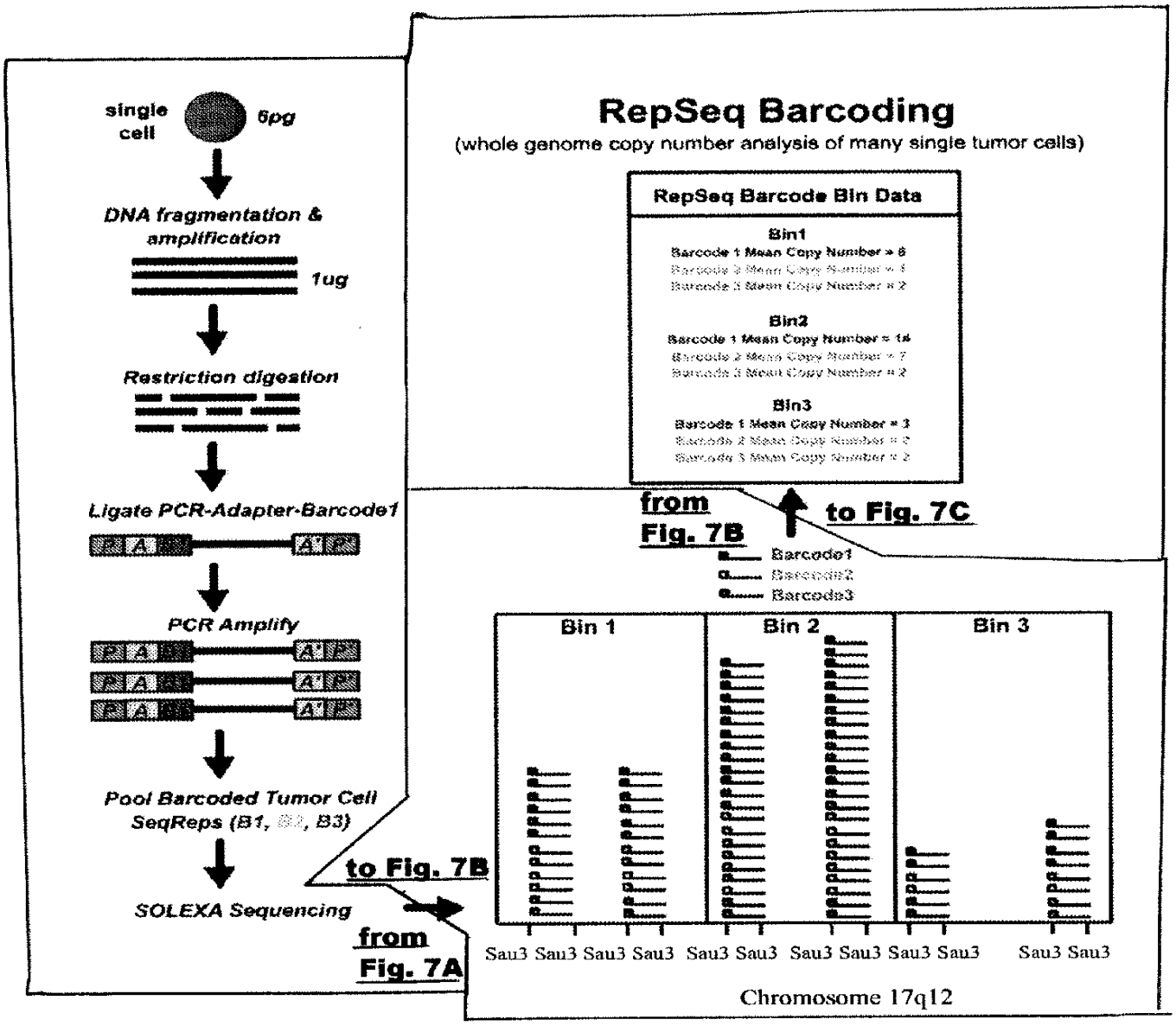



Fig. 7

split into 3 sheets

Applicant(s): James Hicks et al.
U.S. Serial No.: 13/278,333
Filing Date: October 21, 2011
Exhibit B

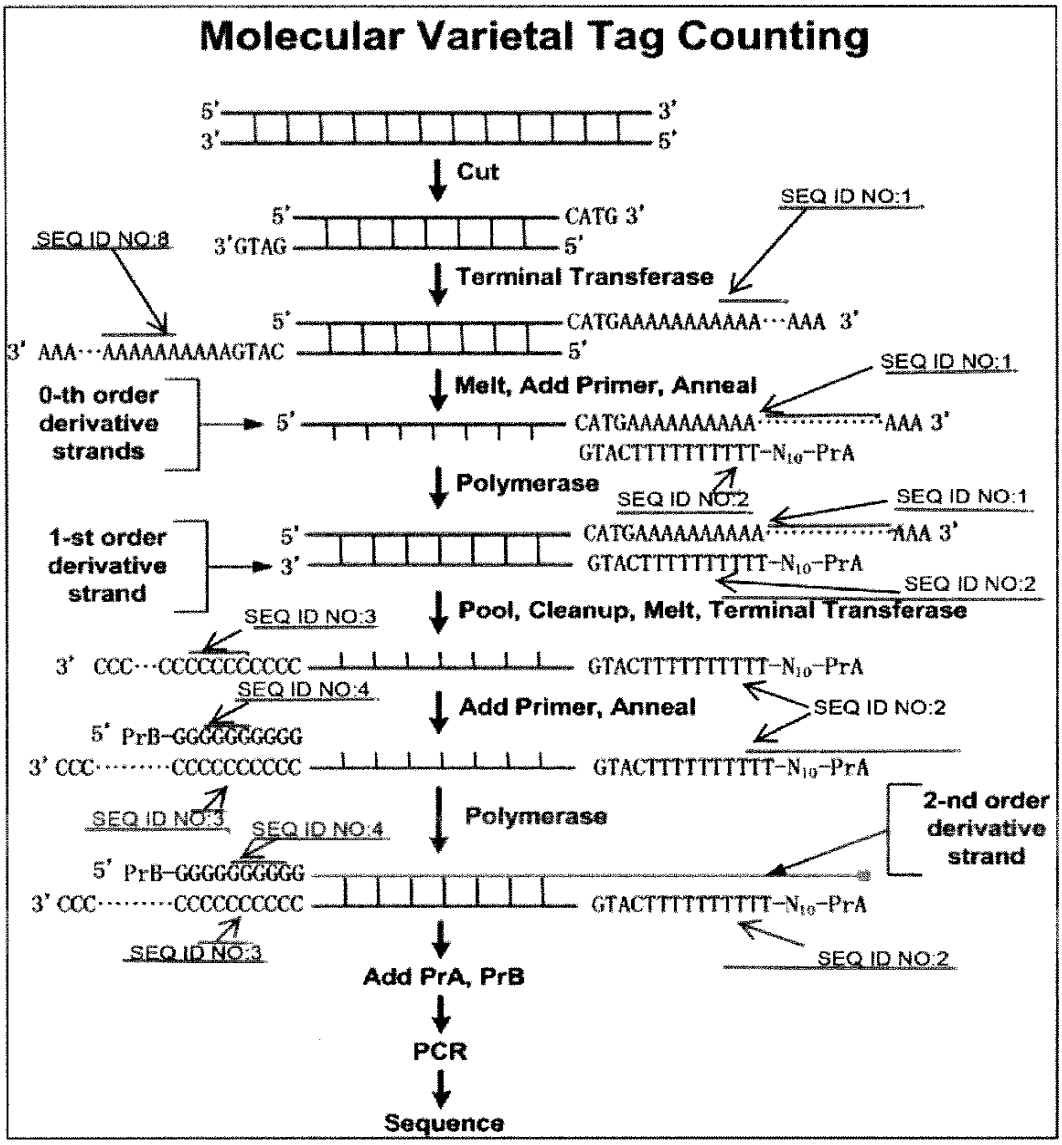

Molecular Varietal Tag Counting
Cut
SEQ ID NO:8
SEQ ID NO:1
Terminal Transferase
Melt, Add Primer, Anneal
0-th order derivative strands
Polymerase
SEQ ID NO:2
1-st order derivative strand
Pool, Cleanup, Melt, Terminal Transferase
SEQ ID NO:3
SEQ ID NO:4
Add Primer, Anneal
2-nd order derivative strand
Add PrA, PrB
PCR
Sequence

Fig. 8

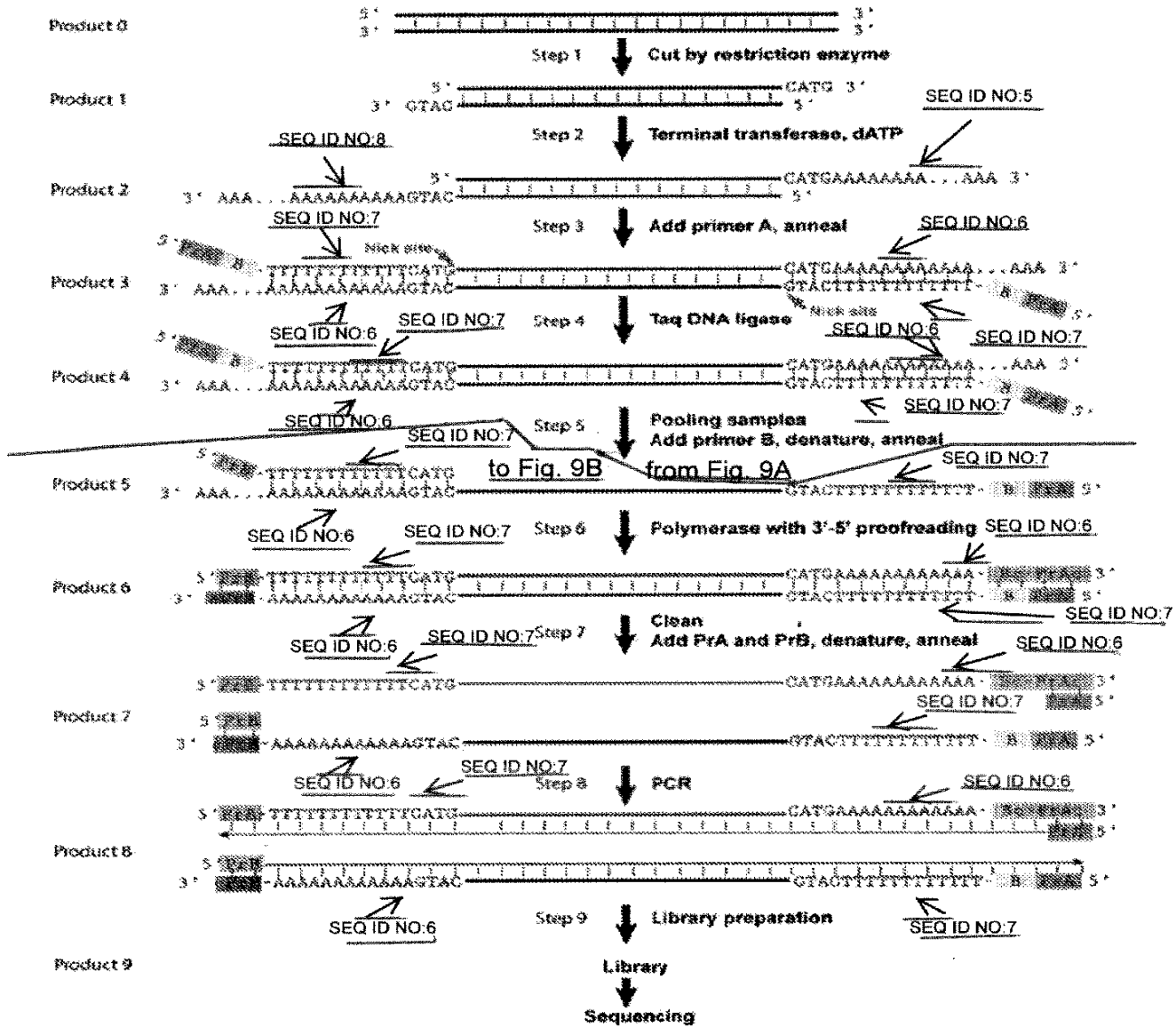

Product 0
Step 1
Cut by restriction enzyme
Product 1
SEQ ID NO:5
SEQ ID NO:8
Step 2
Terminal transferase, dATP
Product 2
SEQ ID NO:7
Step 3
Add primer A, anneal
SEQ ID NO:6
Nick site
Product 3
SEQ ID NO:6
SEQ ID NO:7
Step 4
Taq DNA ligase
Nick site
SEQ ID NO:6
SEQ ID NO:7
Product 4
SEQ ID NO:7
split to multiple sheets
SEQ ID NO:6
SEQ ID NO:7
Step 5
Pooling samples
Add primer B, denature, anneal
to Fig. 9B
from Fig. 9A
SEQ ID NO:7
Product 5
SEQ ID NO:6
SEQ ID NO:7
Step 6
Polymerase with 3'-5' proofreading
SEQ ID NO:6
Product 6
Clean
SEQ ID NO:7
SEQ ID NO:6
SEQ ID NO:7
Step 7
Add PrA and PrB, denature, anneal
SEQ ID NO:6
SEQ ID NO:7
Product 7
SEQ ID NO:6
SEQ ID NO:7
Step 8
PCR
SEQ ID NO:6
Product 8
SEQ ID NO:6
Step 9
Library preparation
SEQ ID NO:7
Product 9
Library
Sequencing

Fig. 9

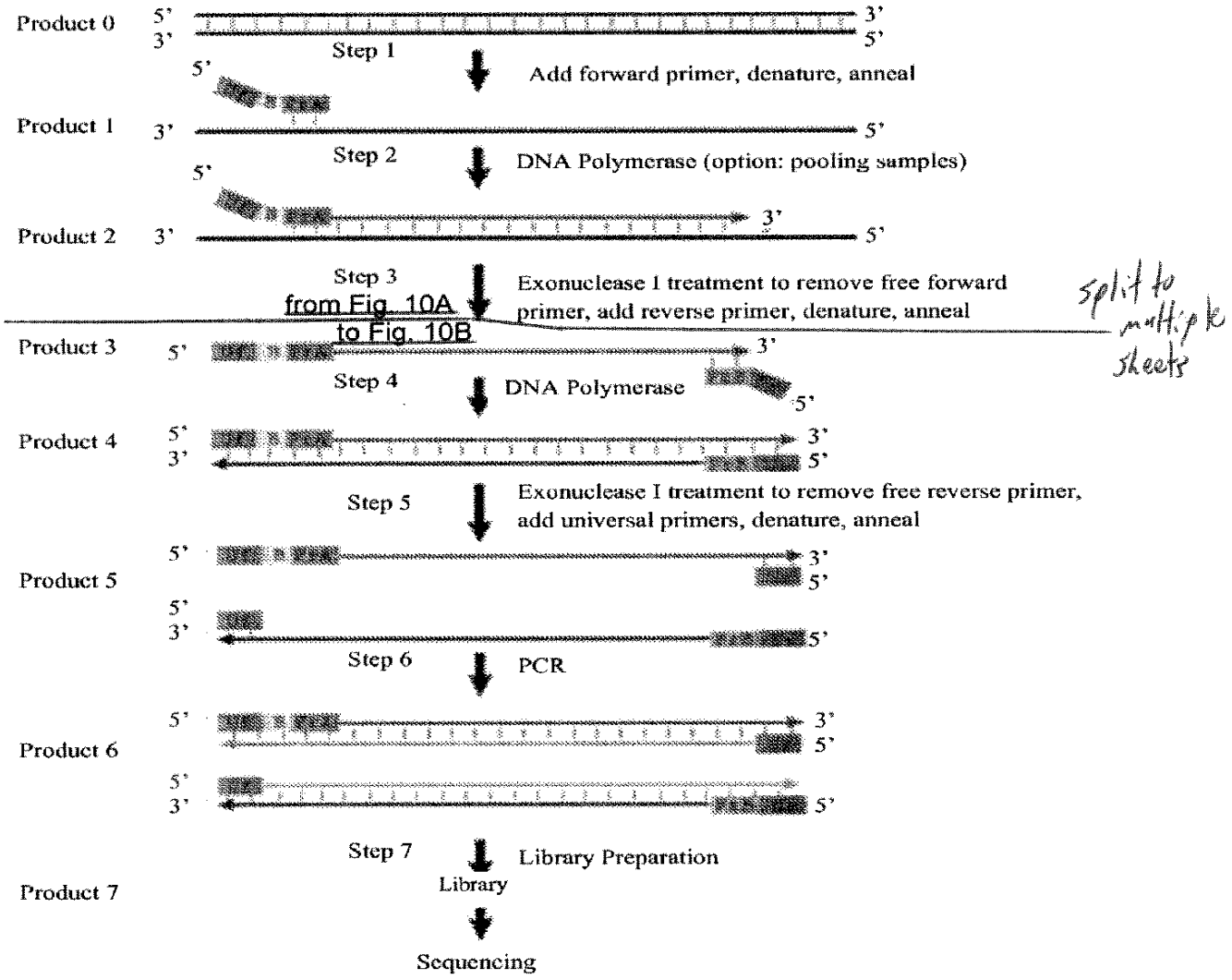

Product 0
5'
3'
3'
5'
Step 1
Add forward primer, denature, anneal
5'
Product 1
3'
5'
Step 2
DNA Polymerase (option: pooling samples)
5'
Product 2
3'
3'
5'
Step 3
from Fig. 10A
to Fig. 10B
Exonuclease I treatment to remove free forward primer, add reverse primer, denature, anneal
split to multiple sheets
Product 3
5'
3'
Step 4
DNA Polymerase
5'
Product 4
5'
3'
3'
5'
Step 5
Exonuclease I treatment to remove free reverse primer, add universal primers, denature, anneal
5'
3'
Product 5
5'
3'
5'
5'
3'
Step 6
PCR
5'
3'
Product 6
5'
3'
5'
5'
3'
Step 7
Library Preparation
Library
Product 7
Sequencing

Fig. 10

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | Application Number | 13/278,333 |
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| | 1 | 8,663,917 | 03-4-2014 | Wigler et al. |
| | 2 | 8,694,263 | 04-08-2014 | Wigler et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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 of 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.

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 13278333 |
| **Filing Date:** | 21-Oct-2011 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | John P. White/Robert Bolcome |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |

Filed as Small Entity

## Utility under 35 USC 111(a) Filing Fees

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| **Extension-of-Time:** | | | | |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Miscellaneous:** | | | | |
| Submission- Information Disclosure Stmt | 2806 | 1 | 90 | 90 |
| | | **Total in USD ($)** | | **90** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 19724817 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Robert Bolcome |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 30-JUL-2014 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 19:03:18 |
| **Application Type:** | Utility under 35 USC 111(a) |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | Deposit Account |
| Payment was successfully received in RAM | $90 |
| RAM confirmation Number | 4969 |
| Deposit Account | 033125 |
| Authorized User | |

## File Listing:

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



IPR2025-01353, IPR2025-01355
CSHL EX2002

| | | | | | |
|---|---|---|---|---|---|
| 1 | | 140730_5981_81503_A_Response_REB.pdf | 7499609<br>d4a6734961fc8078f0902fbf4e4229f1fd1a2322 | yes | 41 |

| Multipart Description/PDF files in .zip description | | |
|---|---|---|
| **Document Description** | **Start** | **End** |
| Amendment/Req. Reconsideration-After Non-Final Reject | 1 | 1 |
| Applicant Arguments/Remarks Made in an Amendment | 2 | 2 |
| Specification | 3 | 3 |
| Claims | 4 | 10 |
| Applicant Arguments/Remarks Made in an Amendment | 11 | 26 |
| Transmittal Letter | 27 | 27 |
| Applicant Arguments/Remarks Made in an Amendment | 28 | 28 |
| Drawings-only black and white line drawings | 29 | 36 |
| Applicant Arguments/Remarks Made in an Amendment | 37 | 40 |
| Information Disclosure Statement (IDS) Form (SB08) | 41 | 41 |

| | | | | | |
|---|---|---|---|---|---|
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Fee Worksheet (SB06) | fee-info.pdf | 30531<br>6799f1de95e33b067c90d25065d98d1663757c1f | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | | **Total Files Size (in bytes):** | 7530140 | |

**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.**

# SCORE Placeholder Sheet for IFW Content

Application Number: 13278333 Document Date: 07/30/2014

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



IPR2025-01353; IPR2025-01355
CSHL EX2002

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<br>Substitute for Form PTO-875 | Application or Docket Number<br>13/278,333 | Filing Date<br>10/21/2011 | ☐ 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 **07/30/2014**

| | (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)) | * 20 | Minus | ** 20 | = 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
/VINCENT BUTLER/

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.*



IPR2025-01353; IPR2025-01355
CSHL EX2002

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 10/08/2014

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER | |
|---|---|
| LU, FRANK WEI MIN | |
| **ART UNIT** | **PAPER NUMBER** |
| 1634 | |
| **MAIL DATE** | **DELIVERY MODE** |
| 10/08/2014 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. 13/278,333 | Applicant(s) HICKS ET AL. | |
|---|---|---|---|
| | Examiner FRANK LU | Art Unit 1634 | AIA (First Inventor to File) Status No |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 2 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *7/30/2014*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims***

5)☒ Claim(s) *1-7,9,11-13,24,26-28,31 and 68-71* is/are pending in the application.

5a) Of the above claim(s) *11* is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☐ Claim(s) _____ is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☒ Claim(s) *1-7,9,12,13,24,26-28,31 and 68-71* are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

**Application Papers**

10)☐ The specification is objected to by the Examiner.

11)☒ The drawing(s) filed on *10/21/2011* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

**Certified copies:**

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date _____.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.


IPR2025-01353; IPR2025-01355
CSHL EX2002

## DETAILED ACTION

### *Response to Amendment*

1. Applicant's response to the office action filed on July 30, 2014 has been entered. The claims pending in this application are claims 1-7, 9, 11-13, 24, 26-28, 31, and 68-71 wherein claim 11 have been withdrawn due to the election of species mailed on December 20, 2013. Claims 1-7, 9, 12, 13, 24, 26-28, 31, and 68-71 will be examined.

### *Election of Species*

2. In view of newly amended claim 24 and newly added claims 68-71, this application contains claims directed to the following patentably distinct species:

(3) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR (claims 24, 68, and 69)

(4) amplified nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d) (claims 24, 70, and 71)

The species are independent or distinct because they are directed to different method steps which have different properties.

In addition, these species are not obvious variants of each other based on the current record.

Applicant is required under 35 U.S.C. 121 to elect a single disclosed species, or a single grouping of patentably indistinct species, for prosecution on the merits to which the claims shall be restricted if no generic claim is finally held to be allowable. Currently, generic claims are 1-7, 9, 12, 13, 26-28, and 31.

There is a search and/or examination burden for the patentably distinct species as set forth above because the searches for species (3) and (4) are not overlap.

**Applicant is advised that the reply to this requirement to be complete must include (i) an election of a species to be examined** even though the requirement may be traversed (37 CFR 1.143) **and (ii) identification of the claims encompassing the elected species or grouping of patentably indistinct species**, including any claims subsequently added. An argument that a claim is allowable or that all claims are generic is considered nonresponsive unless accompanied by an election.

The election may be made with or without traverse. To preserve a right to petition, the election must be made with traverse. If the reply does not distinctly and specifically point out supposed errors in the election of species requirement, the election shall be treated as an election without traverse. Traversal must be presented at the time of election in order to be considered timely. Failure to timely traverse the requirement will result in the loss of right to petition under 37 CFR 1.144. If claims are added after the election, applicant must indicate which of these claims are readable on the elected species or grouping of patentably indistinct species.

Should applicant traverse on the ground that the species, or groupings of patentably indistinct species from which election is required, are not patentably distinct, applicant should submit evidence or identify such evidence now of record showing them to be obvious variants or clearly admit on the record that this is the case. In either instance, if the examiner finds one of the species unpatentable over the prior art, the evidence or admission may be used in a rejection under 35 U.S.C. 103 or pre-AIA 35 U.S.C. 103(a) of the other species.

Upon the allowance of a generic claim, applicant will be entitled to consideration of claims to additional species which depend from or otherwise require all the limitations of an allowable generic claim as provided by 37 CFR 1.141.

A telephone call was made to Mr. Gary Gersnik (Reg. No. 39,992) on October 2, 2014 to request an oral election to the above restriction requirement, but did not result in an election being made.

Applicant is reminded that upon the cancellation of claims to a non-elected invention, the inventorship must be corrected in compliance with 37 CFR 1.48(a) if one or more of the currently named inventors is no longer an inventor of at least one claim remaining in the application. A request to correct inventorship under 37 CFR 1.48(a) must be accompanied by an application data sheet in accordance with 37 CFR 1.76 that identifies each inventor by his or her legal name and by the processing fee required under 37 CFR 1.17(i).

3. Papers related to this application may be submitted to Group 1600 by facsimile transmission. Papers should be faxed to Group 1600 via the PTO Fax Center. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993)(See 37 CAR § 1.6(d)). The CM Fax Center number is (571)273-8300.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Frank Lu, Ph.D., whose telephone number is (571)272-0746. The examiner can normally be reached on Monday-Friday from 9 A.M. to 5 P.M.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dave Nguyen, can be reached on (571)272-0731.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Frank W Lu /
Primary Examiner, Art Unit 1634
October 3, 2014

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
December 8, 2014

BY EFS
Commissioner for Patents
Alexandria, VA

**COMMUNICATION IN RESPONSE TO OCTOBER 8, 2014 OFFICE ACTION**

This Communication is submitted in reply to a October 8, 2014 Office Action issued by the United States Patent and Trademark Office in connection with the above-identified application. A response to the October 8, 2014 Office Action is due December 8, 2014. Accordingly, this response is being timely filed.

**Remarks** begin on page **2** of this paper.

**REMARKS**

Claims 1-7, 9, 11-13, 24, 26-28, 31 and 68-71 are pending in the subject application. Applicants have not herein amended, added, or canceled any claim. Accordingly, claims 1-7, 9, 11-13, 24, 26-28, 31 and 68-71 are pending in the subject application.

**Election of Species**

In the October 8, 2014 Office Action, the Examiner required applicants to elect a single species from among the species listed below for initiating prosecution on the merits and to which the claims shall be restricted if no generic claim is finally held to be allowable:

(3) The tagged nucleic acid molecules are subject to hybrid capture prior to PCR (claims 24, 68, and 69); and

(4) Amplified nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d) (claims 24, 70 and 71).

The Examiner acknowledged that claims 1-7, 9, 12, 13, 26-28, and 31 were generic.

*Applicants' Reply*

In response, applicants elect for initiating examination, Species (4), *i.e.* amplified nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d) (claims 24, 70, and 71).

Claims 1-7, 9, 12, 13, 24, 26-28, 31, 70 and 71 encompass the elected species.

As acknowledged by the Examiner, upon the allowance of a generic claim, applicants will be entitled to consideration of claims to additional species which depend from or otherwise require all the limitations of an allowable generic claim.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicant's undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this response. However, if any fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39.992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

Minerva Y. Wong 12/8/2014
Minerva Y. Wong Date

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 20898240 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Alissa Roland |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 08-DEC-2014 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 18:32:39 |
| **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 | Response to Election / Restriction Filed | 141208_5981_81503_A_Response_MW.pdf | 174820<br>f275cad9dbd9307394d78449d76ae6b01ffaf6d1 | no | 3 |

**Warnings:**

**Information:**



IPR2025-01353; IPR2025-01355
CSHL EX2002

| | |
|---|---|
| **Total Files Size (in bytes):** | 174820 |

**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.**

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 02/04/2015

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER |
|---|
| LU, FRANK WEI MIN |

| ART UNIT | PAPER NUMBER |
|---|---|
| 1634 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 02/04/2015 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. 13/278,333 | Applicant(s) HICKS ET AL. | |
|---|---|---|---|
| | Examiner FRANK LU | Art Unit 1634 | AIA (First Inventor to File) Status No |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *12/8/2014*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims***

5)☒ Claim(s) *1-7,9,11-13,24,26-28 and 68-71* is/are pending in the application.

5a) Of the above claim(s) *11,68 and 69* is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *1-7,9,12,13,24,26-28, 70, and 71* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

**Application Papers**

10)☒ The specification is objected to by the Examiner.

11)☒ The drawing(s) filed on *7/30/2014* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

**Certified copies:**

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)

2) ☒ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date *7/30/2014*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

## DETAILED ACTION

### *Election/Restrictions*

1. Applicant's election of species (4) (amplified nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d), see claims 24, 70, and 71) in the reply filed on December 8, 2014 is acknowledged. Because applicant did not distinctly and specifically point out the supposed errors in the restriction requirement, the election has been treated as an election without traverse (MPEP § 818.03(a)). Claims 1-7, 9,12,13,24,26-28, 70, and 71 will be examined.

### *Specification*

2. The disclosure is objected to because of the following informality: what are reference sources of Craig et al., 2008, Albert et al., 2007 and Okou et al., 2007 in paragraph [0324] of US 2014/0065609 A1, which is a US Publication of this instant application.

### *Claim Objections*

3. Claim 24 or 70 is objected to because of the following informality: "prior to sequencing in step (d)" should be "prior to step (d)".

4. Claim 70 is objected to because of the following informality: "amplified tagged nucleic acid molecules" should be "the amplified tagged nucleic acid molecules".

Appropriate correction is required.

### *Claim Rejections - 35 USC § 112*

5. The following is a quotation of the first paragraph of 35 U.S.C. 112(a):

> (a) IN GENERAL.—The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor or joint inventor of carrying out the invention.

The following is a quotation of the first paragraph of pre-AIA 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor of carrying out his invention.

6. Scope of Enablement

Claims 1-7, 9, 12, 13, 24, 26-28, 31, 70, and 71 are rejected under 35 U.S.C. 112(a) or 35 U.S.C. 112 (pre-AIA), first paragraph, because the specification, while being enabling for generating tagged nucleic acid molecules, does not reasonably provide enablement for performing the methods recited in claims 1-7, 12, 13, 24, 26-28, 31, 70, and 71. The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to use the invention commensurate in scope with these claims.

Factors to be considered in determining whether a disclosure meets the enablement requirement of 35 USC 112, first paragraph, have been described by the court in *In re Wands*, 8 USPQ2d 1400 (CA FC 1988). *Wands* states at page 1404,

"Factors to be considered in determining whether a disclosure would require undue experimentation have been summarized by the board in Ex parte Forman. They include (1) the quantity of experimentation necessary, (2) the amount of direction or guidance presented, (3) the presence or absence of working examples, (4) the nature of the invention, (5) the state of the prior art, (6) the relative skill of those in the art, (7) the predictability or unpredictability of the art, and (8) the breadth of the claims."

<u>The nature of the invention</u>

The claims are drawn to a method for obtaining from genomic material genomic copy number information unaffected by amplification distortion and a method for obtaining from mRNA transcriptions mRNA copy number information unaffected by amplification distortion. The invention is a class of invention which the CAFC has characterized as "the unpredictable arts such as chemistry and biology." *Mycogen Plant Sci., Inc. v. Monsanto Co.,* 243 F.3d 1316, 1330 (Fed. Cir. 2001).

The Breadth of The Claims

Claims 1-7, 12, 13, 24, 26-28, 31, 70, and 71 encompass a method for obtaining from genomic material genomic copy number information unaffected by amplification distortion comprising a) obtaining segments of the genomic material; b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the genomic material from step (a) and a tag; c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR); d) generating tag associated sequence reads by sequencing the product of step (c); e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to one segment of the segments of the genomic material to a location on the genome; and f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome. Claim 9 encompasses a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion comprising a) generating tagged nucleic acid molecules, comprising: i) subjecting the

mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands; ii) adding a polynucleotide tail to the first order derivative strands; and iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands, wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules; b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification; c) generating tag associated sequence reads by sequencing the product of step (b); d) assigning each of the tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each of the tag associated sequence reads corresponding to a mRNA transcript to a location on the cDNA library; and e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the CDNA library.

Working Examples

The specification provides working examples (see pages 56-62) for Methylation sequencing, Comparative genomic hybridization (CGH) by high throughput sequencing, Single cell microarray and sequencing, Sequencing across chromosome breakpoints, and Copy number profiling using varietal tags on restriction enzyme fragments. The specification provides no working example or the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, 31, 70, and 71.

The Amount of Direction or Guidance Provided and The State of The Prior Art

Although the specification provides working examples (see pages 56-62) for Methylation sequencing, Comparative genomic hybridization (CGH) by high throughput sequencing, Single cell microarray and sequencing, Sequencing across chromosome breakpoints, and Copy number profiling using varietal tags on restriction enzyme fragments, the specification does not provide a guidance to show that the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, 31, 70, and 71 can be performed. Furthermore, there is no experimental condition and/or experimental data in the specification to support the claimed invention. During the process of the prior art search, the examiner has not found any prior art which is related to the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, 31, 70, and 71.

Level of Skill in The Art, The Unpredictability of The Art, and The Quantity of Experimentation Necessary

While the relative skill in the art is very high (the Ph.D. degree with laboratory experience), there is no predictability whether the methods recited in claims 1-7, 9, 12, 13, 24, 26-28, 31, 70, and 71 can be performed.

First, since step c) of claim 1 does not require that the product of step (c) is an amplification product of a specific gene amplified using primers only specific for the specific gene and a nucleic acid sequencing only can be performed using a specific nucleic acid and cannot be performed using a nucleic acid mixture comprising a different nucleic acids, if segments of the genomic material are a genomic DNA library and the product of step (c) is an amplification product of the genomic library amplified using a universal primer and one of the

substantially unique tags, the product of step (c) comprises a mixture of different genomic DNAs, it is unclear how the product of step (c) such as a mixture of different genomic DNAs can be sequenced so that tag associated sequence reads cannot be generated and each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material cannot be assigned by mapping the subsequence of each tag associated sequence read corresponding to one segment of the segments of the genomic material to a location on the genome as recited in steps d) and e) of claim 1. Furthermore, since tag associated sequence reads in step c) of claim 1 can be reasonably read as reads of the sequences of the tags, step d) of claim 1 does not require completely sequencing the product of step c). If step d) of claim 1 only require sequencing the tag of the product of step c), tag associated sequence reads only comprise sequence of the tag of the product of step c) and does not comprise other sequences of the tagged nucleic acid molecules. Thus, it is unclear how each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material can be assigned by mapping the subsequence of each tag associated sequence read corresponding to one segment of the segments of the genomic material to a location on the genome so that the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome cannot be counted as recited in steps e) and f) of claim 1. In addition, since claim 1 is directed to a method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, the product of step (c) must contain amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules. Since claim 1 does not indicate how to differentiate the amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged

nucleic acid molecules, it is unclear how to obtain genomic copy number information unaffected by amplification distortion by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome.

Second, in view of step a) of claim 9, tagged nucleic acid molecules are a cDNA library. Since step c) of claim 9 does not require that the product of step (b) is an amplification product of a specific gene amplified using primers only specific for the specific gene, and a nucleic acid sequencing only can be performed using a specific nucleic acid and cannot be performed using a nucleic acid mixture comprising a different nucleic acids, if the product of step (b) is an amplification product of the cDNA library amplified using a universal primer and one of the substantially unique primers, the product of step (b) comprises a mixture of cDNAs, it is unclear how to the product of step (b) such as a mixture of cDNAs can be sequenced so that tag associated sequence reads cannot be generated and each of the tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts cannot be assigned by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript of the mRNA transcripts to a location on the cDNA library as recited in steps c) and d) of claim 9. Furthermore, since tag associated sequence reads in step c) of claim 9 can be reasonably read as reads of the sequences of the tags, step c) of claim 9 does not require completely sequencing the product of step c). If step c) of claim 9 only require sequencing the tag of the product of step b), tag associated sequence reads only comprise sequence of the tag of the product of step b) and does not comprise other sequences of the tagged nucleic acid molecules. Thus, it is unclear how each of the tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts can be assigned by mapping the

subsequence of each tag associated sequence read corresponding to a mRNA transcript of the mRNA transcripts to a location on the cDNA library so that the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library cannot be counted as recited in steps d) and e) of claim 9. In addition, since claim 9 is directed to a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, the product of step (b) must contain amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules. Since claim 9 does not indicate how to differentiate the amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules, it is unclear how to obtain mRNA copy number information unaffected by amplification distortion by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library.

Third, in view of the specification and claim 2, it is unclear how the highest count obtained in step (f) for the locations within the region must be assigned as the copy number of the region and the average count or median count obtained in step (f) for the locations within the region, if present, cannot be assigned as the copy number of the region.

In view of above discussions, the skilled artisan will have no way to predict the experimental results. Accordingly, it is concluded that undue experimentation is required to make the invention as it is claimed. The undue experimentation at least includes to test whether the method recited in claims 1-7, 12, 13, 24, 26-28, 31, 70, and 71 can be performed.

Conclusion

In the instant case, as discussed above, the level of unpredictability in the art is high, the specification provides one with no guidance that leads one to claimed methods. One of skill in the art cannot readily anticipate the effect of a change within the subject matter to which the claimed invention pertains. Thus given the broad claims in an art whose nature is identified as unpredictable, the unpredictability of that art, the large quantity of research required to define these unpredictable variables, the lack of guidance provided in the specification, the absence of any working example related to claimed invention and the no teaching in the prior art balanced only against the high skill level in the art, it is the position of the examiner that it would require undue experimentation for one of skill in the art to perform the method of the claim as broadly written.

***Response to Arguments***

I. In page 16, fifth paragraph bridging to page 18, first paragraph of applicant's remarks filed on July 30, 2014, applicant argues that "[P]aragraph [0010] of the subject specification states that a method 'that allows for copy number determination free of distortions caused by amplification steps and which allows for accurate and efficient copy number determination of complex samples' using existing methodologies for high volume, massively parallel sequencing' was needed. The subject specification goes on to explain that it provides such methods, and that steps of the methods provided may use well known and understood high-throughput sequencing and mapping techniques. *See, e.g.,* paragraphs [0285]-[0288], [0321], [0322], [0332] and [0336], and Figure 7. Additionally, paragraph [0332] discloses that applicants successfully obtained 'genome-wide' copy number data. *See also,* Figure IIA. Paragraph [0285] shows that applicants successfully sequenced and mapped hundreds of thousands of genomic fragments. *See also,* Figure 7.[2] Paragraphs [0331] and [0324] of the subject application describe the parallel

sequencing of many tagged nucleic acid molecules. For example, paragraph [0324] describes an available sequencing technique which 'can provide millions of individual sequence reads randomly distributed across any input DNA' (citation omitted). With respect to mRNA, paragraph [0322] explains that '[s]everal methods have now been published whereby researchers have profiled the RNA content of single cells by sequencing' (citation omitted). The subject application also describes how to map multiple sequence reads to a genome or cDNA library. *See,* e.g., paragraph [0255], which states 'the term 'mapping' refers to identifying on a or cDNA library' (emphasis added)). Paragraph [0282] explains that mapping can readily be achieved by aligning sequence reads to specific locations on a map of the human genome".

These arguments have been fully considered but they are not persuasive toward the withdrawal of the rejection. First, since step c) of claim 1 does not require that the product of step (c) is an amplification product of a specific gene amplified using primers only specific for the specific gene and a nucleic acid sequencing only can be performed using a specific nucleic acid (eg., see abstract and Figure 1 of Craig et al., Nat. Methods. 5(10), 887-893, October 2008, the reference cited in paragraph [0324] of US 2014/0065609 A1 which is a US Publication of this instant application) and cannot be performed using a nucleic acid mixture comprising a different nucleic acids, if segments of the genomic material are a genomic DNA library and the product of step (c) is an amplification product of the genomic library amplified using a universal primer and one of the substantially unique tags, the product of step (c) comprises a mixture of different genomic DNAs, it is unclear how the product of step (c) such as a mixture of different genomic DNAs can be sequenced so that tag associated sequence reads cannot be generated and each of the tagged nucleic acid molecules to a location on a genome associated with the genomic

material cannot be assigned by mapping the subsequence of each tag associated sequence read corresponding to one segment of the segments of the genomic material to a location on the genome as recited in steps d) and e) of claim 1. Second, Since step c) of claim 9 does not require that the product of step (b) is an amplification product of a specific gene amplified using primers only specific for the specific gene, and a nucleic acid sequencing only can be performed using a specific nucleic acid (eg., see abstract and Figure 1 of Craig et al., Nat. Methods. 5(10), 887-893, October 2008) and cannot be performed using a nucleic acid mixture comprising a different nucleic acids, if the product of step (b) is an amplification product of the cDNA library amplified using a universal primer and one of the substantially unique primers, the product of step (b) comprises a mixture of cDNAs, it is unclear how to the product of step (b) such as a mixture of cDNAs can be sequenced so that tag associated sequence reads cannot be generated and each of the tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts cannot be assigned by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript of the mRNA transcripts to a location on the cDNA library as recited in steps c) and d) of claim 9.

II. In page 18, second to fourth paragraphs of applicant's remarks filed on July 30, 2014, applicant argues that "neither claim 1 nor claim 9 recites the term 'mutant'. Accordingly, applicants are unsure how the Examiner's statements relate to the claims. To the extent the Examiner was referring to mutations introduced during the performance of claim steps, paragraph [0311] explains: 'Through sequence error during processing, the tag may very occasionally mutate, creating the appearance of a new tagged nucleic acid molecule. But if the set of tags are sufficiently large, the tagged nucleic acid molecules will generally differ at more

than one nucleotide, and so tagged nucleic acid molecules that differ by a single nucleotide can (optionally) be ignored when counting.' Accordingly, the subject application provided ample guidance to enable a POSA to account for unwanted mutations which might unexpectedly and rarely occur".

These arguments have been fully considered but they are not persuasive toward the withdrawal of the rejection. First, since claim 1 is directed to a method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, the product of step (c) must contain amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules. Since claim 1 does not indicate how to differentiate the amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules, it is unclear how to obtain genomic copy number information unaffected by amplification distortion by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome. Second, since claim 9 is directed to a method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, the product of step (b) must contain amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules. Since claim 9 does not indicate how to differentiate the amplification products having original tagged nucleic acid molecules and the amplification products having mutant tagged nucleic acid molecules, it is unclear how to obtain mRNA copy number information unaffected by amplification distortion by counting the number of tagged

nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library.

III. In page 19, fifth paragraph bridging to page 20, first paragraph of applicant's remarks filed on July 30, 2014, applicant argues that "[S]tep d) of claim 1 recites 'generating tag associated sequence reads by sequencing the product of step (c)'. The Examiner's alleged embodiment of sequencing only tags (without also sequencing segments of genetic material) does not fall within the scope of step d). Step d) recites generating 'tag associated sequence reads' (emphases added). Step d) does not recite generating mere 'tag sequence reads'.

These arguments have been fully considered but they are not persuasive toward the withdrawal of the rejection. Since the phrase "tag associated sequence reads" can be reasonably read as reads of the sequences of the tags, step d) of claim 1 does not require completely sequencing the product of step c) or step c) of claim 9 does not require completely sequencing the product of step (b).

7. The following is a quotation of 35 U.S.C. 112(b):

> (b) CONCLUSION.—The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention.

> The following is a quotation of 35 U.S.C. 112 (pre-AIA), second paragraph:
> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

8. Claims 1-7, 9, 12, 13, 24, 26-28, 31, 70, and 71 are rejected under 35 U.S.C. 112(b) or 35 U.S.C. 112 (pre-AIA), second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which the inventor or a joint inventor, or for pre-AIA the applicant regards as the invention.

9. Claim 1 is rejected as vague and indefinite in view of step b) because it is unclear what is the relationship between a tag and substantially unique tags. Please clarify. Note that applicant has not addressed this issue.

10. Claim 1 is rejected as vague and indefinite in view of step e) because it is unclear what is the relationship between tag associated sequence reads in step d) and each tag associated sequence read in step e). Please clarify. Note that applicant has not addressed this issue.

11. Claim 1 is rejected as vague and indefinite in view of step f) because it is unclear what is the relationship between a different tag in step f) and substantially unique tags in step b). Please clarify. Note that applicant has not addressed this issue.

12. Claim 7 is rejected as vague and indefinite because it is unclear what kind of sum means the sum in line 3. Please clarify. Note that applicant has not addressed this issue.

13. Claim 9 is rejected as vague and indefinite in view of step d) because it is unclear what is the relationship between tag associated sequence reads in step c) and each tag associated sequence read in step d). Please clarify. Note that applicant has not addressed this issue.

14. Claim 9 is rejected as vague and indefinite in view of step e). Since step a) does not require that the tagged nucleic acid molecules have different tag, it is unclear where a different tag in step e) comes from. Please clarify. Note that applicant has not addressed this issue.

15. Claim 12 is rejected as vague and indefinite in view of step a) because it is unclear what is the relationship between substantially unique tags in step b) of claim 1 and a polynucleotide tail in step a) of claim 12. Please clarify.

16. Claim 26 recites the limitation "the tag sequences" in the claim. There is insufficient antecedent basis for this limitation in the claim because there is no phrase "tag sequences" in claim 1. Please clarify.

***Conclusion***

17. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

18. No claim is allowed.

19. Papers related to this application may be submitted to Group 1600 by facsimile transmission. Papers should be faxed to Group 1600 via the PTO Fax Center. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993)(See 37 CAR § 1.6(d)). The CM Fax Center number is (571)273-8300.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Frank Lu, Ph.D., whose telephone number is (571)272-0746. The examiner can normally be reached on Monday-Friday from 9 A.M. to 5 P.M.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dave Nguyen, can be reached on (571)272-0731.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Frank W Lu /
Primary Examiner, Art Unit 1634
January 30, 2015

| Search Notes (continued) | Application/Control No. | Applicant(s)/Patent under Reexamination | |
|---|---|---|---|
| | 13/278,333 | HICKS ET AL. | |
| | Examiner | Art Unit | |
| | FRANK LU | 1634 | |

| SEARCHED | | | |
|---|---|---|---|
| Class | Subclass | Date | Examiner |
| Update | | 1/30/2015 | /FL/ |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| INTERFERENCE SEARCHED | | | |
|---|---|---|---|
| Class | Subclass | Date | Examiner |
| | | | |
| | | | |
| | | | |
| | | | |

| SEARCH NOTES (INCLUDING SEARCH STRATEGY) | | |
|---|---|---|
| | DATE | EXMR |
| Update | 1/30/2015 | /FL/ |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

| Notice of References Cited | Application/Control No. 13/278,333 | Applicant(s)/Patent Under Reexamination HICKS ET AL. | |
|---|---|---|---|
| | Examiner FRANK LU | Art Unit 1634 | Page 1 of 1 |

**U.S. PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date MM-YYYY | Name | Classification |
|---|---|---|---|---|---|
| | A | US- | | | |
| | B | US- | | | |
| | C | US- | | | |
| | D | US- | | | |
| | E | US- | | | |
| | F | US- | | | |
| | G | US- | | | |
| | H | US- | | | |
| | I | US- | | | |
| | J | US- | | | |
| | K | US- | | | |
| | L | US- | | | |
| | M | US- | | | |

**FOREIGN PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date MM-YYYY | Country | Name | Classification |
|---|---|---|---|---|---|---|
| | N | | | | | |
| | O | | | | | |
| | P | | | | | |
| | Q | | | | | |
| | R | | | | | |
| | S | | | | | |
| | T | | | | | |

**NON-PATENT DOCUMENTS**

| * | | Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages) |
|---|---|---|
| | U | Craig et al., IDENTIFICATION OF GENETIC VARIANTS USING BARCODED MULTIPLEXED SEQUENCING. Nat. Methods. 5(10), 887-893, October 2008. |
| | V | |
| | W | |
| | X | |

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | First Named Inventor | James Hicks |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| /F.L./ | 1 | 8,663,917 | 03-4-2014 | Wigler et al. |
| /F.L./ | 2 | 8,694,263 | 04-08-2014 | Wigler et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 01/31/2015 |
|---|---|---|---|

*EXAMINER: Initial if citation 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 of 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.

RECEIVED
CENTRAL FAX CENTER

APR 20 2015

Applicant : James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed : October 21, 2011


**AGENDA FOR TELEPHONIC EXAMINER INTERVIEW**

**Rejections under 35 U.S.C. § 112, paragraph 1 (enablement)**

I **A person having ordinary skill in the art (POSA) would be enabled to amplify and sequence tagged nucleic acid molecules.**

In response to the Examiner's comments on pages 11-12 of the February 4, 2015 Final Office Action:

The Examiner's assertion that universal primers cannot be used with the claimed method is incorrect. Either specific or universal primers capable of hybridizing to sequences of the genomic material can be used in the claimed method.

Figure 8 of the subject specification, described in paragraphs [0040], [0041], [0042], [303], and [304] illustrates one embodiment of the method described in step (c). Figure 8 shows that following addition of terminal transferase, a poly-A tail is added to the ends of the segments of the genomic material to generate $0^{th}$ order derivative strands. A primer (PrA) comprising a tag ($N_{10}$) hybridizes to the poly-A tail added to the segment of the genomic material. Thus, in Figure 8, the hybridization of primer (PrA) with the poly-A tail is in effect analogous to the effect of using a universal primer. Yet, the method still works. A copy of Figure 8 is attached hereto as **Exhibit A**.

A schematic example of the claim is shown in **Exhibit B**.

**II A POSA could perform the claimed methods even if the amplification product of step (c) contained mutant tagged nucleic acid molecules, as alleged by the Examiner.**

In response to the Examiner's comments on pages 12-14 of the February 4, 2015 Final Office Action:

a) While mutations can occur in PCR, mutations were known to occur at a rate of $10^{-5}$ to $10^{-6}$ mutations/base pair (see i.e. Table 1 of Cline et al., Nucleic Acids Research 1996; attached hereto as **Exhibit C**).

b) Even if a mutation was introduced in 1 out of 100,000 base pairs, it was well understood in the art and well within the skill of a POSA to compensate for such mutation.

c) A POSA could 1) recognize an aberrant "mutant" subsequence during the mapping of step (e); and 2) discard the aberrant "mutant" subsequence during the counting of step (f).

d) The subject specification provides guidance for a POSA to account for unwanted mutations which might unexpectedly and rarely occur. Paragraph [0311] explains that:

> "tagged nucleic acid molecules that differ by a single nucleotide can (optionally) be ignored when counting".





IPR2025-01353, IPR2025-01355
CSHL EX2002

**III A "tag associated sequence read" is defined to require sequencing of the tag and at least a portion of the nucleic acid.**

In response to the Examiner's comments on page 14 of the February 4, 2015 Final Office Action:

a) "Sequence reads" when referring to tagged nucleic acid molecules are defined in the specification as containing (1) a subsequence of a tag, and (2) a subsequence of a nucleic acid in the sample:

> "In the disclosed method, each sequence read generated from a tagged nucleic acid molecule comprises **two parts**, or subsequences. The **first part corresponds to the tag**, which identifies the tag (and the sample tag if samples are pooled). The **second part corresponds to a nucleic acid in the sample** which is mapped to a location in the genome (or to a transcript if RNA molecules are being counted)." (paragraph [0292], emphases added).

b) Tag associated sequence reads are sequence reads associated with a tag.





IPR2025-01353, IPR2025-01355
CSHL EX2002

Applicant : James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed : October 21, 2011
Page 4 of 4 of Agenda for April 21, 2015 Telephonic Examiner Interview

**Rejections under 35 U.S.C. § 112, paragraph 2 (clarity)**

In response to the Examiner's comments on pages 15-16 of the February 4, 2015 Final Office Action:

**I** Items 9, 11, 12 and 16

Applicants propose to amend claims 1, 7 and 26. A copy of the proposed amendments is attached hereto as **Exhibit D**.

**II** Items 10 and 13

Applicants would like clarification from the Examiner.

**III** Item 14

The subject specification in paragraph [0272] describes:

> "As used herein, the term 'substantially unique primers' refers to a plurality of primers, wherein each primer comprises a tag, and wherein at least 50% of the tags of the plurality are unique to the plurality of tags" (emphasis added).

Therefore, claim 9 step (a) requires that the tagged nucleic acid molecules have different tags.

**IV** Item 15

Applicants would like clarification from the Examiner. An embodiment of claim 12 is described in Figure 8.


IPR2025-01353, IPR2025-01355
CSHL EX2002

SHEET 8/8

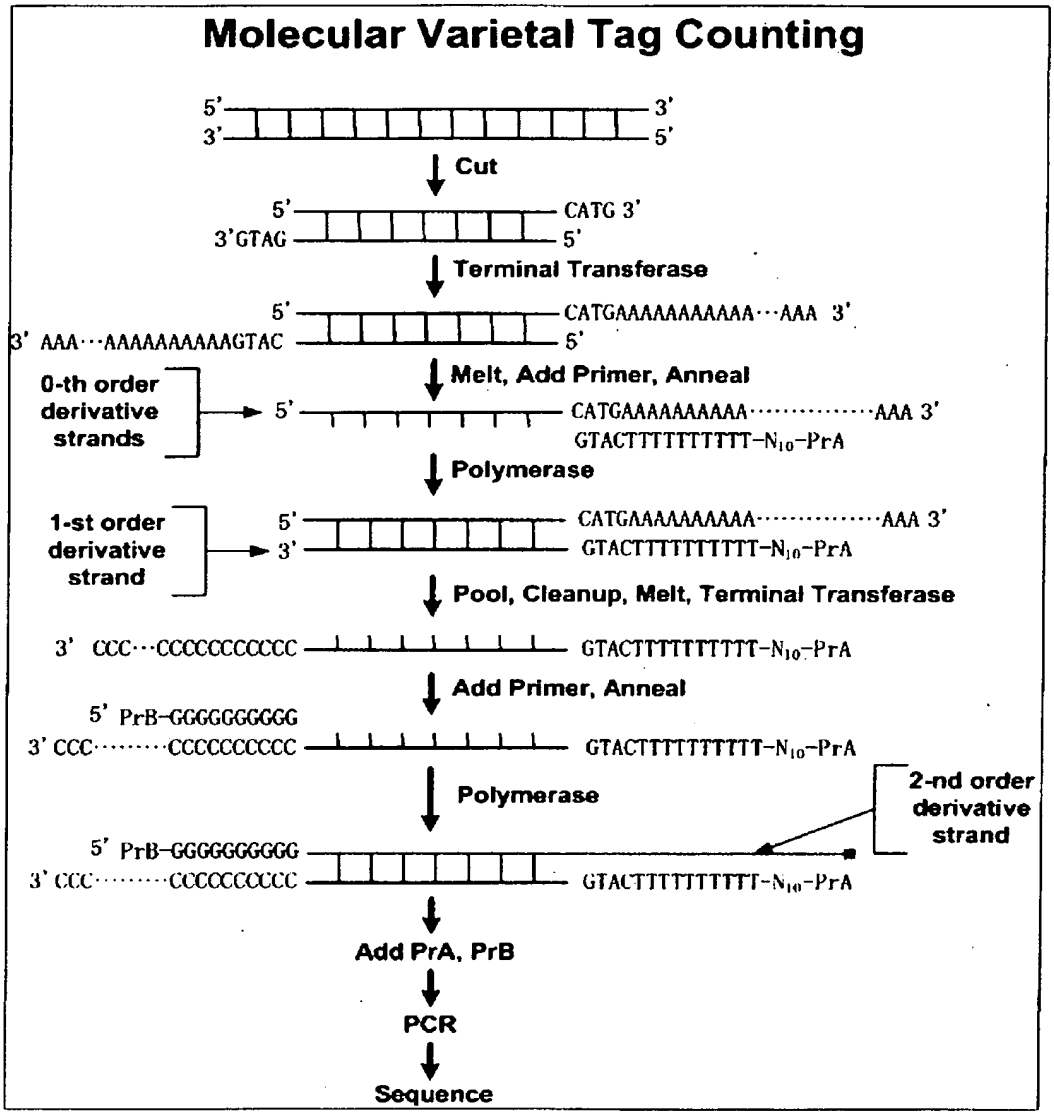



Fig. 8

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit A

12123910630 cooper dunham 04-20-2015 07:40:01 p.m. 7/23

# CLAIM 1

a) obtaining segments of genomic material



b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic material from step (a) and a substantially unique tag




Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit B

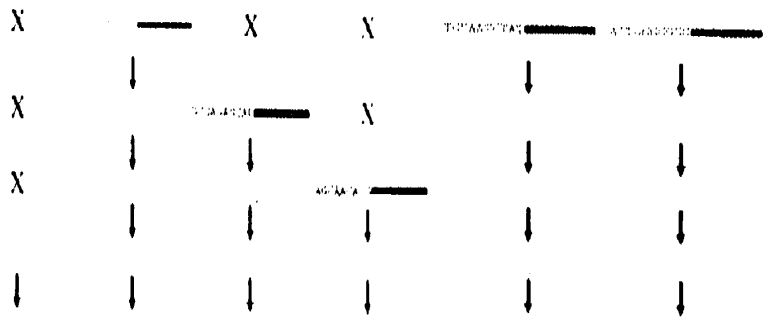
PAGE 7/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



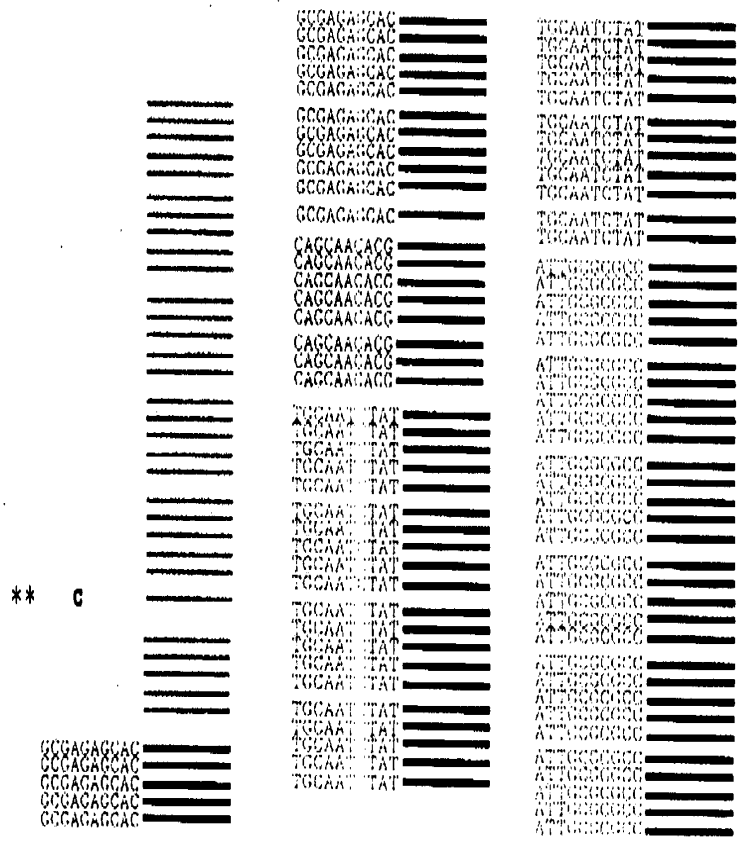
IPR2025-01353; IPR2025-01355
CSHL EX2002

8/23
04-20-2015
07:40:06 p.m.
cooper dunham
1212391063O

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR) *



d) Generating tag associated sequence reads by sequencing the product of step (c)




*Amplification distortion, refers to non-uniform amplification.

** A 1-base mismatch, highlighted in red, represents the 1/1000 frequency of sequence error in PCR

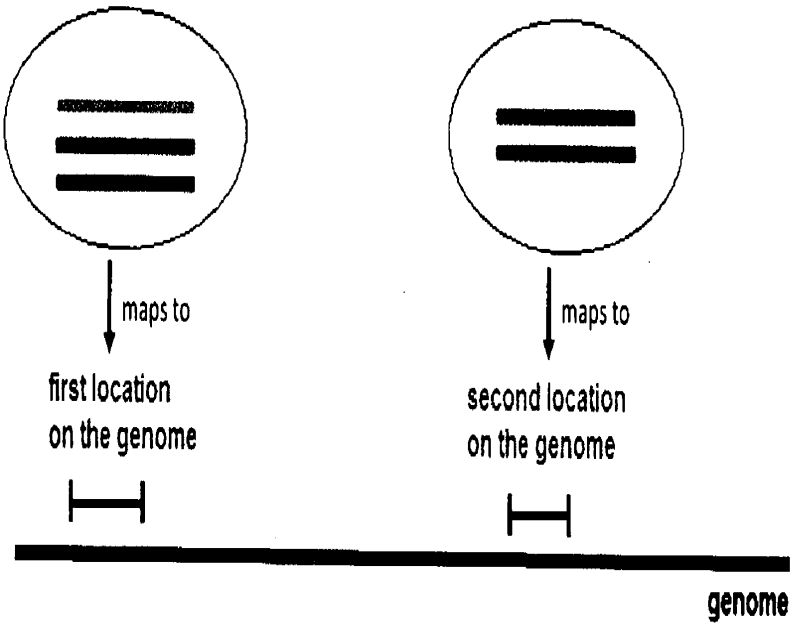
PAGE 8/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



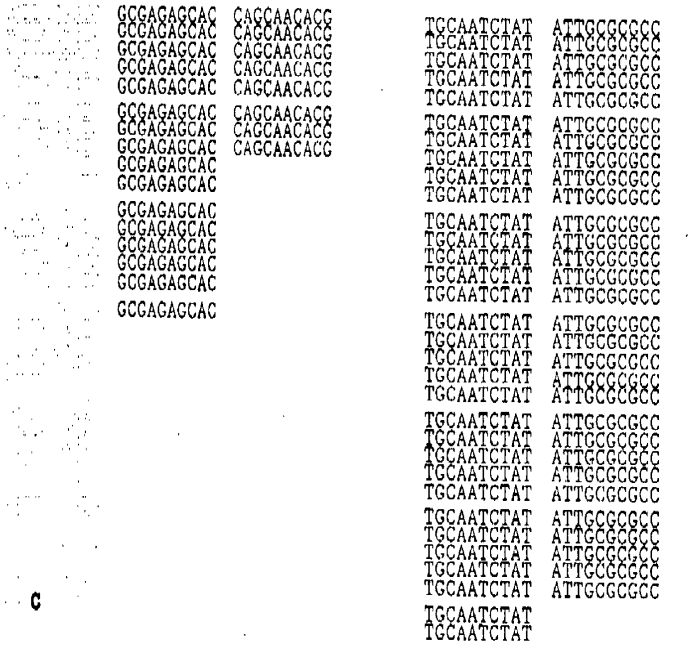
IPR2025-01353; IPR2025-01355
CSHL EX2002

07:40:17 p.m. 04-20-2015 9 /23
cooper dunham
12123910630

e) Assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read...




f) Counting the number of tagged nucleic acid molecules having a substantially unique tag that have been assigned to the same location on the genome,...




First location
Total unique tag counts = 4

Second location
Total unique tag counts = 2

PAGE 9/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24


IPR2025-01353; IPR2025-01355
CSHL EX2002

10/23
04-20-2015
07:40:27 p.m.
cooper dunham
12123910630

Table 1. Average error rates of thermostable DNA polymerases during PCR[a]

| DNA polymerase | No. of PCRs | Target (ng) | Template doublings[b] | *lacI*⁻ plaques[c] (% ± SD) | Error rate[d] (×10⁻⁶ ± SD) |
|---|---|---|---|---|---|
| *Pfu* | 10 | 24 | 9.7 | 0.42 ± 0.08 | 1.3 ± 0.2 |
| | 2 | 2 | 12.7 | 0.30 ± 0.06[e,f] | 0.7 ± 0.1[e] |
| | 2 | 0.2 | 16.0 | 0.47 ± 0.03[e,f] | 0.8 ± 0.02[e] |
| | 2 | 0.02 | 19.4 | 0.66 ± 0.03[e,f] | 1.0 ± 0.04[e] |
| *Deep Vent* | 4 | 24 | 9.7–10 | 0.9 ± 0.1 | 2.7 ± 0.2 |
| *Vent* | 6 | 24 | 8.7–10 | 0.9 ± 0.3 | 2.8 ± 0.9 |
| *Taq* | 11 | 24 | 8.7–11 | 2.7 ± 1.1 | 8.0 ± 3.9 |
| *UITma* | 2 | 24 | 9.7 | 18.8 ± 0.8[e] | 55 ± 2[e] |

[a]PCRs were performed in each manufacturers' recommended buffer (all pH 8.8) in the presence of 200 μM each dNTP and 2 mM $MgSO_4$ (*Pfu*, *Deep Vent* and *Vent*), 2 mM $MgCl_2$ (*UITma*) or 1.5 mM $MgCl_2$ (*Taq*).

[b]Template doublings (*d*) were determined using the equation $2^d$ = (amount of PCR product)/(amount of starting target). 24 ng *lacI* target corresponds to 30 ng *lacIOZα* plasmid template. The range of *d* obtained is indicated.

[c]Mutant frequencies (*mf*) were determined by dividing the total number of blue plaques (*lacI*⁻ mutants) on X-gal plates by the total number of plaques containing a functional *lacZα* sequence (blue plaques on X-gal plus IPTG plates).

[d]Error rates were calculated using the equation $ER = mf/(bp \times d)$, where *mf* is the mutation frequency, *bp* is the number of detectable sites in *lacI* (=349; 15) and *d* is the number of template doublings.

[e]± indicates range of duplicate measurements.

[f]Mutation frequencies for *Pfu* amplification of 0.02–2 ng target were normalized such that the mean mutation frequency for *Pfu* amplification of 24 ng target (assay internal control) was 0.42%.

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit C

PAGE 10/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



IPR2025-01353; IPR2025-01355
CSHL EX2002

Best Available Copy



3546–3551 Nucleic Acids Research, 1996, Vol 24, No. 18 © 1996 Oxford University Press

# PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases

**Janice Cline, Jeffery C. Braman and Holly H. Hogrefe***

Stratagene Cloning Systems, 11011 North Torrey Pines Road, La Jolla, CA 92037, USA

Received June 12, 1996; Revised and Accepted July 29, 1996

## ABSTRACT

**The replication fidelities of *Pfu*, *Taq*, *Vent*, *Deep Vent* and *UlTma* DNA polymerases were compared using a PCR-based forward mutation assay. Average error rates (mutation frequency/bp/duplication) increased as follows: *Pfu* ($1.3 \times 10^{-6}$) < *Deep Vent* ($2.7 \times 10^{-6}$) < *Vent* ($2.8 \times 10^{-6}$) < *Taq* ($8.0 \times 10^{-6}$) << exo⁻ *Pfu* and *UlTma* (~$5 \times 10^{-5}$). Buffer optimization experiments indicated that *Pfu* fidelity was highest in the presence of 2–3 mM $MgSO_4$ and 100–300 µM each dNTP and at pH 8.5–9.1. Under these conditions, the error rate of exo⁻ *Pfu* was ~40-fold higher ($5 \times 10^{-6}$) than the error rate of *Pfu*. As the reaction pH was raised from pH 8 to 9, the error rate of *Pfu* decreased ~2-fold, while the error rate of exo⁻ *Pfu* increased ~9-fold. An increase in error rate with pH has also been noted for the exonuclease-deficient DNA polymerases *Taq* and exo⁻ Klenow, suggesting that the parameters which influence replication error rates may be similar in pol I- and α-like polymerases. Finally, the fidelity of 'long PCR' DNA polymerase mixtures was examined. The error rates of a *Taq*/*Pfu* DNA polymerase mixture and a Klen*taq*/*Pfu* DNA polymerase mixture were found to be less than the error rate of *Taq* DNA polymerase, but ~3–4-fold higher than the error rate of *Pfu* DNA polymerase.**

## INTRODUCTION

The use of high fidelity DNA polymerases in the polymerase chain reaction (PCR) is essential for reducing the introduction of amplification errors in PCR products that will be cloned, sequenced and expressed. Several thermostable DNA polymerases with 3′→5′ exonuclease-dependent proofreading activity (*Pfu*, *Vent*, *Deep Vent* and *UlTma*) have been introduced for high fidelity PCR amplification (1–3). Flaman *et al.* have reported that the error rate of *Pfu* was 5- and 30-fold lower than the error rates of the proofreading enzymes *Deep Vent* and *UlTma*, respectively (4). Using several different fidelity assays, the error rate of *Pfu* has been found to be ~10-fold lower than that of the non-proofreading enzyme *Taq* (1,4,5).

The parameters which contribute to the replication fidelity of DNA polymerases need to be investigated, as very little is known about the molecular features of these enzymes which give rise to variations in replication fidelity and mutational spectra. A number of factors are thought to contribute to the overall fidelity of a DNA polymerase (reviewed in 6–8). These parameters include the tendency of a polymerase to incorporate incorrect nucleotides and the presence of an integral 3′→5′ exonuclease activity which can remove mispaired bases (proofreading activity).

The importance of proofreading activity to replication fidelity has been demonstrated for both the Klenow fragment (9) and for *Vent* polymerase (10), which exhibit 10- and 5-fold increases in error rates, respectively, when the associated 3′→5′ exonuclease activity is inactivated. The contribution of proofreading activity to DNA polymerase fidelity is also evident when the error rates of proofreading and non-proofreading enzymes are compared. Kunkel has noted that the average base substitution error rates exhibited by non-proofreading DNA polymerases range from $10^{-2}$ to $\geq 10^{-6}$, while the error rates of proofreading enzymes range from $10^{-6}$ to $10^{-7}$ (7). The parameters which contribute to error rate variations among proofreading enzymes may reflect inherent differences in 3′→5′ exonuclease activity, the tendency to discriminate mispaired versus correctly paired bases and/or the efficiency of shuttling between polymerization and proofreading modes.

Recently, mixtures of non-proofreading and proofreading DNA polymerases have been reported to synthesize higher yields of PCR product and to allow amplification of longer templates than is possible with single enzyme formulations ('long PCR') (5). The addition of a low level of a proofreading enzyme (e.g. *Pfu* DNA polymerase) to PCR reaction mixtures has been proposed to improve the performance of non-proofreading polymerases (e.g. *Taq* DNA polymerase) by correcting mismatches introduced during PCR which prevent the efficient synthesis of full-length products (5). The PCR fidelity of DNA polymerase mixtures has not yet been determined, but error rates are likely to reflect the fidelity of the component polymerases and the ratio of non-proofreading to proofreading enzyme activities.

*Pfu* DNA polymerase has been found to be useful in high fidelity amplifications (1,4) of DNA targets up to 25 kb (K. Nielson, personal communication). In this report we use the previously described *lacI* PCR mutation assay (1) to compare the error rate of *Pfu* with an expanded number of PCR polymerases, including exo⁻ *Pfu*, *Deep Vent*, *Vent*, *UlTma* and *Taq*, as well as 'long PCR' DNA polymerase mixtures. Polymerase error rates have been found to vary with buffer composition, including pH, $Mg^{2+}$ concentration and nucleotide concentration (11–13). PCR reaction conditions have been optimized with respect to fidelity for

* To whom correspondence should be addressed

PAGE 11/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



IPR2025-01353, IPR2025-01355
CSHL EX2002

Best Available Copy

12123910630 cooper dunham 07:40:55 p.m. 04-20-2015 12 /23



both *Vent* and *Taq* DNA polymerases (11). Buffer optimization studies with *Pfu* DNA polymerase were performed here to assess whether the fidelity of *Pfu* DNA polymerase could be further enhanced. Error rate comparisons between *Pfu* and exo⁻ *Pfu* are expected to illuminate the contribution of proofreading activity to the fidelity of *Pfu* DNA polymerase. Finally, PCR fidelity comparisons between *Pfu* DNA polymerase and *Pfu*-containing DNA polymerase mixtures will allow evaluation of the contribution of the predominant non-proofreading enzyme to the error rate of 'long PCR' mixtures.

## MATERIALS AND METHODS

### DNA polymerases

Cloned *Pfu*, exo⁻ *Pfu* and *Taq* DNA polymerases were prepared at Stratagene. *Deep Vent* and *Vent* polymerases were purchased from New England BioLabs. *UlTma* was obtained from Perkin-Elmer and KlentaqLA (KTLA) was provided by Wayne Barnes (Washington University School of Medicine, St Louis, MO). Except where indicated, PCR amplifications were performed in the presence of buffers supplied by the manufacturers. The KTLA PCR buffer used was buffer PC2 of Barnes (5).

### PCR fidelity assay

The fidelity of DNA replication during PCR was measured using a previously described assay (1,14). Briefly, a 1.9 kb sequence encoding *lacIOZ*α was PCR amplified as described below with oligonucleotide primers containing 5′ *Eco*RI restriction sites (1). The amplified fragments were digested with *Eco*RI, purified by gel electrophoresis and ligated into λgt10 arms. The ligation reactions were packaged and the λ phage used to infect an α-complementing *Escherichia coli* host strain. Aliquots of infected cells were plated on LB plates with top agar containing either X-gal (1 mg/ml) or X-gal plus IPTG (1.5 mM). Error rates were calculated as described in the legend to Table 1.

### PCR amplifications

Except where indicated, PCR amplifications were performed in 100 µl reaction volumes in the presence of the appropriate Tris-based buffer, using 5 U polymerase, 200 µM each dNTP, 250 ng each primer and 24 ng *lacIOZ*α target (50 ng *lacIOZ*α plasmid template). The PCR mixtures were denatured by heating at 95°C for 30 s. Thirty cycles of amplification were performed using the following conditions: 5 s at 95°C; 1 min at 55°C; 2.5 min at 72°C.

## RESULTS

### PCR fidelity of thermostable DNA polymerases

Replication fidelities of thermostable DNA polymerases were compared using a previously described assay (1) which measures the frequency of mutations introduced into the *lacI* target gene during PCR amplification. PCR amplification was performed in the presence of each enzyme's optimal PCR buffer. All other PCR parameters remained constant, including the dNTP, primer and template concentrations, the PCR cycling parameters and the number of PCR cycles performed.

*Pfu* DNA polymerase exhibited the greatest PCR fidelity, with an average error rate of $1.3 \times 10^{-6}$ mutation frequency/bp/duplication (Table 1). The *lacI* target size used in these calculations was estimated to be 349 bp, based upon the most recent analysis of *lacI*⁻ mutant DNA sequences (15). Previous error rate calculations assumed a *lacI* target size of 182 bp (1). After recalculating error rates based on a *lacI* target size of 349 bp, the mean error rate of *Pfu* DNA polymerase obtained in this study ($1.3 \times 10^{-6}$ mutation frequency/bp/duplication) was found to be similar to previous estimates obtained using an identical assay ($0.8 \times 10^{-6}$; 1) or an alternative PCR-based assay employing a p53 target gene ($\leq 1.0 \times 10^{-6}$; 4).

**Table 1.** Average error rates of thermostable DNA polymerases during PCR[a]

| DNA polymerase | No. of PCRs | Target (ng) | Template doublings[b] | *lacI*⁻ plaques[c] (% ± SD) | Error rate[d] (×10⁻⁶ ± SD) |
|---|---|---|---|---|---|
| *Pfu* | 10 | 24 | 9.7 | 0.42 ± 0.08 | 1.3 ± 0.2 |
| | 2 | 2 | 12.7 | 0.30 ± 0.06[e,f] | 0.7 ± 0.1[e] |
| | 2 | 0.2 | 16.0 | 0.47 ± 0.03[e,f] | 0.8 ± 0.02[e] |
| | 2 | 0.02 | 19.4 | 0.66 ± 0.03[e,f] | 1.0 ± 0.04[e] |
| *Deep Vent* | 4 | 24 | 9.7–10 | 0.9 ± 0.1 | 2.7 ± 0.2 |
| *Vent* | 6 | 24 | 8.7–10 | 0.9 ± 0.3 | 2.8 ± 0.9 |
| *Taq* | 11 | 24 | 8.7–11 | 2.7 ± 1.1 | 8.0 ± 3.9 |
| *UlTma* | 2 | 24 | 9.7 | 18.8 ± 0.8[e] | 55 ± 2[e] |

[a]PCRs were performed in each manufacturers' recommended buffer (all pH 8.8) in the presence of 200 µM each dNTP and 2 mM $MgSO_4$ (*Pfu*, *Deep Vent* and *Vent*), 2 mM $MgCl_2$ (*UlTma*) or 1.5 mM $MgCl_2$ (*Taq*).

[b]Template doublings (*d*) were determined using the equation $2^d$ = (amount of PCR product)/(amount of starting target). 24 ng *lacI* target corresponds to 50 ng *lacIOZ*α plasmid template. The range of *d* obtained is indicated.

[c]Mutant frequencies (*mf*) were determined by dividing the total number of blue plaques (*lacI*⁻ mutants) on X-gal plates by the total number of plaques containing a functional *lacZ*α sequence (blue plaques on X-gal plus IPTG plates).

[d]Error rates were calculated using the equation $ER = mf/(bp \times d)$, where *mf* is the mutation frequency, *bp* is the number of detectable sites in *lacI* (=349; 15) and *d* is the number of template doublings.

[e]± indicates range of duplicate measurements.

[f]Mutation frequencies for *Pfu* amplification of 0.02–2 ng target were normalized such that the mean mutation frequency for *Pfu* amplification of 24 ng target (assay internal control) was 0.42%.

PAGE 12/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24

IPR2025-01353, IPR2025-01355
463 of 819 CSHL EX2002

Best Available Copy





Average error rates of thermostable DNA polymerases were found to increase in the following order: *Pfu* ($1.3 \times 10^{-6}$) < *Deep Vent* ($2.7 \times 10^{-6}$) < *Vent* ($2.8 \times 10^{-6}$) < *Taq* ($8.0 \times 10^{-6}$) << *UlTma* ($5.5 \times 10^{-5}$). These results are in excellent agreement with the relative error rates measured by Flaman *et al.* (1), who reported that *Pfu* exhibits the greatest PCR fidelity, followed by *Deep Vent*, *Taq* and *UlTma* DNA polymerases. The relative error rates obtained here are also consistent with DGGE analyses showing that *Pfu* exhibits a lower error rate than *Vent* and *Taq* DNA polymerases (16). We found that relative error rates observed using the *lacI* screening assay were consistent from PCR reaction to PCR reaction.

The influence of template doublings (*d*) on error rate estimates of *Pfu* DNA polymerase was also examined (Table 1). Amplification reactions described above and resulting in the *Pfu* error rate of $1.3 \times 10^{-6}$ employed 24 ng *lacI* target ($10^{10}$ copies). Approximately 10 doublings were observed in 30 PCR cycles. When the input *lacI* target DNA was decreased from $10^{10}$ copies (24 ng) to $10^{7}$ copies (0.02 ng), the number of template doublings increased from 9.7 (~900-fold amplification) to 19.4 (~700 000-fold amplification) after 30 cycles of PCR. The error rate of *Pfu* DNA polymerase varied from 0.7 to $1.3 \times 10^{-6}$ over the 1000-fold range of DNA target concentrations tested. Flaman *et al.* have also reported that polymerase error rates do not appear to be significantly influenced by the number of template doublings (1).

### Optimization of the PCR fidelity of *Pfu*

We attempted to further improve the fidelity of *Pfu* by optimizing PCR reaction conditions. PCR error rates were measured at varying concentrations of $MgSO_4$ (Fig. 1) and dNTPs (Fig. 2) and at varying pHs (Fig. 3). The indicated pH values are those measured at room temperature. Where noted, the pH of Tris buffers at elevated temperatures was estimated using the formula: $pH_T = pH_{25°C} + [(T°C - 25°C) \times (-0.03\ pH\ U/°C)]$ (where $T$ is the reaction temperature; 17). The lowest error rates for *Pfu* were observed when PCR amplifications were performed in the presence of 2–3 mM $MgSO_4$, 100–300 µM each dNTP and in a pH range between 8.5 and 9.1 (pH ~7.1–7.7 at 72°C). These conditions have been found to give optimal yield of PCR product as well (18).

In the presence of 1 mM $MgSO_4$ and 800 µM total dNTPs, the error rate of *Pfu* was ~3-fold higher than when PCR amplifications were performed in 2 mM $MgSO_4$ at the same dNTP concentration (Fig. 1). The error rate did not vary significantly as the $MgSO_4$ concentration was increased from 2 to 10 mM (~1.2–9.2 mM free $Mg^{2+}$). The error rate of *Vent* polymerase has also been shown to decrease significantly between 0.5 and 2 mM $MgSO_4$ in the presence of 2 mM total dNTPs and thereafter remains constant with increasing concentrations of free $Mg^{2+}$ (11). These results are in contrast to those reported for *Taq*, in which error rates are lowest at equimolar concentrations of $MgCl_2$ and dNTPs (1 mM) and increase with increasing concentration of free $Mg^{2+}$ (12). Error rate variations of *Pfu* and *Vent* likely reflect the $Mg^{2+}$ dependency of both proofreading and polymerase activities.

In Figure 2, the error rate of *Pfu* was found to increase 2.4-fold as the total dNTP concentration was raised from 0.4 to 4 mM in the presence of a constant amount of free $Mg^{2+}$ (~1.2 mM). These results are consistent with the observations of Clayton *et al.* (19), who report that high concentrations of dNTPs diminish the proofreading ability of exonuclease-proficient polymerases by increasing the efficiency of mispair extension. It is likely that the fidelity of *Pfu* DNA polymerase could be further increased by reducing the total dNTP concentration below 0.4 mM total dNTPs. However, using lower dNTP concentrations to increase the fidelity of PCR amplification reactions is not practical, as PCR product yields decrease significantly below 0.4 mM total dNTPs.

In Figure 3 (inset), the error rate of *Pfu* was measured as a function of pH. The error rate of *Pfu* was found to decrease 4-fold between pH 7.5 and 8.5 in the presence of 2 mM $MgSO_4$ and 0.8 mM total dNTPs. *Vent* polymerase has also been reported to

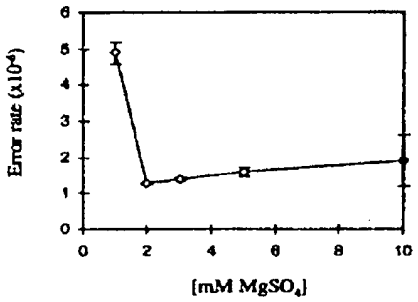



Figure 1. Variation of the PCR error rates of *Pfu* DNA polymerase with $MgSO_4$ concentration. PCR amplification was performed in buffer containing 20 mM Tris–HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 µg/ml BSA, 200 µM each dNTP and varying concentrations of $MgSO_4$ (1–10 mM). Error rates shown are the average (± range) values obtained from two independent PCR amplifications.

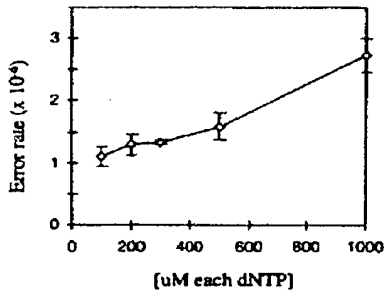



Figure 2. Variation of the PCR error rates of *Pfu* DNA polymerase with dNTP concentration. PCR amplification was performed in buffer containing 20 mM Tris–HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100µg/ml BSA and varying concentrations of dNTPs (100–1000 µM each). The $MgSO_4$ concentration of the PCR mixtures was adjusted to give a constant free $Mg^{2+}$ concentration (1.2 mM) at each dNTP concentration. 1.6 mM $MgSO_4$/ 0.4 mM total dNTPs; 2 mM $MgSO_4$/0.8 mM total dNTPs; 2.4 mM $MgSO_4$/ 1.2 mM total dNTPs; 3.2 mM $MgSO_4$/2 mM total dNTPs; 5.2 mM $MgSO_4$/4 mM total dNTPs. Error rates shown were normalized such that the mean mutation frequency for *Pfu* amplifications with 0.8 mM total dNTPs (assay internal control) was $1.3 \times 10^{-6}$ mutation frequency/bp/duplication. The average error rates (± range) from two independent PCR amplifications are shown.

PAGE 13/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



IPR2025-01353, IPR2025-01355
CSHL EX2002

Best Available Copy

12123910630 cooper dunham 07:41:31 p.m. 04-20-2015 14 /23



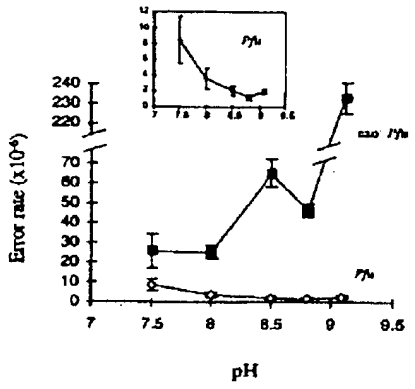



Figure 3. Variation of the PCR error rates of *Pfu* and exo⁻ *Pfu* DNA polymerases with pH. PCR amplification was performed in 20 mM Tris–HCl buffers whose pH values ranged from 7.5 to 9.1 (pH at 25°C). In addition, the buffer contained 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 μg/ml BSA and 200 μM each dNTP. The average error rate of *Pfu* (open diamonds) is shown in comparison with exo⁻ *Pfu* (filled diamonds) and in the accompanying inset. Error rates shown are the average (± range) values obtained from two independent PCR amplifications.

exhibit a significant decrease in error rate as the pH is increased from 7 to 8 in the presence of 2 mM $MgSO_4$ (11). For *Taq* DNA polymerase, a 2-fold increase in error rate was observed when the reaction pH was raised from 8 to 9 (11). Eckert and Kunkel have also reported that the base substitution and frameshift error rates of *Taq* (12) and exo⁻ Klenow (13) increase >10-fold as the reaction pH is raised from ~6.5 to 9.5 (25°C estimates of pH from 12,13).

### pH dependency of the fidelity of *Pfu* and exo⁻ *Pfu*

The error rates of *Pfu* and exo⁻ *Pfu* were compared to assess the contribution of 3′→5′ exonuclease activity to fidelity. In the presence of *Pfu* PCR buffer (2 mM $MgSO_4$, 200 μM each dNTP, pH 8.8), exo⁻ *Pfu* exhibited an error rate of $4.7 \times 10^{-5}$ mutation frequency/bp/duplication, which is ~40-fold higher than that determined for exonuclease-proficient *Pfu*.

Figure 3 shows the error rate variation of exo⁻ *Pfu* and *Pfu* as a function of pH. Exo⁻ *Pfu* shows a dramatic increase in error rate (~9-fold) as the reaction pH is raised from pH 8 to 9.1 (or from 6.6 to 7.7 at 72°C). In contrast to exo⁻ *Pfu*, the error rate of *Pfu* decreased ~2-fold in this pH range. Presumably, the fidelity of *Pfu* is maintained at high pH (pH 9) by enhanced proofreading activity, which accompanies the dramatic increase in nucleotide misincorporation occurring between pH 8 and 9.1 (identified using exo⁻ *Pfu*). These results and those reported by others for *Taq* and exo⁻ Klenow (11–13) indicate that the error rates of exonuclease-deficient enzymes, *Taq*, exo⁻ Klenow and exo⁻ *Pfu*, are similarly increased by pH. The significance of the apparent biphasic relationship between error rate and pH is currently under investigation.

### PCR fidelity of 'long PCR' DNA polymerase mixtures

The fidelities of *Pfu* and *Taq* DNA polymerases were compared with the fidelities of two *Pfu*-containing DNA polymerase mixtures (Table 2). A *Taq*/*Pfu* (16 U:1 U) mixture was prepared and shown to amplify DNA targets >30 kb (data not shown). The *Taq*/*Pfu* mixture exhibited an average error rate of $5.6 \times 10^{-6}$ mutation frequency/bp/duplication when amplifications were performed in *Taq* PCR buffer. The mean error rate of the *Taq*/*Pfu* mixture was 30% lower than the mean error rate of *Taq* DNA polymerase when amplifications were conducted in *Taq* PCR buffer. When compared with the error rate of *Pfu* DNA polymerase in the same buffer system, the error rate of the *Taq*/*Pfu* mixture was found to be 6-fold higher.

Similar observations were made for a second 'long PCR' mixture, KTLA, which consists of Klentaq (N-terminally truncated *Taq*) and *Pfu* DNA polymerases (5). When PCR amplifications were conducted as described in this report, KTLA exhibited a mean error rate of $3.9 \times 10^{-6}$ mutation frequency/bp/duplication, which was 3-fold higher than the error rate of *Pfu* DNA polymerase (Table 2). When PCR conditions from Barnes (5) were used (Table 2, condition 2), KTLA exhibited a mean error rate ($9.4 \times 10^{-6}$) which was 4-fold higher than the error rate of *Pfu* DNA polymerase assayed under identical conditions.

Table 2. Error rate comparisons of DNA polymerases and 'long PCR' DNA polymerase mixtures

| PCR condition[a] | DNA polymerase | No. of PCRs | Template doublings[b] | Error rate (× $10^{-6}$ ± SD)[b] |
|---|---|---|---|---|
| 1 | *Pfu* | 10 | 8.7–9.7 | 1.3 ± 0.2 |
| | *Taq* | 11 | 8.7–11 | 8.0 ± 3.9 |
| | *Taq*/*Pfu* (16:1)[c] | | | |
| | *Taq* buffer | 8 | 9.7–10 | 5.6 ± 1.6 |
| | *Pfu* buffer | 11 | 9.7–11 | 7.6 ± 1.2 |
| | KTLA | 2 | 9.7 | 3.9 ± 0.1[d] |
| 2 | *Pfu* | 2 | 8.1 | 2.3 ± 0.2[d] |
| | KTLA | 2 | 8.1 | 9.4 ± 0.9[d] |

[a]PCR condition 1 is described in Materials and Methods, PCR amplification. PCR condition 2 is from Barnes (5) and differs in the following respects. PCR amplifications were performed on a Robocycler 40 (Stratagene) using thin-walled PCR tubes and 7.2 ng target (15 ng *lacI*/*Z*α plasmid). Sixteen cycles of amplification were performed using the following conditions: 30 s at 99°C, 30 s at 67°C, 3 min at 68°C.
[b]Defined in the legend to Table 1.
[c]The *Taq*:*Pfu* mixture consists of *Taq* (5 U/μl) and *Pfu* (0.31 U/μl) DNA polymerases.
[d]± indicates range of duplicate measurements.

PAGE 14/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24

IPR2025-01353, IPR2025-01355
465 of 819 CSHL EX2002

Best Available Copy





## DISCUSSION

The intrinsic properties of thermostable DNA polymerases which contribute to variation in PCR fidelity are not fully understood. In general, enzymes which possess an associated 3′→5′ exonuclease-dependent proofreading activity are thought to exhibit higher replication fidelity than non-proofreading DNA polymerases (7). Variation in fidelity among proofreading enzymes, such as *Pfu*, *Vent* and *Deep Vent*, may reflect differences in the rate of mispair excision, the level of discrimination between mispaired and correctly paired bases, the rate of mispair extension and/or the efficiency of shuttling the 3′ primer terminus between the polymerase and exonuclease active sites.

The contribution of 3′→5′ exonuclease activity to the PCR fidelity of *Pfu* was demonstrated directly by comparing the error rates of *Pfu* and exo⁻ *Pfu*. The error rate of exo⁻ *Pfu* was found to be 7-fold higher than the error rate of exo⁺ *Pfu* at pH 8.0 and 40-fold higher at pH 8.8 (*Pfu* PCR buffer).

Despite the importance of proofreading activity to the fidelity of *Pfu* and *Vent* (10), the presence of 3′→5′ exonuclease activity does not necessarily guarantee high fidelity DNA synthesis, as illustrated by *UlTma* DNA polymerase. The poor fidelity of *UlTma* DNA polymerase may be related to the relatively low level of 3′→5′ exonuclease activity exhibited by this enzyme. In a preliminary analysis of exonuclease activity, *UlTma* was found to exhibit significantly lower levels of 3′→5′ exonuclease activity than *Pfu*, *Deep Vent* and *Vent* DNA polymerases (A. Lovejoy, personal communication). However, other parameters are likely to contribute to low fidelity, since *UlTma*, an N-terminally deleted version of *Thermatoga maritima* DNA polymerase (20), exhibits an ~7-fold higher error rate than *Taq*, which is completely devoid of proofreading activity.

In the absence of proofreading activity, a DNA polymerase like *Taq* is thought to accomplish high fidelity DNA synthesis by inefficient incorporation of non-complementary dNTPs and a reduced tendency to extend from mismatched 3′ primer termini. Huang *et al.* (21) have shown that, with the exception of C–T mispairs, *Taq* polymerase exhibits ~100–1000-fold greater discrimination against mispair extension, as compared with avian myeloblastosis and HIV-1 reverse transcriptases, which extend most mispairs permissively. The rate at which DNA polymerases extend from mispaired 3′ primer termini, however, does not contribute to the actual fidelity of non-proofreading enzymes. The mismatch extension rate only contributes to fidelity in the sense that if the mismatch is extended inefficiently, the DNA will not be replicated to completion and the mutation will not be scored. Therefore, the mispair extension rate influences the number of detected mutants, rather than reflecting the inherent fidelity of a non-proofreading DNA polymerase.

The observed 6-fold difference in error rate between *Taq* ($8 \times 10^{-6}$) and exo⁻ *Pfu* ($4.7 \times 10^{-5}$) suggests that the misincorporation and/or misextension rates of *Pfu* (as measured with exo⁻ *Pfu*) are significantly higher than those of *Taq*. Apparently, a lower degree of discrimination against misinsertion or mispair extension errors can be tolerated when an associated proofreading activity is present, as is the case with exonuclease-proficient *Pfu*.

Further fidelity measurements with exo⁻ *Pfu* revealed that the fidelity of dNTP incorporation was significantly influenced by the pH of the PCR buffer. The error rate increased by ~9-fold as the pH was raised from pH 6.6 to 7.7 (pH at 72°C). The error rates of both *Taq* (12) and exo⁻ Klenow (13) increase similarly at higher pH. Eckert and Kunkel have attributed the lower fidelity of exo⁻ Klenow at high pH to an increase in both nucleotide misinsertion and mispair extension (13).

It is tempting to speculate that the lower fidelity of exo⁻ *Pfu* at high pH may also reflect increased misinsertion and mispair extension, analogous to the observations made for exo⁻ Klenow (13). If so, it would suggest that the parameters which contribute to fidelity are similar, despite the structural differences which are thought to exist between the α-like (exo⁻ *Pfu*; 22) and pol I-like (exo⁻ Klenow and *Taq*) DNA polymerases. For example, the observed variation in error rates with pH suggests that an active site histidine residue may play a role in fidelity, possibly in the discrimination of mismatched 3′ primer termini. Alternatively, protonation of the primer, template or substrate dNTP may enhance error discrimination (13). Finally, pol I- and α-like polymerases may undergo a similar conformational change at low pH which may alter template binding properties, thereby improving error discrimination. Such a mechanism was proposed for exo⁻ Klenow by Eckert and Kunkel (13) and was supported by additional data showing that lower error rates at low pH were accompanied by an increase in polymerase processivity.

The relative error rates for *Pfu*, *Vent* and *Taq* were found to parallel the terminal transferase activities of DNA polymerases. Hu (23) has compared the tendency of DNA polymerases to catalyze the addition of non-template-directed bases to the 3′-end of a DNA fragment (terminal transferase activity). Terminal transferase activity is high in *Taq* but low (Klenow and *Vent*) or absent (*Pfu*, *T4* and *T7*) in proofreading enzymes, which presumably edit the misextended base. The absence of terminal transferase activity appears to correlate with high fidelity. Fidelity measurements compiled by Cha and Thilly show that the error rates of *Pfu*, *T4* and *T7* DNA polymerases are lower than the error rates of *Vent* and Klenow (16). Thus, the parameters which give rise to terminal transferase activity may be similar to those which contribute to lower fidelity. The lower error rate and lack of terminal transferase activity for *Pfu* (as compared with *Vent*) may be the result of a reduced tendency of *Pfu* to incorporate a mismatch or a base opposite an abasic site. Alternatively, *Pfu* may excise misincorporated bases more readily or shuttle between the exonuclease and polymerase active sites more efficiently.

Finally, fidelity comparisons with *Pfu*-containing 'long PCR' DNA polymerase mixtures have shown that the error rate of mixtures appears to be intermediate between the error rate of *Pfu* and the non-proofreading DNA polymerase. The lower error rate of a *Taq*/*Pfu* mixture, as compared with *Taq* alone, suggests that *Pfu* is editing a certain percentage of mismatches that have been introduced by *Taq* during the PCR process. Editing may occur at the 3′-terminus after *Taq* has introduced a mismatch and dissociated from the incomplete PCR product (5). In the absence of *Pfu*, *Taq* presumably extends some of these putative stalling mismatches during the course of the PCR process; otherwise the mutations would not be scored in the *lacI*⁻ screening assay and there would be no apparent difference in error rate between *Taq* and the *Taq*/*Pfu* mixture. *Pfu* may also reduce the overall error rate of *Taq* DNA polymerase by degrading *Taq*-generated duplex DNA containing mismatches and resynthesizing the correct sequence.

Although the error rate of the *Taq*/*Pfu* mixture is somewhat lower than the error rate of *Taq* alone, it is still 4–6-fold higher than the error rate of *Pfu* alone (Table 2). These results indicate that the majority of PCR products are synthesized by *Taq*. This

PAGE 15/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



IPR2025-01353, IPR2025-01355
CSHL EX2002

Best Available Copy

12123910630 cooper dunham 07:42:16 p.m. 04-20-2015 16 /23



result is not surprising, since *Taq* is present in this particular mixture at a 16-fold higher polymerase unit concentration than *Pfu* DNA polymerase. Hence, the misincorporation rate of *Taq* DNA polymerase contributes significantly to the error rate of *Taq*/*Pfu* DNA polymerase mixtures.

KTLA, a 'long PCR' mixture of Klentaq and *Pfu* DNA polymerases, was also found to exhibit an error rate significantly higher than the error rate of *Pfu*. Our results are inconsistent with the results of Barnes (5), who has compared the error rates of *Pfu*, Klentaq and KTLA-64 (~640 U Klentaq:1 U *Pfu*) using a similar PCR forward mutation assay based on the mutational target gene *lacZ*. Barnes reported that the error rate of the KTLA mixture was 2-fold lower than the error rate of *Pfu* DNA polymerase (5). There are several differences between the Barnes assay and the assay performed here, including PCR amplification conditions (see Table 2 legend), number of clones screened [500–4200 clones/1 PCR in Barnes (5) versus 10 000–50 000 clones/PCR/4 PCRs in this study] and the mutational target gene used (*lacZ* versus *lacI*), as well as possible unknown variations in the KTLA mixtures. The results in Table 2 demonstrate that differences in the PCR amplification conditions employed are not likely to contribute to the differences in relative error rates observed in the two studies. Fidelity analyses of additional DNA polymerase mixtures are currently under way to help elucidate the role of component enzymes and buffer composition in the fidelity of 'long PCR' amplifications.

## REFERENCES


1 Lundberg,K.S., Shoemaker,D.D., Adams,M.W.W., Short,J.M., Sorge,J.A. and Mathur,E.J. (1991) *Gene*, **108**, 1–6.
2 Kong,H., Kucera,R.B. and Jack,W.E. (1993) *J. Biol. Chem.*, **268**, 1965–1975.
3 Abramson,R.D. (1995) In Innis,M.A., Gelfand,D.H. and Sninsky,J.J. (eds), *PCR Strategies*. Academic Press, San Diego, CA.
4 Flaman,J.-M., Frebourg,T., Moreau,V., Charbonnier,F., Martin,C., Ishioka,C., Friend,S.H. and Iggo,R. (1994) *Nucleic Acids Res.*, **22**, 3259–3260.
5 Barnes,W.M. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 2216–2220.
6 Echols,H. and Goodman,M.F. (1991) *Annu. Rev. Biochem.*, **60**, 477–511.
7 Kunkel,T.A. (1992) *J. Biol. Chem.*, **267**, 18251–18254.
8 Goodman,M.F., Creighton,S., Bloom,L.B. and Petruska,J. (1993) *Crit. Rev. Biochem. Mol. Biol.*, **28**, 83–126.
9 Kunkel,T.A. (1988) *Cell*, **53**, 837–840.
10 Mantila,P., Korpela,J., Tenkanen,T. and Pitkanen,K. (1991) *Nucleic Acids Res.*, **19**, 4967–4973.
11 Lung,L.L., Keohavong,P., Dias,C. and Thilly,W.G. (1991) *PCR Methods Applicat.*, **1**, 63–69.
12 Eckert,K.A. and Kunkel,T.A. (1990) *Nucleic Acids Res.*, **18**, 3739–3744.
13 Eckert,K.A. and Kunkel,T.A. (1993) *J. Biol. Chem.*, **268**, 13462–13471.
14 Cline,J., Braman,J. and Kretz,K. (1995) *Strategies Mol. Biol.*, **8**, 24–25.
15 Provost,G.S., Kretz,P.L., Hamner,R.T., Matthews,C.D., Rogers,B.J., Lundberg,K.S., Dycaico,M.J. and Short,J.M. (1993) *Mutat. Res.*, **288**, 133–149.
16 Cha,R.S. and Thilly,W.G. (1995) In Dieffenbach,C.W. and Dveksler,G.S. (eds), *PCR Primer*. Cold Spring Harbor Laboratory Press, Plainview, NY.
17 Good,N.E. (1986) *Biochemistry*, **5**, 467–476.
18 Nielson,K.B., Costa,G.L. and Braman,J. (1996) *Strategies Mol. Biol.*, **9**, 24–25.
19 Clayton,L.K., Goodman,M.F., Branscomb,E.W. and Galas,D.J. (1979) *J. Biol. Chem.*, **254**, 1902–1912.
20 Innis,M.A., Myambo,K.B., Gelfand,D.H. and Brow,M.A.D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 9436–9440.
21 Huang,M.-M., Arnheim,N. and Goodman,M.F. (1992) *Nucleic Acids Res.*, **20**, 4567–4573.
22 Mathur,E.J., Adams,M.W.W., Callen,W.N. and Cline,J.M. (1991) *Nucleic Acids Res.*, **19**, 6952.
23 Hu,G. (1993) *DNA Cell Biol.*, **12**, 763–770.


PAGE 16/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



IPR2025-01353, IPR2025-01355
CSHL EX2002

RECEIVED
CENTRAL FAX CENTER
APR 20 2015

Applicant : James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed : October 21, 2011
Page 1 of 7 of Proposed claim amendments for April 21, 2015 Telephonic Examiner Interview

Docket No. 5981/81503-A/GJG/MW

**PROPOSED AMENDMENTS TO THE CLAIMS**

**Listing of Claims**

1. (Currently Amended) A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic material from step (a) and a substantially unique tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to [[a]]one segment of the segments of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.



PAGE 17/23 * RCVD AT 4/20/2015 8:38:46 PM [Eastern Daylight Time] * SVR:W-PTOFAX-003/17 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):04-24



IPR2025-01353, IPR2025-01355
CSHL EX2002

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Previously Presented) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising i) summing the counts obtained in step (f) for locations which comprise a region of the genome~~, and comparing the sum to a sum~~; ii) summing the counts obtained from a reference sample for the same region of the genome~~,~~; ~~and~~ iii) comparing the sums obtained in step (i) and (ii), thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Previously Presented) A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;



IPR2025-01353, IPR2025-01355
CSHL EX2002

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. (Cancelled)

11. (Previously Presented) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:





IPR2025-01353, IPR2025-01355
CSHL EX2002

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Previously Presented) The method of claim 1, wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands





IPR2025-01353, IPR2025-01355
CSHL EX2002

added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein adding [[a]]the polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to ~~sequencing in~~ step (d).

25. (Cancelled)

26. (Currently Amended) The method of claim 1, wherein the substantially unique tags ~~sequenced~~ further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag



IPR2025-01353, IPR2025-01355
CSHL EX2002

prior to PCR amplification or prior to said sequencing of the product of step (c).

28. (Currently Amended) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.





IPR2025-01353, IPR2025-01355
CSHL EX2002

RECEIVED
CENTRAL FAX CENTER

APR 20 2015

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 02/04/2015

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER | |
|---|---|
| LU, FRANK WEI MIN | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 1634 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 02/04/2015 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

PTOL-90A (Rev. 04/07)

RECEIVED
CENTRAL FAX CENTER
APR 20 2015

Applicant : James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed : October 21, 2011


**AGENDA FOR TELEPHONIC EXAMINER INTERVIEW**

**Rejections under 35 U.S.C. § 112, paragraph 1 (enablement)**

I **A person having ordinary skill in the art (POSA) would be enabled to amplify and sequence tagged nucleic acid molecules.**

In response to the Examiner's comments on pages 11-12 of the February 4, 2015 Final Office Action:

The Examiner's assertion that universal primers cannot be used with the claimed method is incorrect. Either specific or universal primers capable of hybridizing to sequences of the genomic material can be used in the claimed method.

Figure 8 of the subject specification, described in paragraphs [0040], [0041], [0042], [303], and [304] illustrates one embodiment of the method described in step (c). Figure 8 shows that following addition of terminal transferase, a poly-A tail is added to the ends of the segments of the genomic material to generate $0^{th}$ order derivative strands. A primer (PrA) comprising a tag ($N_{10}$) hybridizes to the poly-A tail added to the segment of the genomic material. Thus, in Figure 8, the hybridization of primer (PrA) with the poly-A tail is in effect analogous to the effect of using a universal primer. Yet, the method still works. A copy of Figure 8 is attached hereto as **Exhibit A**.

A schematic example of the claim is shown in **Exhibit B**.


IPR2025-01353, IPR2025-01355
CSHL EX2002

Best Available Copy



RECEIVED
CENTRAL FAX CENTER
APR 20 2015

Applicant : James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed : October 21, 2011
Page 1 of 7 of Proposed claim amendments for April 21, 2015 Telephonic Examiner Interview

Docket No. 5981/81503-A/GJG/MW

**PROPOSED AMENDMENTS TO THE CLAIMS**

**Listing of Claims**

1. (Currently Amended) A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic material from step (a) and a substantially unique tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to [[a]]one segment of the segments of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.



PAGE 17/23 * RCVD AT 4/20/2015 8:56:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):10-30



IPR2025-01353, IPR2025-01355
CSHL EX2002

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Previously Presented) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (b) further comprises


IPR2025-01353, IPR2025-01355
CSHL EX2002

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising i) summing the counts obtained in step (f) for locations which comprise a region of the genome, ~~and comparing the sum to a sum~~; ii) summing the counts obtained from a reference sample for the same region of the genome; ~~and iii) comparing the sums obtained in step (i) and (ii),~~ thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Previously Presented) A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;



IPR2025-01353, IPR2025-01355
CSHL EX2002

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. (Cancelled)

11. (Previously Presented) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:





IPR2025-01353, IPR2025-01355
CSHL EX2002

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Previously Presented) The method of claim 1, wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands





IPR2025-01353, IPR2025-01355
CSHL EX2002

added in step (a) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein adding [[a]]the polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to ~~sequencing in~~ step (d).

25. (Cancelled)

26. (Currently Amended) The method of claim 1, wherein the substantially unique tags ~~sequences~~ further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag

prior to PCR amplification or prior to said sequencing of the product of step (c).

28. (Currently Amended) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.





IPR2025-01353, IPR2025-01355
CSHL EX2002

II **A POSA could perform the claimed methods even if the amplification product of step (c) contained mutant tagged nucleic acid molecules, as alleged by the Examiner.**

In response to the Examiner's comments on pages 12-14 of the February 4, 2015 Final Office Action:

a) While mutations can occur in PCR, mutations were known to occur at a rate of $10^{-5}$ to $10^{-6}$ mutations/base pair (see i.e. Table 1 of Cline et al., Nucleic Acids Research 1996; attached hereto as **Exhibit C**).

b) Even if a mutation was introduced in 1 out of 100,000 base pairs, it was well understood in the art and well within the skill of a POSA to compensate for such mutation.

c) A POSA could 1) recognize an aberrant "mutant" subsequence during the mapping of step (e); and 2) discard the aberrant "mutant" subsequence during the counting of step (f).

d) The subject specification provides guidance for a POSA to account for unwanted mutations which might unexpectedly and rarely occur. Paragraph [0311] explains that:

> "tagged nucleic acid molecules that differ by a single nucleotide can (optionally) be ignored when counting".





IPR2025-01353, IPR2025-01355
CSHL EX2002

**III A "tag associated sequence read" is defined to require sequencing of the tag and at least a portion of the nucleic acid.**

In response to the Examiner's comments on page 14 of the February 4, 2015 Final Office Action:

a) "Sequence reads" when referring to tagged nucleic acid molecules are defined in the specification as containing (1) a subsequence of a tag, and (2) a subsequence of a nucleic acid in the sample:

> "In the disclosed method, each sequence read generated from a tagged nucleic acid molecule comprises **two parts**, or subsequences. The **first part corresponds to the tag**, which identifies the tag (and the sample tag if samples are pooled). The **second part corresponds to a nucleic acid in the sample** which is mapped to a location in the genome (or to a transcript if RNA molecules are being counted)." (paragraph [0292], emphases added).

b) Tag associated sequence reads are sequence reads associated with a tag.





IPR2025-01353, IPR2025-01355
CSHL EX2002

Applicant : James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed : October 21, 2011
Page 4 of 4 of Agenda for April 21, 2015 Telephonic Examiner Interview

**Rejections under 35 U.S.C. § 112, paragraph 2 (clarity)**

In response to the Examiner's comments on pages 15-16 of the February 4, 2015 Final Office Action:

**I** Items 9, 11, 12 and 16

Applicants propose to amend claims 1, 7 and 26. A copy of the proposed amendments is attached hereto as **Exhibit D**.

**II** Items 10 and 13

Applicants would like clarification from the Examiner.

**III** Item 14

The subject specification in paragraph [0272] describes:

> "As used herein, the term 'substantially unique primers' refers to a plurality of primers, wherein each primer comprises a tag, and wherein at least 50% of the tags of the plurality are unique to the plurality of tags" (emphasis added).

Therefore, claim 9 step (a) requires that the tagged nucleic acid molecules have different tags.

**IV** Item 15

Applicants would like clarification from the Examiner. An embodiment of claim 12 is described in Figure 8.

07:59:18 p.m. 04-20-2015 7/23
12123910630 cooper dunham

# CLAIM 1

a) obtaining segments of genomic material



b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic material from step (a) and a substantially unique tag



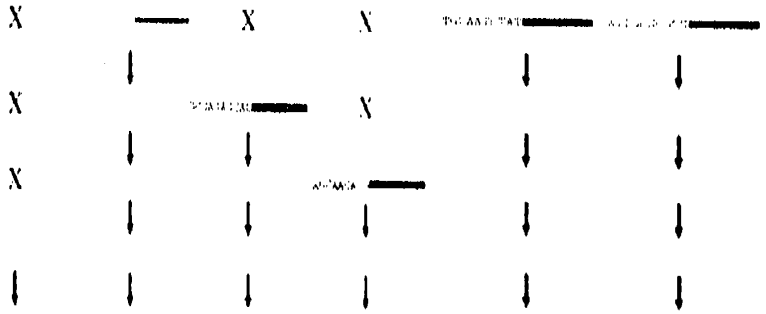

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit B

PAGE 7/23 * RCVD AT 4/20/2015 8:58:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):10-30

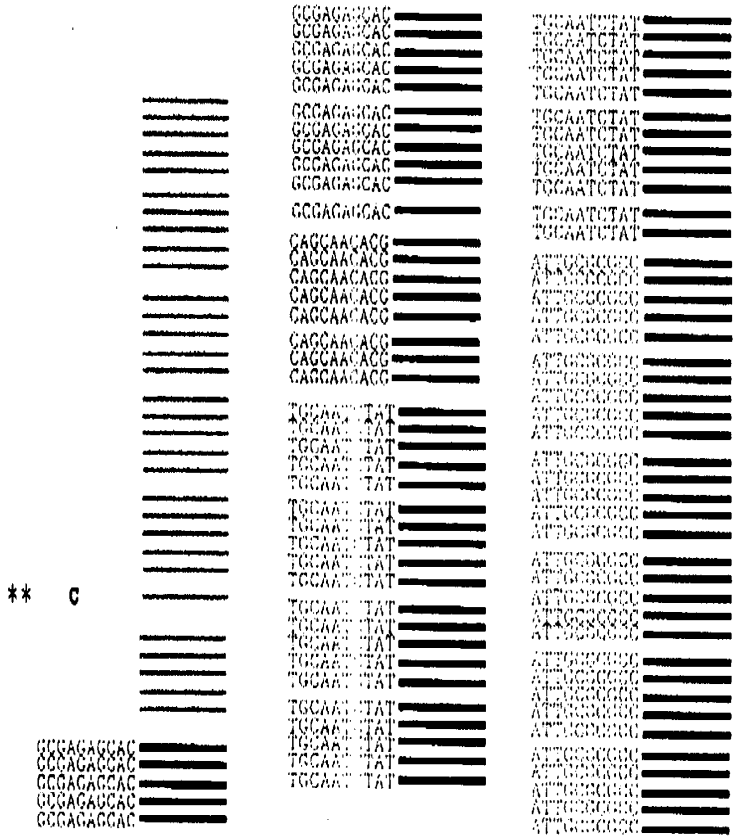


IPR2025-01353; IPR2025-01355
CSHL EX2002

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR) *



d) Generating tag associated sequence reads by sequencing the product of step (c)




*Amplification distortion, refers to non-uniform amplification.

** A 1-base mismatch, highlighted in red, represents the 1/1000 frequency of sequence error in PCR

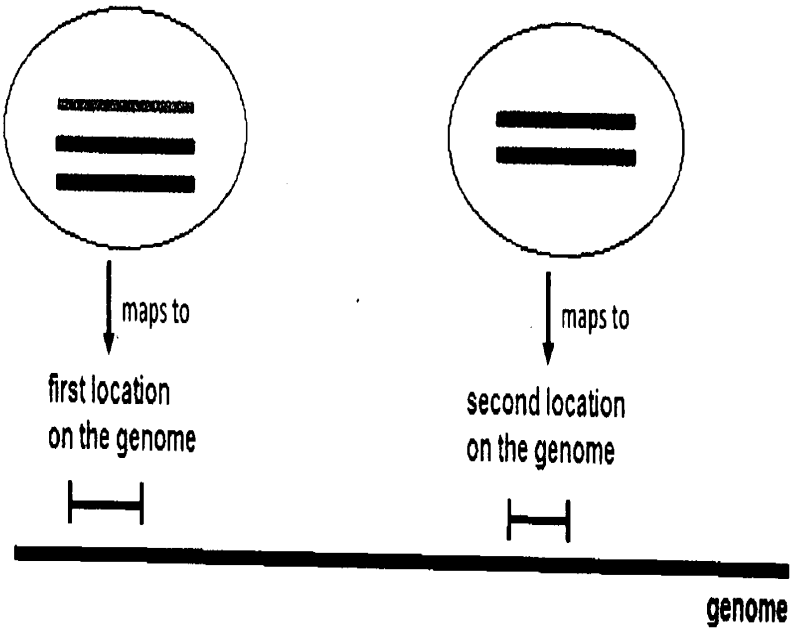
8/23
04-20-2015
07:59:28 p.m.
cooper dunham
1212391063O
PAGE 8/23 * RCVD AT 4/20/2015 8:58:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):10-30



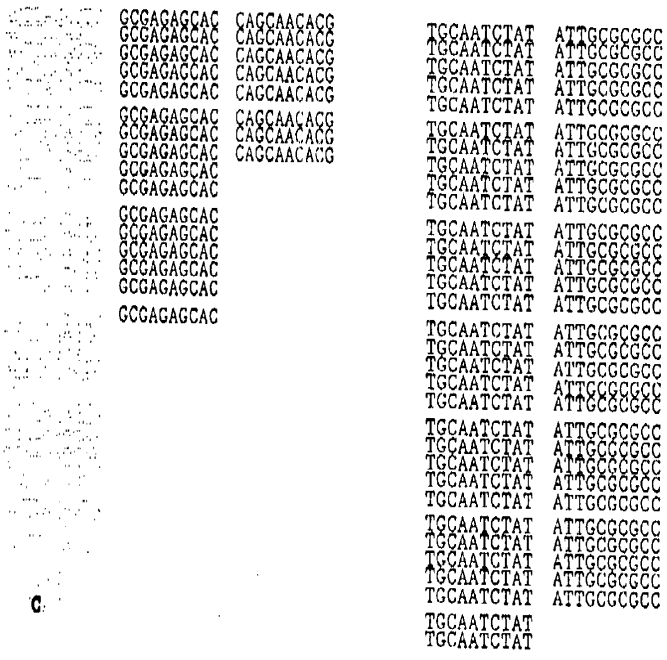
IPR2025-01353; IPR2025-01355
CSHL EX2002

07:59:58 p.m. 04-20-2015 9 /23

e) Assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read...




f) Counting the number of tagged nucleic acid molecules having a substantially unique tag that have been assigned to the same location on the genome,...



First location
Total unique tag counts = 4

Second location
Total unique tag counts = 2

12123910630 cooper dunham

PAGE 9/23 * RCVD AT 4/20/2015 8:56:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):10-30



IPR2025-01353; IPR2025-01355
CSHL EX2002

Table 1. Average error rates of thermostable DNA polymerases during PCR[a]

| DNA polymerase | No. of PCRs | Target (ng) | Template doublings[b] | *lacI*⁻ plaques[c] (% ± SD) | Error rate[d] (×10⁻⁶ ± SD) |
|---|---|---|---|---|---|
| *Pfu* | 10 | 24 | 9.7 | 0.42 ± 0.08 | 1.3 ± 0.2 |
| | 2 | 2 | 12.7 | 0.30 ± 0.06[e,f] | 0.7 ± 0.1[e] |
| | 2 | 0.2 | 16.0 | 0.47 ± 0.03[e,f] | 0.8 ± 0.02[e] |
| | 2 | 0.02 | 19.4 | 0.66 ± 0.03[e,f] | 1.0 ± 0.04[e] |
| *Deep Vent* | 4 | 24 | 9.7–10 | 0.9 ± 0.1 | 2.7 ± 0.2 |
| *Vent* | 6 | 24 | 8.7–10 | 0.9 ± 0.3 | 2.8 ± 0.9 |
| *Taq* | 11 | 24 | 8.7–11 | 2.7 ± 1.1 | 8.0 ± 3.9 |
| *UlTma* | 2 | 24 | 9.7 | 18.8 ± 0.8[e] | 55 ± 2[e] |

[a]PCRs were performed in each manufacturers' recommended buffer (all pH 8.8) in the presence of 200 μM each dNTP and 2 mM $MgSO_4$ (*Pfu*, *Deep Vent* and *Vent*), 2 mM $MgCl_2$ (*UlTma*) or 1.5 mM $MgCl_2$ (*Taq*).

[b]Template doublings ($d$) were determined using the equation $2^d$ = (amount of PCR product)/(amount of starting target). 24 ng *lacI* target corresponds to 50 ng *lacIOZα* plasmid template. The range of $d$ obtained is indicated.

[c]Mutant frequencies ($mf$) were determined by dividing the total number of blue plaques (*lacI*⁻ mutants) on X-gal plates by the total number of plaques containing a functional *lacZα* sequence (blue plaques on X-gal plus IPTG plates).

[d]Error rates were calculated using the equation $ER = mf/(bp \times d)$, where $mf$ is the mutation frequency, $bp$ is the number of detectable sites in *lacI* (=349; 15) and $d$ is the number of template doublings.

[e]± indicates range of duplicate measurements.

[f]Mutation frequencies for *Pfu* amplification of 0.02–2 ng target were normalized such that the mean mutation frequency for *Pfu* amplification of 24 ng target (assay internal control) was 0.42%.

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit C

10/23 04-20-2015 08:00:21 p.m. cooper dunham 12123910630
PAGE 10/23 * RCVD AT 4/20/2015 8:56:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):10-30



IPR2025-01353; IPR2025-01355
CSHL EX2002

3546–3551 Nucleic Acids Research, 1996, Vol. 24, No. 18 © 1996 Oxford University Press

# PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases

**Janice Cline, Jeffery C. Braman and Holly H. Hogrefe***

Stratagene Cloning Systems, 11011 North Torrey Pines Road, La Jolla, CA 92037, USA

Received June 12, 1996; Revised and Accepted July 29, 1996

## ABSTRACT

**The replication fidelities of *Pfu*, *Taq*, *Vent*, *Deep Vent* and *UlTma* DNA polymerases were compared using a PCR-based forward mutation assay. Average error rates (mutation frequency/bp/duplication) increased as follows: *Pfu* ($1.3 \times 10^{-6}$) < *Deep Vent* ($2.7 \times 10^{-6}$) < *Vent* ($2.8 \times 10^{-6}$) < *Taq* ($8.0 \times 10^{-6}$) << exo⁻ *Pfu* and *UlTma* (~$5 \times 10^{-5}$). Buffer optimization experiments indicated that *Pfu* fidelity was highest in the presence of 2–3 mM $MgSO_4$ and 100–300 μM each dNTP and at pH 8.5–9.1. Under these conditions, the error rate of exo⁻ *Pfu* was ~40-fold higher ($5 \times 10^{-5}$) than the error rate of *Pfu*. As the reaction pH was raised from pH 8 to 9, the error rate of *Pfu* decreased ~2-fold, while the error rate of exo⁻ *Pfu* increased ~9-fold. An increase in error rate with pH has also been noted for the exonuclease-deficient DNA polymerases *Taq* and exo⁻ Klenow, suggesting that the parameters which influence replication error rates may be similar in pol I- and α-like polymerases. Finally, the fidelity of 'long PCR' DNA polymerase mixtures was examined. The error rates of a *Taq*/*Pfu* DNA polymerase mixture and a Klen*taq*/*Pfu* DNA polymerase mixture were found to be less than the error rate of *Taq* DNA polymerase, but ~3–4-fold higher than the error rate of *Pfu* DNA polymerase.**

## INTRODUCTION

The use of high fidelity DNA polymerases in the polymerase chain reaction (PCR) is essential for reducing the introduction of amplification errors in PCR products that will be cloned, sequenced and expressed. Several thermostable DNA polymerases with 3′→5′ exonuclease-dependent proofreading activity (*Pfu*, *Vent*, *Deep Vent* and *UlTma*) have been introduced for high fidelity PCR amplification (1–3). Flaman *et al.* have reported that the error rate of *Pfu* was 5- and 30-fold lower than the error rates of the proofreading enzymes *Deep Vent* and *UlTma*, respectively (4). Using several different fidelity assays, the error rate of *Pfu* has been found to be ~10-fold lower than that of the non-proofreading enzyme *Taq* (1,4,5).

The parameters which contribute to the replication fidelity of DNA polymerases need to be investigated, as very little is known about the molecular features of these enzymes which give rise to variations in replication fidelity and mutational spectra. A number of factors are thought to contribute to the overall fidelity of a DNA polymerase (reviewed in (6–8). These parameters include the tendency of a polymerase to incorporate incorrect nucleotides and the presence of an integral 3′→5′ exonuclease activity which can remove mispaired bases (proofreading activity).

The importance of proofreading activity to replication fidelity has been demonstrated for both the Klenow fragment (9) and for *Vent* polymerase (10), which exhibit 10- and 5-fold increases in error rates, respectively, when the associated 3′→5′ exonuclease activity is inactivated. The contribution of proofreading activity to DNA polymerase fidelity is also evident when the error rates of proofreading and non-proofreading enzymes are compared. Kunkel has noted that the average base substitution error rates exhibited by non-proofreading DNA polymerases range from $10^{-2}$ to ≥ $10^{-6}$, while the error rates of proofreading enzymes range from $10^{-6}$ to $10^{-7}$ (7). The parameters which contribute to error rate variations among proofreading enzymes may reflect inherent differences in 3′→5′ exonuclease activity, the tendency to discriminate mispaired versus correctly paired bases and/or the efficiency of shuttling between polymerization and proofreading modes.

Recently, mixtures of non-proofreading and proofreading DNA polymerases have been reported to synthesize higher yields of PCR product and to allow amplification of longer templates than is possible with single enzyme formulations ('long PCR') (5). The addition of a low level of a proofreading enzyme (e.g. *Pfu* DNA polymerase) to PCR reaction mixtures has been proposed to improve the performance of non-proofreading polymerases (e.g. *Taq* DNA polymerase) by correcting mismatches introduced during PCR which prevent the efficient synthesis of full-length products (5). The PCR fidelity of DNA polymerase mixtures has not yet been determined, but error rates are likely to reflect the fidelity of the component polymerases and the ratio of non-proofreading to proofreading enzyme activities.

*Pfu* DNA polymerase has been found to be useful in high fidelity amplifications (1,4) of DNA targets up to 25 kb (K. Nielson, personal communication). In this report we use the previously described *lacI* PCR mutation assay (1) to compare the error rate of *Pfu* with an expanded number of PCR polymerases, including exo⁻ *Pfu*, *Deep Vent*, *Vent*, *UlTma* and *Taq*, as well as 'long PCR' DNA polymerase mixtures. Polymerase error rates have been found to vary with buffer composition, including pH, $Mg^{2+}$ concentration and nucleotide concentration (11–13). PCR reaction conditions have been optimized with respect to fidelity for

* To whom correspondence should be addressed

PAGE 11/23 * RCVD AT 4/20/2015 8:58:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):18-30
IPR2025-01353, IPR2025-01355
491 of 819
CSHL EX2002

both *Vent* and *Taq* DNA polymerases (11). Buffer optimization studies with *Pfu* DNA polymerase were performed here to assess whether the fidelity of *Pfu* DNA polymerase could be further enhanced. Error rate comparisons between *Pfu* and exo⁻ *Pfu* are expected to illuminate the contribution of proofreading activity to the fidelity of *Pfu* DNA polymerase. Finally, PCR fidelity comparisons between *Pfu* DNA polymerase and *Pfu*-containing DNA polymerase mixtures will allow evaluation of the contribution of the predominant non-proofreading enzyme to the error rate of 'long PCR' mixtures.

## MATERIALS AND METHODS

### DNA polymerases

Cloned *Pfu*, exo⁻ *Pfu* and *Taq* DNA polymerases were prepared at Stratagene. *Deep Vent* and *Vent* polymerases were purchased from New England BioLabs. *UlTma* was obtained from Perkin-Elmer and KlentaqLA (KTLA) was provided by Wayne Barnes (Washington University School of Medicine, St Louis, MO). Except where indicated, PCR amplifications were performed in the presence of buffers supplied by the manufacturers. The KTLA PCR buffer used was buffer PC2 of Barnes (5).

### PCR fidelity assay

The fidelity of DNA replication during PCR was measured using a previously described assay (1,14). Briefly, a 1.9 kb sequence encoding *lacIOZ*α was PCR amplified as described below with oligonucleotide primers containing 5′ *Eco*RI restriction sites (1). The amplified fragments were digested with *Eco*RI, purified by gel electrophoresis and ligated into λgt10 arms. The ligation reactions were packaged and the λ phage used to infect an α-complementing *Escherichia coli* host strain. Aliquots of infected cells were plated on LB plates with top agar containing either X-gal (1 mg/ml) or X-gal plus IPTG (1.5 mM). Error rates were calculated as described in the legend to Table 1.

### PCR amplifications

Except where indicated, PCR amplifications were performed in 100 µl reaction volumes in the presence of the appropriate Tris-based buffer, using 5 U polymerase, 200 µM each dNTP, 250 ng each primer and 24 ng *lacIOZ*α target (50 ng *lacIOZ*α plasmid template). The PCR mixtures were denatured by heating at 95°C for 30 s. Thirty cycles of amplification were performed using the following conditions: 5 s at 95°C; 1 min at 55°C; 2.5 min at 72°C.

## RESULTS

### PCR fidelity of thermostable DNA polymerases

Replication fidelities of thermostable DNA polymerases were compared using a previously described assay (1) which measures the frequency of mutations introduced into the *lacI* target gene during PCR amplification. PCR amplification was performed in the presence of each enzyme's optimal PCR buffer. All other PCR parameters remained constant, including the dNTP, primer and template concentrations, the PCR cycling parameters and the number of PCR cycles performed.

*Pfu* DNA polymerase exhibited the greatest PCR fidelity, with an average error rate of $1.3 \times 10^{-6}$ mutation frequency/bp/duplication (Table 1). The *lacI* target size used in these calculations was estimated to be 349 bp, based upon the most recent analysis of *lacI*⁻ mutant DNA sequences (15). Previous error rate calculations assumed a *lacI* target size of 182 bp (1). After recalculating error rates based on a *lacI* target size of 349 bp, the mean error rate of *Pfu* DNA polymerase obtained in this study ($1.3 \times 10^{-6}$ mutation frequency/bp/duplication) was found to be similar to previous estimates obtained using an identical assay ($0.8 \times 10^{-6}$; 1) or an alternative PCR-based assay employing a p53 target gene ($\leq 1.0 \times 10^{-6}$; 4).

**Table 1.** Average error rates of thermostable DNA polymerases during PCR[a]

| DNA polymerase | No. of PCRs | Target (ng) | Template doublings[b] | *lacI*⁻ plaques[c] (% ± SD) | Error rate[d] (×10⁻⁶ ± SD) |
|---|---|---|---|---|---|
| *Pfu* | 10 | 24 | 9.7 | 0.42 ± 0.08 | 1.3 ± 0.2 |
| | 2 | 2 | 12.7 | 0.30 ± 0.06[e,f] | 0.7 ± 0.1[e] |
| | 2 | 0.2 | 16.0 | 0.47 ± 0.03[e,f] | 0.8 ± 0.02[e] |
| | 2 | 0.02 | 19.4 | 0.66 ± 0.03[e,f] | 1.0 ± 0.04[e] |
| *Deep Vent* | 4 | 24 | 9.7–10 | 0.9 ± 0.1 | 2.7 ± 0.2 |
| *Vent* | 6 | 24 | 8.7–10 | 0.9 ± 0.3 | 2.8 ± 0.9 |
| *Taq* | 11 | 24 | 8.7–11 | 2.7 ± 1.1 | 8.0 ± 3.9 |
| *UlTma* | 2 | 24 | 9.7 | 18.8 ± 0.8[e] | 55 ± 2[e] |

[a]PCRs were performed in each manufacturers' recommended buffer (all pH 8.8) in the presence of 200 µM each dNTP and 2 mM $MgSO_4$ (*Pfu*, *Deep Vent* and *Vent*), 2 mM $MgCl_2$ (*UlTma*) or 1.5 mM $MgCl_2$ (*Taq*).

[b]Template doublings (*d*) were determined using the equation $2^d$ = (amount of PCR product)/(amount of starting target). 24 ng *lacI* target corresponds to 50 ng *lacIOZ*α plasmid template. The range of *d* obtained is indicated.

[c]Mutant frequencies (*mf*) were determined by dividing the total number of blue plaques (*lacI*⁻ mutants) on X-gal plates by the total number of plaques containing a functional *lacZ*α sequence (blue plaques on X-gal plus IPTG plates).

[d]Error rates were calculated using the equation $ER = mf/(bp \times d)$, where *mf* is the mutation frequency, *bp* is the number of detectable sites in *lacI* (=349; 15) and *d* is the number of template doublings.

[e]± indicates range of duplicate measurements.

[f]Mutation frequencies for *Pfu* amplification of 0.02–2 ng target were normalized such that the mean mutation frequency for *Pfu* amplification of 24 ng target (assay internal control) was 0.42%.





IPR2025-01353, IPR2025-01355
CSHL EX2002

Average error rates of thermostable DNA polymerases were found to increase in the following order: *Pfu* ($1.3 \times 10^{-6}$) < *Deep Vent* ($2.7 \times 10^{-6}$) < *Vent* ($2.8 \times 10^{-6}$) < *Taq* ($8.0 \times 10^{-6}$) << *UlTma* ($5.5 \times 10^{-5}$). These results are in excellent agreement with the relative error rates measured by Flaman *et al.* (4), who reported that *Pfu* exhibits the greatest PCR fidelity, followed by *Deep Vent*, *Taq* and *UlTma* DNA polymerases. The relative error rates obtained here are also consistent with DGGE analyses showing that *Pfu* exhibits a lower error rate than *Vent* and *Taq* DNA polymerases (16). We found that relative error rates observed using the *lacI* screening assay were consistent from PCR reaction to PCR reaction.

The influence of template doublings (*d*) on error rate estimates of *Pfu* DNA polymerase was also examined (Table 1). Amplification reactions described above and resulting in the *Pfu* error rate of $1.3 \times 10^{-6}$ employed 24 ng *lacI* target ($10^{10}$ copies). Approximately 10 doublings were observed in 30 PCR cycles. When the input *lacI* target DNA was decreased from $10^{10}$ copies (24 ng) to $10^{7}$ copies (0.02 ng), the number of template doublings increased from 9.7 (~900-fold amplification) to 19.4 (~700 000-fold amplification) after 30 cycles of PCR. The error rate of *Pfu* DNA polymerase varied from 0.7 to $1.3 \times 10^{-6}$ over the 1000-fold range of DNA target concentrations tested. Flaman *et al.* have also reported that polymerase error rates do not appear to be significantly influenced by the number of template doublings (4).

### Optimization of the PCR fidelity of *Pfu*

We attempted to further improve the fidelity of *Pfu* by optimizing PCR reaction conditions. PCR error rates were measured at varying concentrations of $MgSO_4$ (Fig. 1) and dNTPs (Fig. 2) and at varying pHs (Fig. 3). The indicated pH values are those measured at room temperature. Where noted, the pH of Tris buffers at elevated temperatures was estimated using the formula: $pH_T = pH_{25°C} + [(T°C - 25°C) \times (-0.03\ pH\ U/°C)]$ (where $T$ is the reaction temperature; 17). The lowest error rates for *Pfu* were observed when PCR amplifications were performed in the presence of 2–3 mM $MgSO_4$, 100–300 µM each dNTP and in a pH range between 8.5 and 9.1 (pH ~7.1–7.7 at 72°C). These conditions have been found to give optimal yield of PCR product as well (18).

In the presence of 1 mM $MgSO_4$ and 800 µM total dNTPs, the error rate of *Pfu* was ~3-fold higher than when PCR amplifications were performed in 2 mM $MgSO_4$ at the same dNTP concentration (Fig. 1). The error rate did not vary significantly as the $MgSO_4$ concentration was increased from 2 to 10 mM (~1.2–9.2 mM free $Mg^{2+}$). The error rate of *Vent* polymerase has also been shown to decrease significantly between 0.5 and 2 mM $MgSO_4$ in the presence of 2 mM total dNTPs and thereafter remains constant with increasing concentrations of free $Mg^{2+}$ (11). These results are in contrast to those reported for *Taq*, in which error rates are lowest at equimolar concentrations of $MgCl_2$ and dNTPs (1 mM) and increase with increasing concentration of free $Mg^{2+}$ (12). Error rate variations of *Pfu* and *Vent* likely reflect the $Mg^{2+}$ dependency of both proofreading and polymerase activities.

In Figure 2, the error rate of *Pfu* was found to increase 2.4-fold as the total dNTP concentration was raised from 0.4 to 4 mM in the presence of a constant amount of free $Mg^{2+}$ (~1.2 mM). These results are consistent with the observations of Clayton *et al.* (19), who report that high concentrations of dNTPs diminish the proofreading ability of exonuclease-proficient polymerases by increasing the efficiency of mispair extension. It is likely that the fidelity of *Pfu* DNA polymerase could be further increased by reducing the total dNTP concentration below 0.4 mM total dNTPs. However, using lower dNTP concentrations to increase the fidelity of PCR amplification reactions is not practical, as PCR product yields decrease significantly below 0.4 mM total dNTPs.

In Figure 3 (inset), the error rate of *Pfu* was measured as a function of pH. The error rate of *Pfu* was found to decrease 4-fold between pH 7.5 and 8.5 in the presence of 2 mM $MgSO_4$ and 0.8 mM total dNTPs. *Vent* polymerase has also been reported to

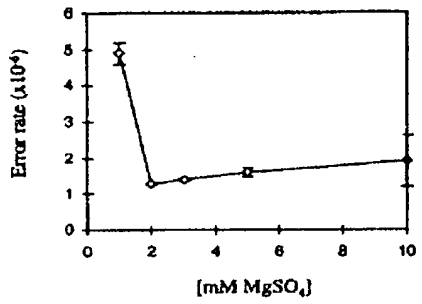



**Figure 1.** Variation of the PCR error rates of *Pfu* DNA polymerase with $MgSO_4$ concentration. PCR amplification was performed in buffer containing 20 mM Tris–HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 µg/ml BSA, 200 µM each dNTP and varying concentrations of $MgSO_4$ (1–10 mM). Error rates shown are the average (± range) values obtained from two independent PCR amplifications.

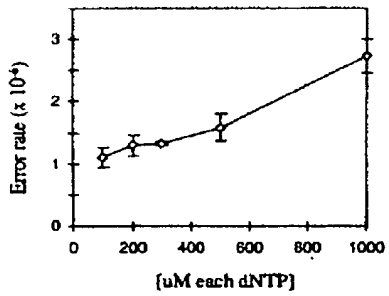



**Figure 2.** Variation of the PCR error rates of *Pfu* DNA polymerase with dNTP concentration. PCR amplification was performed in buffer containing 20 mM Tris–HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100µg/ml BSA and varying concentrations of dNTPs (100–1000 µM each). The $MgSO_4$ concentration of the PCR mixtures was adjusted to give a constant free $Mg^{2+}$ concentration (1.2 mM) at each dNTP concentration. 1.6 mM $MgSO_4$/0.4 mM total dNTPs; 2 mM $MgSO_4$/0.8 mM total dNTPs; 2.4 mM $MgSO_4$/1.2 mM total dNTPs; 3.2 mM $MgSO_4$/2 mM total dNTPs; 5.2 mM $MgSO_4$/4 mM total dNTPs. Error rates shown were normalized such that the mean mutation frequency for *Pfu* amplifications with 0.8 mM total dNTPs (assay internal control) was $1.3 \times 10^{-6}$ mutation frequency/bp/duplication. The average error rates (± range) from two independent PCR amplifications are shown.





IPR2025-01353, IPR2025-01355
CSHL EX2002

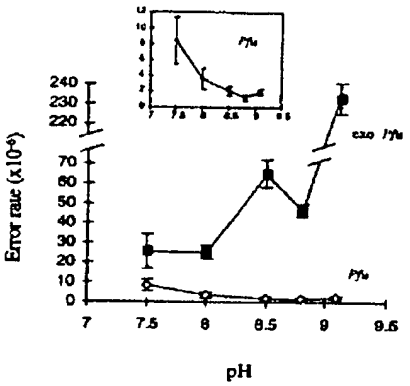



**Figure 3.** Variation of the PCR error rates of *Pfu* and exo⁻ *Pfu* DNA polymerases with pH. PCR amplification was performed in 20 mM Tris–HCl buffers whose pH values ranged from 7.5 to 9.1 (pH at 25°C). In addition, the buffer contained 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 μg/ml BSA and 200 μM each dNTP. The average error rate of *Pfu* (open diamonds) is shown in comparison with exo⁻ *Pfu* (filled diamonds) and in the accompanying inset. Error rates shown are the average (± range) values obtained from two independent PCR amplifications.

exhibit a significant decrease in error rate as the pH is increased from 7 to 8 in the presence of 2 mM $MgSO_4$ (11). For *Taq* DNA polymerase, a 2-fold increase in error rate was observed when the reaction pH was raised from 8 to 9 (11). Eckert and Kunkel have also reported that the base substitution and frameshift error rates of *Taq* (12) and exo⁻ Klenow (13) increase >10-fold as the reaction pH is raised from ~6.5 to 9.5 (25°C estimates of pH from 12, 13).

### pH dependency of the fidelity of *Pfu* and exo⁻ *Pfu*

The error rates of *Pfu* and exo⁻ *Pfu* were compared to assess the contribution of 3′→5′ exonuclease activity to fidelity. In the presence of *Pfu* PCR buffer (2 mM $MgSO_4$, 200 μM each dNTP, pH 8.8), exo⁻ *Pfu* exhibited an error rate of $4.7 \times 10^{-5}$ mutation frequency/bp/duplication, which is ~40-fold higher than that determined for exonuclease-proficient *Pfu*.

Figure 3 shows the error rate variation of exo⁻ *Pfu* and *Pfu* as a function of pH. Exo⁻ *Pfu* shows a dramatic increase in error rate (~9-fold) as the reaction pH is raised from pH 8 to 9.1 (or from 6.6 to 7.7 at 72°C). In contrast to exo⁻ *Pfu*, the error rate of *Pfu* decreased ~2-fold in this pH range. Presumably, the fidelity of *Pfu* is maintained at high pH (pH 9) by enhanced proofreading activity, which accompanies the dramatic increase in nucleotide misincorporation occurring between pH 8 and 9.1 (identified using exo⁻ *Pfu*). These results and those reported by others for *Taq* and exo⁻ Klenow (11–13) indicate that the error rates of exonuclease-deficient enzymes, *Taq*, exo⁻ Klenow and exo⁻ *Pfu*, are similarly increased by pH. The significance of the apparent biphasic relationship between error rate and pH is currently under investigation.

### PCR fidelity of 'long PCR' DNA polymerase mixtures

The fidelities of *Pfu* and *Taq* DNA polymerases were compared with the fidelities of two *Pfu*-containing DNA polymerase mixtures (Table 2). A *Taq*/*Pfu* (16 U:1 U) mixture was prepared and shown to amplify DNA targets >30 kb (data not shown). The *Taq*/*Pfu* mixture exhibited an average error rate of $5.6 \times 10^{-6}$ mutation frequency/bp/duplication when amplifications were performed in *Taq* PCR buffer. The mean error rate of the *Taq*/*Pfu* mixture was 30% lower than the mean error rate of *Taq* DNA polymerase when amplifications were conducted in *Taq* PCR buffer. When compared with the error rate of *Pfu* DNA polymerase in the same buffer system, the error rate of the *Taq*/*Pfu* mixture was found to be 6-fold higher.

Similar observations were made for a second 'long PCR' mixture, KTLA, which consists of Klentaq (N-terminally truncated *Taq*) and *Pfu* DNA polymerases (5). When PCR amplifications were conducted as described in this report, KTLA exhibited a mean error rate of $3.9 \times 10^{-6}$ mutation frequency/bp/duplication, which was 3-fold higher than the error rate of *Pfu* DNA polymerase (Table 2). When PCR conditions from Barnes (5) were used (Table 2, condition 2), KTLA exhibited a mean error rate ($9.4 \times 10^{-6}$) which was 4-fold higher than the error rate of *Pfu* DNA polymerase assayed under identical conditions.

**Table 2.** Error rate comparisons of DNA polymerases and 'long PCR' DNA polymerase mixtures

| PCR condition[a] | DNA polymerase | No. of PCRs | Template doublings[b] | Error rate (× $10^{-6}$ ± SD)[b] |
|---|---|---|---|---|
| 1 | *Pfu* | 10 | 8.7–9.7 | 1.3 ± 0.2 |
| | *Taq* | 11 | 8.7–11 | 8.0 ± 3.9 |
| | *Taq*/*Pfu* (16:1)[c] | | | |
| | *Taq* buffer | 8 | 9.7–10 | 5.6 ± 1.6 |
| | *Pfu* buffer | 11 | 9.7–11 | 7.6 ± 3.2 |
| | KTLA | 2 | 9.7 | 3.9 ± 0.1[d] |
| 2 | *Pfu* | 2 | 8.1 | 2.3 ± 0.2[d] |
| | KTLA | 2 | 8.1 | 9.4 ± 0.9[d] |

[a]PCR condition 1 is described in Materials and Methods, PCR amplification. PCR condition 2 is from Barnes (5) and differs in the following respects. PCR amplifications were performed on a Robocycler 40 (Stratagene) using thin-walled PCR tubes and 7.2 ng target (15 ng *lacI*/*OZα* plasmid). Sixteen cycles of amplification were performed using the following conditions: 30 s at 99°C, 30 s at 67°C, 3 min at 68°C.
[b]Defined in the legend to Table 1.
[c]The *Taq*:*Pfu* mixture consists of *Taq* (5 U/μl) and *Pfu* (0.31 U/μl) DNA polymerases.
[d]± indicates range of duplicate measurements.

PAGE 14/23 * RCVD AT 4/20/2015 8:56:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):18-30



IPR2025-01353, IPR2025-01355
CSHL EX2002

## DISCUSSION

The intrinsic properties of thermostable DNA polymerases which contribute to variation in PCR fidelity are not fully understood. In general, enzymes which possess an associated 3′→5′ exonuclease-dependent proofreading activity are thought to exhibit higher replication fidelity than non-proofreading DNA polymerases (7). Variation in fidelity among proofreading enzymes, such as *Pfu*, *Vent* and *Deep Vent*, may reflect differences in the rate of mispair excision, the level of discrimination between mispaired and correctly paired bases, the rate of mispair extension and/or the efficiency of shuttling the 3′ primer terminus between the polymerase and exonuclease active sites.

The contribution of 3′→5′ exonuclease activity to the PCR fidelity of *Pfu* was demonstrated directly by comparing the error rates of *Pfu* and exo⁻ *Pfu*. The error rate of exo⁻ *Pfu* was found to be 7-fold higher than the error rate of exo⁺ *Pfu* at pH 8.0 and 40-fold higher at pH 8.8 (*Pfu* PCR buffer).

Despite the importance of proofreading activity to the fidelity of *Pfu* and *Vent* (10), the presence of 3′→5′ exonuclease activity does not necessarily guarantee high fidelity DNA synthesis, as illustrated by *UlTma* DNA polymerase. The poor fidelity of *UlTma* DNA polymerase may be related to the relatively low level of 3′→5′ exonuclease activity exhibited by this enzyme. In a preliminary analysis of exonuclease activity, *UlTma* was found to exhibit significantly lower levels of 3′→5′ exonuclease activity than *Pfu*, *Deep Vent* and *Vent* DNA polymerases (A. Lovejoy, personal communication). However, other parameters are likely to contribute to low fidelity, since *UlTma*, an N-terminally deleted version of *Thermatoga maritima* DNA polymerase (20), exhibits an ~7-fold higher error rate than *Taq*, which is completely devoid of proofreading activity.

In the absence of proofreading activity, a DNA polymerase like *Taq* is thought to accomplish high fidelity DNA synthesis by inefficient incorporation of non-complementary dNTPs and a reduced tendency to extend from mismatched 3′ primer termini. Huang *et al.* (21) have shown that, with the exception of C–T mispairs, *Taq* polymerase exhibits ~100–1000-fold greater discrimination against mispair extension, as compared with avian myeloblastosis and HIV-1 reverse transcriptases, which extend most mispairs permissively. The rate at which DNA polymerases extend from mispaired 3′ primer termini, however, does not contribute to the actual fidelity of non-proofreading enzymes. The mismatch extension rate only contributes to fidelity in the sense that if the mismatch is extended inefficiently, the DNA will not be replicated to completion and the mutation will not be scored. Therefore, the mispair extension rate influences the number of detected mutants, rather than reflecting the inherent fidelity of a non-proofreading DNA polymerase.

The observed 6-fold difference in error rate between *Taq* ($8 \times 10^{-6}$) and exo⁻ *Pfu* ($4.7 \times 10^{-5}$) suggests that the misincorporation and/or misextension rates of *Pfu* (as measured with exo⁻ *Pfu*) are significantly higher than those of *Taq*. Apparently, a lower degree of discrimination against misinsertion or mispair extension errors can be tolerated when an associated proofreading activity is present, as is the case with exonuclease-proficient *Pfu*.

Further fidelity measurements with exo⁻ *Pfu* revealed that the fidelity of dNTP incorporation was significantly influenced by the pH of the PCR buffer. The error rate increased by ~9-fold as the pH was raised from pH 6.6 to 7.7 (pH at 72°C). The error rates of both *Taq* (12) and exo⁻ Klenow (13) increase similarly at higher pH. Eckert and Kunkel have attributed the lower fidelity of exo⁻ Klenow at high pH to an increase in both nucleotide misinsertion and mispair extension (13).

It is tempting to speculate that the lower fidelity of exo⁻ *Pfu* at high pH may also reflect increased misinsertion and mispair extension, analogous to the observations made for exo⁻ Klenow (13). If so, it would suggest that the parameters which contribute to fidelity are similar, despite the structural differences which are thought to exist between the α-like (exo⁻ *Pfu*; 22) and pol I-like (exo⁻ Klenow and *Taq*) DNA polymerases. For example, the observed variation in error rates with pH suggests that an active site histidine residue may play a role in fidelity, possibly in the discrimination of mismatched 3′ primer termini. Alternatively, protonation of the primer, template or substrate dNTP may enhance error discrimination (13). Finally, pol I- and α-like polymerases may undergo a similar conformational change at low pH which may alter template binding properties, thereby improving error discrimination. Such a mechanism was proposed for exo⁻ Klenow by Eckert and Kunkel (13) and was supported by additional data showing that lower error rates at low pH were accompanied by an increase in polymerase processivity.

The relative error rates for *Pfu*, *Vent* and *Taq* were found to parallel the terminal transferase activities of DNA polymerases. Hu (23) has compared the tendency of DNA polymerases to catalyze the addition of non-template-directed bases to the 3′-end of a DNA fragment (terminal transferase activity). Terminal transferase activity is high in *Taq* but low (Klenow and *Vent*) or absent (*Pfu*, *T4* and *T7*) in proofreading enzymes, which presumably edit the misextended base. The absence of terminal transferase activity appears to correlate with high fidelity. Fidelity measurements compiled by Cha and Thilly show that the error rates of *Pfu*, *T4* and *T7* DNA polymerases are lower than the error rates of *Vent* and Klenow (16). Thus, the parameters which give rise to terminal transferase activity may be similar to those which contribute to lower fidelity. The lower error rate and lack of terminal transferase activity for *Pfu* (as compared with *Vent*) may be the result of a reduced tendency of *Pfu* to incorporate a mismatch or a base opposite an abasic site. Alternatively, *Pfu* may excise misincorporated bases more readily or shuttle between the exonuclease and polymerase active sites more efficiently.

Finally, fidelity comparisons with *Pfu*-containing 'long PCR' DNA polymerase mixtures have shown that the error rate of mixtures appears to be intermediate between the error rate of *Pfu* and the non-proofreading DNA polymerase. The lower error rate of a *Taq*/*Pfu* mixture, as compared with *Taq* alone, suggests that *Pfu* is editing a certain percentage of mismatches that have been introduced by *Taq* during the PCR process. Editing may occur at the 3′-terminus after *Taq* has introduced a mismatch and dissociated from the incomplete PCR product (5). In the absence of *Pfu*, *Taq* presumably extends some of these putative stalling mismatches during the course of the PCR process; otherwise the mutations would not be scored in the *lacI*⁻ screening assay and there would be no apparent difference in error rate between *Taq* and the *Taq*/*Pfu* mixture. *Pfu* may also reduce the overall error rate of *Taq* DNA polymerase by degrading *Taq*-generated duplex DNA containing mismatches and resynthesizing the correct sequence.

Although the error rate of the *Taq*/*Pfu* mixture is somewhat lower than the error rate of *Taq* alone, it is still 4–6-fold higher than the error rate of *Pfu* alone (Table 2). These results indicate that the majority of PCR products are synthesized by *Taq*. This



IPR2025-01353, IPR2025-01355
495 of 819
CSHL EX2002

result is not surprising, since *Taq* is present in this particular mixture at a 16-fold higher polymerase unit concentration than *Pfu* DNA polymerase. Hence, the misincorporation rate of *Taq* DNA polymerase contributes significantly to the error rate of *Taq*/*Pfu* DNA polymerase mixtures.

KTLA, a 'long PCR' mixture of Klentaq and *Pfu* DNA polymerases, was also found to exhibit an error rate significantly higher than the error rate of *Pfu*. Our results are inconsistent with the results of Barnes (5), who has compared the error rates of *Pfu*, Klentaq and KTLA-64 (~640 U Klentaq:1 U *Pfu*) using a similar PCR forward mutation assay based on the mutational target gene *lacZ*. Barnes reported that the error rate of the KTLA mixture was 2-fold lower than the error rate of *Pfu* DNA polymerase (5). There are several differences between the Barnes assay and the assay performed here, including PCR amplification conditions (see Table 2 legend), number of clones screened [500–4200 clones/1 PCR in Barnes (5) versus 10 000–50 000 clones/PCR/4 PCRs in this study] and the mutational target gene used (*lacZ* versus *lacI*), as well as possible unknown variations in the KTLA mixtures. The results in Table 2 demonstrate that differences in the PCR amplification conditions employed are not likely to contribute to the differences in relative error rates observed in the two studies. Fidelity analyses of additional DNA polymerase mixtures are currently under way to help elucidate the role of component enzymes and buffer composition in the fidelity of 'long PCR' amplifications.

## REFERENCES


1 Lundberg,K.S., Shoemaker,D.D., Adams,M.W.W., Short,J.M., Sorge,J.A. and Mathur,E.J. (1991) *Gene*, **180**, 1–6.
2 Kong,H., Kucera,R.B. and Jack,W.E. (1993) *J. Biol. Chem.*, **268**, 1965–1975.
3 Abramson,R.D. (1995) In Innis,M.A., Gelfand,D.H. and Sninsky,J.J. (eds), *PCR Strategies*. Academic Press, San Diego, CA.
4 Flaman,J.-M., Frebourg,T., Moreau,V., Charbonnier,F., Martin,C., Ishioka,C., Friend,S.H. and Iggo,R. (1994) *Nucleic Acids Res.*, **22**, 3259–3260.
5 Barnes,W.M. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 2216–2220.
6 Echols,H. and Goodman,M.F. (1991) *Annu. Rev. Biochem.*, **60**, 477–511.
7 Kunkel,T.A. (1992) *J. Biol. Chem.*, **267**, 18251–18254.
8 Goodman,M.F., Creighton,S., Bloom,L.B. and Petruska,J. (1993) *Crit. Rev. Biochem. Mol. Biol.*, **28**, 83–126.
9 Kunkel,T.A. (1988) *Cell*, **53**, 837–840.
10 Mattila,P., Korpela,J., Tenkanen,T. and Pitkanen,K. (1991) *Nucleic Acids Res.*, **19**, 4967–4973.
11 Ling,L.L., Keohavong,P., Dias,C. and Thilly,W.G. (1991) *PCR Methods Applicat.*, **1**, 63–69.
12 Eckert,K.A. and Kunkel,T.A. (1990) *Nucleic Acids Res.*, **18**, 3739–3744.
13 Eckert,K.A. and Kunkel,T.A. (1993) *J. Biol. Chem.*, **268**, 13462–13471.
14 Cline,J., Braman,J. and Kretz,K. (1995) *Strategies Mol. Biol.*, **8**, 24–25.
15 Provost,G.S., Kretz,P.L., Hamner,R.T., Matthews,C.D., Rogers,B.J., Lundberg,K.S., Dycaico,M.J. and Short,J.M. (1993) *Mutat. Res.*, **288**, 133–149.
16 Cha,R.S. and Thilly,W.G. (1995) In Dieffenbach,C.W. and Dveksler,G.S. (eds), *PCR Primer*. Cold Spring Harbor Laboratory Press, Plainview, NY.
17 Good,N.E. (1986) *Biochemistry*, **5**, 467–476.
18 Nielson,K.B., Costa,G.L. and Braman,J. (1996) *Strategies Mol. Biol.*, **9**, 24–25.
19 Clayton,L.K., Goodman,M.F., Branscomb,E.W. and Galas,D.J. (1979) *J. Biol. Chem.*, **254**, 1902–1912.
20 Innis,M.A., Myambo,K.B., Gelfand,D.H. and Brow,M.A.D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 9436–9440.
21 Huang,M.-M., Arnheim,N. and Goodman,M.F. (1992) *Nucleic Acids Res.*, **20**, 4567–4573.
22 Mathur,E.J., Adams,M.W.W., Cullen,W.N. and Cline,J.M. (1991) *Nucleic Acids Res.*, **19**, 6952.
23 Hu,G. (1993) *DNA Cell Biol.*, **12**, 763–770.


PAGE 16/23 * RCVD AT 4/20/2015 8:56:39 PM [Eastern Daylight Time] * SVR:W-PTOFAX-001/31 * DNIS:2738300 * CSID:12123910630 * DURATION (mm-ss):16-30
IPR2025-01353, IPR2025-01355
496 of 819
CSHL EX2002

SHEET 8/8

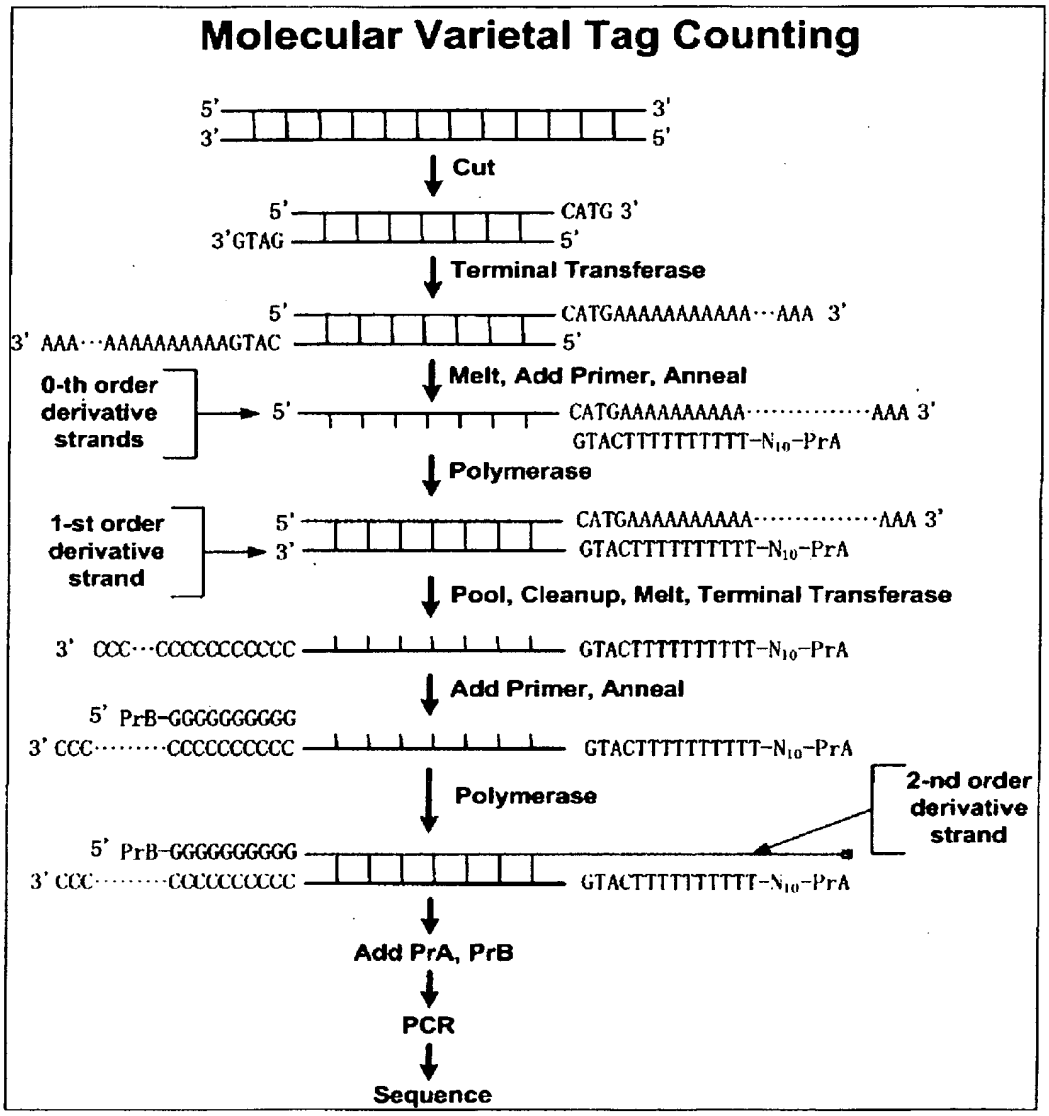



Fig. 8

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit A

[redacted]

RECEIVED
CENTRAL FAX CENTER
APR 20 2015

COOPER & DUNHAM LLP
ATTORNEYS AT LAW
30 ROCKEFELLER PLAZA, NEW YORK, NEW YORK 10112
TELEPHONE: (212) 278-0400
FACSIMILE: (212) 391-0525; 391-0526; 391-0630
www.cooperdunham.com

JOHN P. WHITE
GARY J. GERSHIK
ROBERT T. MALDONADO
JEFFREY L. SNOW
ERIC M. EISENBERG
ALISSA M. PACCHIOLI
JASON A. KEITH
ELANA B. ARAJ
DARREN HABER*
ASHWAT RISHI*

PAUL TENG
WENDY E. MILLER
TONIA A. SAYOUR
CINDY S. SHU
JAMES T. CORCORAN
AARON P. SELIKSON
JOSEPH V. MICALI
JOONG Y. CHO
LINSHENG LI*

*LAW CLERKS

OF COUNSEL
CHRISTOPHER C. DUNHAM
NORMAN H. ZIVIN
IVAN S. KAVRUKOV
PETER D. MURRAY
WILLIAM E. PELTON

SCIENTIFIC ADVISORS
ROBERT E. BOLCOME, III, PH.D.
LAWRENCE P. TARDIBONO, JR., PH.D.
PHILIP E. ALFORD, PH.D.
MINERVA Y. WONG, PH.D.

REGISTERED PATENT AGENTS
ANDREW CHUNG, B.E.
REGINA CHENG, B.S.

FOUNDED 1887

**FACSIMILE TRANSMISSION**

PLEASE DELIVER THE FOLLOWING PAGES

TO : Examiner Frank Wei Min Lu

FACSIMILE NO.: 571-273-8300

FROM : Gary J. Gershik/MW

TOTAL NUMBER OF PAGES, INCLUDING COVER PAGE: 23 Pages

DATE : April 20, 2015

SUBJECT : Agenda for April 21, 2015 Telephonic Examiner Interview for U.S. Serial No. 13/278,333

CLIENT/DOCKET: Our Docket 5981/81503-A/GJG/MW

IF YOU DO NOT RECEIVE ALL THE PAGES, PLEASE CALL BACK AS SOON AS POSSIBLE TO (212) 278-0400.

==========

**THE INFORMATION CONTAINED IN THIS FACSIMILE TRANSMISSION IS INTENDED SOLELY FOR THE PERSONAL AND CONFIDENTIAL USE OF THE DESIGNATED RECIPIENT(S) NAMED ABOVE. THIS TRANSMISSION MAY BE AN ATTORNEY-CLIENT COMMUNICATION CONTAINING INFORMATION THAT IS PRIVILEGED AND CONFIDENTIAL. IF THE READER OF THIS MESSAGE IS NOT A DESIGNATED RECIPIENT OR AN AGENT RESPONSIBLE FOR DELIVERING IT TO A DESIGNATED RECIPIENT, YOU ARE HEREBY NOTIFIED THAT YOU HAVE RECEIVED THIS DOCUMENT IN ERROR, AND THAT ANY REVIEW, DISTRIBUTION, OR COPYING OF THIS MESSAGE IS STRICTLY PROHIBITED. IF YOU HAVE RECEIVED THIS COMMUNICATION IN ERROR, OR IF UPON READING THIS DOCUMENT YOU HAVE REASON TO BELIEVE THAT THE DOCUMENT WAS INADVERTENTLY SENT TO YOU, PLEASE NOTIFY US IMMEDIATELY BY COLLECT TELEPHONE CALL AND RETURN THE ORIGINAL MESSAGE TO US BY MAIL. THANK YOU.**

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
June 4, 2015

BY EFS
Commissioner for Patents
Alexandria, VA

**AMENDMENT UNDER 37 C.F.R. § 1.116 IN RESPONSE TO FEBRUARY 4, 2015 FINAL OFFICE ACTION, PETITION FOR ONE-MONTH EXTENSION OF TIME, AND SUMMARY OF APRIL 21, 2015 EXAMINER INTERVIEW**

This Amendment is a submission under 37 C.F.R. § 1.116 and responds to the issues raised in a February 4, 2015 Final Office Action issued by the U.S. Patent and Trademark Office in connection with the above-identified application. A response to the February 4, 2015 Final Office Action was due May 4, 2015. Applicants hereby petition for a one-month extension of time. The fee for a one-month extension of time for a small entity is ONE HUNDRED DOLLARS ($100.00), and authorization is hereby given to charge this amount of fee to Deposit Account No. 03-3125. With a one-month extension of time, a response is due June 4, 2015. Accordingly, this response is being timely filed.

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

**Remarks** begin on page **10** of this paper.

A **Summary of April 21, 2015 Examiner Interview** begins on page **11** of this paper.



IPR2025-01353; IPR2025-01355
CSHL EX2002

**AMENDMENTS TO THE CLAIMS**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims**

1. (Currently Amended) A method for obtaining from genomic ~~material~~ nucleic acids genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic ~~material~~ nucleic acids;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic ~~material~~ nucleic acids from step (a) and a substantially unique tag having a unique nucleotide sequence, so as to permit differentiation from another tag having a unique nucleotide sequence, and which tag is part of another tagged nucleic acid molecule;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic ~~material~~ nucleic acids by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (d) corresponding to one segment of the segments of the

genomic ~~material~~ nucleic acids to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Previously Presented) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is

comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising i) summing the counts obtained in step (f) for locations which comprise a region of the genome~~, and comparing the sum to a sum~~; ii) summing the counts obtained from a reference sample for the same region of the genome[[,]]; and iii) comparing the sums obtained in step (i) and (ii), thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Currently Amended) A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag

associated sequence read <u>of said tag associated sequence reads of step (c)</u> corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. (Cancelled)

11. (Currently Amended) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic ~~material~~ <u>nucleic acids</u> to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Currently Amended) The method of claim 1 step (b), wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic ~~material~~ nucleic acids to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein adding the polynucleotide tail comprises the use of a terminal transferase.


IPR2025-01353; IPR2025-01355
CSHL EX2002

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to ~~sequencing in~~ step (d).

25. (Cancelled)

26. (Currently Amended) The method of claim 1, wherein the substantially unique tags ~~sequences~~ further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to said sequencing of the product of step (c).

28. (Previously Presented) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture

has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.

**REMARKS**

Claims 1-7, 9, 11-13, 24, 26-28, 31 and 68-71 are pending in the subject application. Applicants have amended claims 1, 7, 9, 11, 12, 24, 26, and 70 for clarity. Accordingly, upon entry of this Amendment claims 1-7, 9, 11-13, 24, 26-28, 31 and 68-71 will be pending and under examination in the subject application.

The claim amendments herein are appropriate for entry under 37 C.F.R. § 1.116 because the amendments (1) comply with requirements of form expressly set forth in the February 4, 2015 Final Office Action, and (2) do not raise new issues that would require further consideration and/or search. Specifically, claims 1, 7, 9, 11, 12, 24, 26, and 70 were amended for clarity according to suggestions by Examiner Lu during the April 21, 2015 Examiner Interview, summarized below, and do not introduce any new features which would require additional search. Therefore, the amendments are proper and must be entered.

**Summary of April 21, 2015 Examiner Interview**

On April 21, 2015, a telephonic Examiner Interview was conducted in connection with the subject application. Present at the interview were Examiner Frank Wei Min Lu of the U.S. Patent and Trademark Office, and applicants' representative Mr. Vladimir Drozdoff. and applicants' undersigned representative. Applicants acknowledge with appreciation the courtesy that Examiner Lu extended during the April 21, 2015 Interview.

On April 21, 2015, applicants sent an Agenda to Examiner Lu by facsimile. A copy of the April 21, 2015 Agenda is attached hereto as **Exhibit 1.**

During the April 21, 2015 Interview, the rejections of record under 35 U.S.C. § 112 were discussed.

*Rejection of claims under 35 U.S.C. § 112, paragraph 1 (enablement)*
Applicants' representative explained that the claims were not subject to the rejections under 35 U.S.C. § 112, first paragraph (enablement), set forth in the February 4, 2015 Final Office Action. Applicants' representative explained that on the effective filing date of the subject application, a person having ordinary skill in the art ("POSA") could readily amplify and sequence tagged nucleic acid molecules as described in the specification and illustrated in Figure 8 (Exhibit A of the April 21, 2015 Agenda). Applicants' representative also directed the Examiner's attention to a schematic example of claim 1 (Exhibit B of the April 21, 2015 Agenda).

Examiner Lu stated that he understood with the examples, figures and schematic prepared. However, Examiner Lu requested applicants to amend claim 1 for clarity (A) to clarify that one unique tag differs from another unique tag; and (B) to recite "genomic nucleic acids" instead of "genomic material".

Accordingly, solely for the purpose of advancing prosecution, applicants have herein (A) amended claim 1 step (b) as suggested by the Examiner to recite "a substantially unique tag having a unique nucleotide sequence, so as to permit differentiation from another tag having a unique nucleotide sequence, and which tag is part of another tagged nucleic acid molecule"; and (B) amended the preamble of claim 1, claim 11 step (a), and claim 12 step (a) to recite "genomic nucleic acids".

Examiner Lu also requested applicants to amend claim 1 to clarify the each sequence read is generated from a tag and a portion of nucleic acid. In response, applicants' representatives directed the Examiner's attention to paragraph [0292] of the subject specification as published in which "sequence reads" of tagged nucleic acid molecules are defined to contain (1) a subsequence of a tag, and (2) a subsequence of a nucleic acid in the sample:

> "In the disclosed method, each sequence read generated from a tagged nucleic acid molecule comprises **two parts**, or subsequences. The **first part corresponds to the tag**, which identifies the tag (and the sample tag if samples are pooled). The **second part corresponds to a nucleic acid in the sample** which is mapped to a location in the genome (or to a transcript if RNA molecules are being counted)." (paragraph [0292], emphases added).

Therefore, the subject specification makes clear that each sequence read is generated from a tag and a portion of the nucleic acid.

After further discussion, during the April 21, 2015 Interview, Examiner Lu indicated that the recitation of "amplification distortion" in the pending claims may be unclear. In response, applicants' representatives explained that on the effective filing date of the subject application, a POSA would have understood the term "amplification distortion" (1) based on the description in paragraphs [0243] and [310] of the subject specification as published, and (2) based on knowledge in the art at the relevant time.

*Rejection of claims under 35 U.S.C. § 112, paragraph 2 (clarity)*

During the April 21, 2015 Interview, applicants' representatives respectfully requested clarification of the rejection of the claims under 35 U.S.C. § 112, second paragraph (clarity), set forth in the February 4, 2015 Final Office Action.

Examiner Lu stated that to address item 10 in the February 4, 2015 Final Office Action, applicants should amend claim 1 to recite antecedent basis making clear that each tagged sequence in claim 1 step (e) refers to the tagged sequences of step (d). Accordingly, applicants have amended claim 1 step (e) as suggested by the Examiner.

Examiner Lu also stated that to address item 13 in the February 4, 2015 Final Office Action, applicants should amend claim 9 to recite antecedent basis making clear that each tagged sequence in claim 9 step (d) refers to the tagged sequences of step (c). Accordingly, applicants have amended claim 9 step (d) as suggested by the Examiner.

Examiner Lu also stated that to address item 15 in the February 4, 2015 Final Office Action, applicants should amend claim 12 to make clear the relationship between claim 12 and claim 1. Accordingly, applicants have amended the preamble of claim 12 for clarity.

Examiner Lu indicated during the April 21, 2015 Interview that the he would consider applicants' amendments to the claims upon filing. This Amendment is being filed consistent with the April 21, 2015 Interview.

**Response to February 4, 2015 Final Office Action**

**Claim Objections**

In the February 4, 2015 Final Office Action, the Examiner alleged that claims 24 and 70 were unclear. The Examiner's rationale is set forth on page 2 of the February 4, 2015 Office Action.

*Applicants' reply*

In response, applicants respectfully disagree. However, in accordance with the April 21, 2015 Examiner Interview, applicants have amended claims 24 and 70 for clarity.

Accordingly, applicants respectfully request that this objection be withdrawn.

**Rejection of claims under 35 U.S.C. § 112, First Paragraph (Enablement)**

The February 4, 2015 Final Office Action rejected claims 1-7, 9, 11-13, 24, 26-28, 31 and 68-71 under 35 U.S.C. § 112, first paragraph, because the specification allegedly did not enable any person skilled in the art to which it pertains, or with which it is most clearly connected, to make and use the invention commensurate in scope with the claims. The Examiner's rationale is set forth on pages 2-14 of the February 4, 2015 Office Action.

*Applicants' reply*

In response, applicants respectfully disagree. However, in accordance with the April 21, 2015 Examiner Interview, as explained hereinabove, applicants have amended claims 1 and 9 without prejudice or disclaimer. The pending claims comply with 35 U.S.C. § 112, first paragraph (Enablement).

Accordingly, applicants respectfully request that this rejection be withdrawn.


IPR2025-01353; IPR2025-01355
CSHL EX2002

**Rejection of claims under 35 U.S.C. § 112, Second Paragraph (Clarity)**

The February 4, 2015 Final Office Action rejected claims 1-7, 9, 11-13, 24, 26-28, 31 and 68-71 under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which the applicants regard as the invention. The Examiner's rationale is set forth on pages 14-16 of the February 4, 2015 Office Action.

*Applicants' reply*

In response, applicants respectfully disagree. However, in accordance with the April 21, 2015 Examiner Interview, applicants have amended claims 1 and 9 as explained hereinabove, and amended claims 7 and 26 without prejudice or disclaimer. The pending claims comply with 35 U.S.C. § 112, second paragraph (Clarity).

Accordingly, applicants respectfully request that this rejection be withdrawn.

Applicants understand that all of the claims in the present application are now in order for allowance. Notice to this effect is respectfully requested.


IPR2025-01353; IPR2025-01355
CSHL EX2002

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicant's undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee, other than $100.00 for a one-month extension of time, is deemed necessary in connection with the filing of this response. However, if any fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39.992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

Minerva Y. Wong 6/4/2015
Minerva Y. Wong Date

**AGENDA FOR TELEPHONIC EXAMINER INTERVIEW**

**Rejections under 35 U.S.C. § 112, paragraph 1 (enablement)**

I **A person having ordinary skill in the art (POSA) would be enabled to amplify and sequence tagged nucleic acid molecules.**

In response to the Examiner's comments on pages 11-12 of the February 4, 2015 Final Office Action:

The Examiner's assertion that universal primers cannot be used with the claimed method is incorrect. Either specific or universal primers capable of hybridizing to sequences of the genomic material can be used in the claimed method.

Figure 8 of the subject specification, described in paragraphs [0040], [0041], [0042], [303], and [304] illustrates one embodiment of the method described in step (c). Figure 8 shows that following addition of terminal transferase, a poly-A tail is added to the ends of the segments of the genomic material to generate $0^{th}$ order derivative strands. A primer (PrA) comprising a tag ($N_{10}$) hybridizes to the poly-A tail added to the segment of the genomic material. Thus, in Figure 8, the hybridization of primer (PrA) with the poly-A tail is in effect analogous to the effect of using a universal primer. Yet, the method still works. A copy of Figure 8 is attached hereto as **Exhibit A.**

A schematic example of the claim is shown in **Exhibit B.**

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Exhibit 1



IPR2025-01353; IPR2025-01355
CSHL EX2002

**II** **A POSA could perform the claimed methods even if the amplification product of step (c) contained mutant tagged nucleic acid molecules, as alleged by the Examiner.**

In response to the Examiner's comments on pages 12-14 of the February 4, 2015 Final Office Action:

a) While mutations can occur in PCR, mutations were known to occur at a rate of $10^{-5}$ to $10^{-6}$ mutations/base pair (see i.e. Table 1 of Cline et al., Nucleic Acids Research 1996; attached hereto as **Exhibit C**).

b) Even if a mutation was introduced in 1 out of 100,000 base pairs, it was well understood in the art and well within the skill of a POSA to compensate for such mutation.

c) A POSA could 1) recognize an aberrant "mutant" subsequence during the mapping of step (e); and 2) discard the aberrant "mutant" subsequence during the counting of step (f).

d) The subject specification provides guidance for a POSA to account for unwanted mutations which might unexpectedly and rarely occur. Paragraph [0311] explains that:

> "tagged nucleic acid molecules that differ by a single nucleotide can (optionally) be ignored when counting".


IPR2025-01353; IPR2025-01355
CSHL EX2002

**III** **A "tag associated sequence read" is defined to require sequencing of the tag and at least a portion of the nucleic acid.**

In response to the Examiner's comments on page 14 of the February 4, 2015 Final Office Action:

a) "Sequence reads" when referring to tagged nucleic acid molecules are defined in the specification as containing (1) a subsequence of a tag, and (2) a subsequence of a nucleic acid in the sample:

> "In the disclosed method, each sequence read generated from a tagged nucleic acid molecule comprises **two parts**, or subsequences. The **first part corresponds to the tag**, which identifies the tag (and the sample tag if samples are pooled). The **second part corresponds to a nucleic acid in the sample** which is mapped to a location in the genome (or to a transcript if RNA molecules are being counted)." (paragraph [0292], emphases added).

b) Tag associated sequence reads are sequence reads associated with a tag.


IPR2025-01353; IPR2025-01355
CSHL EX2002

**Rejections under 35 U.S.C. § 112, paragraph 2 (clarity)**

In response to the Examiner's comments on pages 15-16 of the February 4, 2015 Final Office Action:

**I** Items 9, 11, 12 and 16

Applicants propose to amend claims 1, 7 and 26. A copy of the proposed amendments is attached hereto as **Exhibit D.**

**II** Items 10 and 13

Applicants would like clarification from the Examiner.

**III** Item 14

The subject specification in paragraph [0272] describes:

> "As used herein, the term 'substantially unique primers' refers to a plurality of primers, wherein each primer comprises a tag, and wherein at least 50% of the tags of the plurality are unique to the plurality of tags" (emphasis added).

Therefore, claim 9 step (a) requires that the tagged nucleic acid molecules have different tags.

**IV** Item 15

Applicants would like clarification from the Examiner. An embodiment of claim 12 is described in Figure 8.

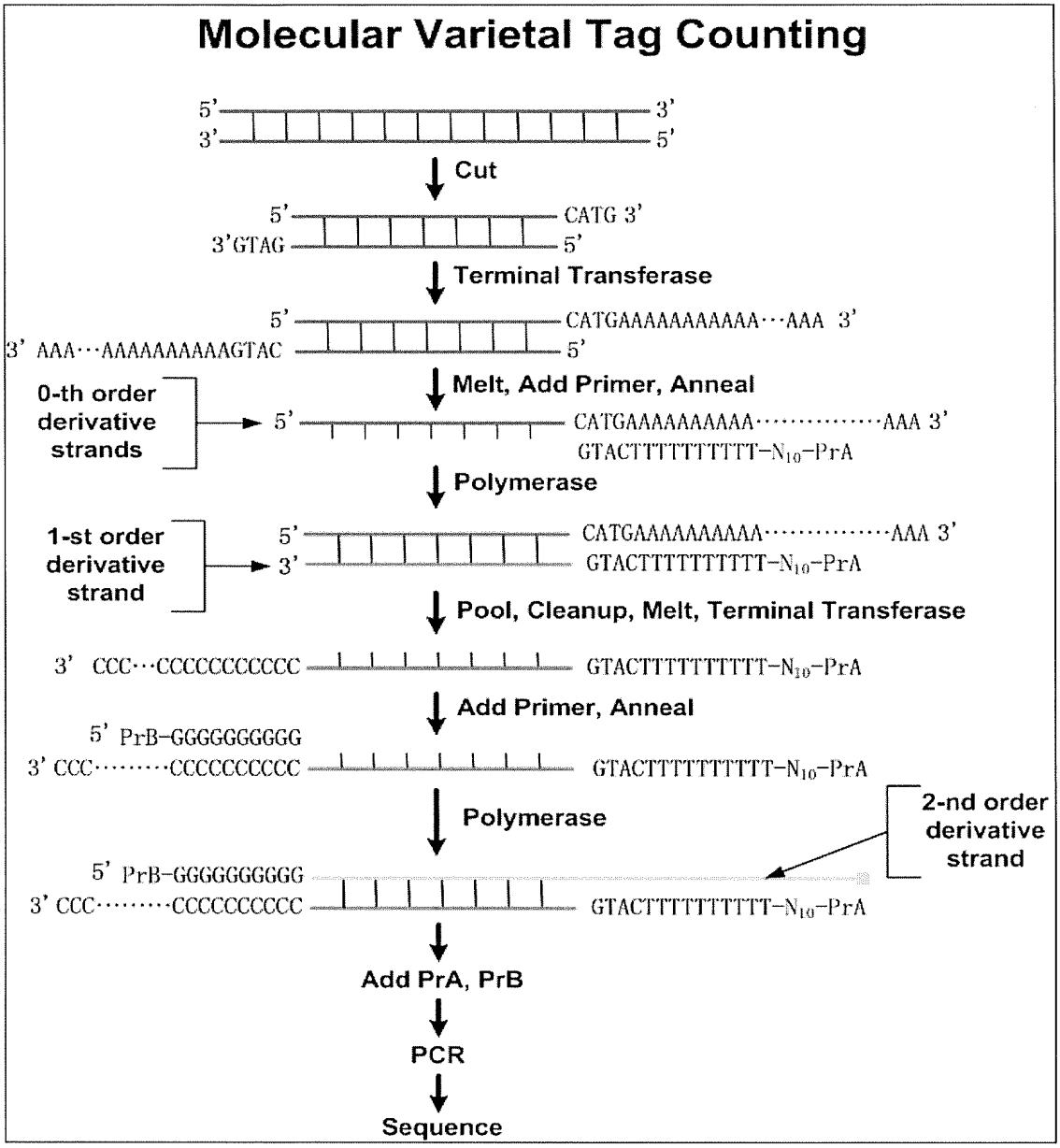



Fig. 8

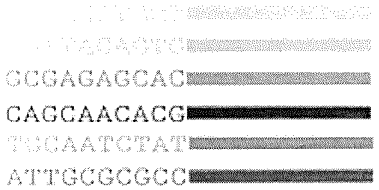
Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit A

## CLAIM 1

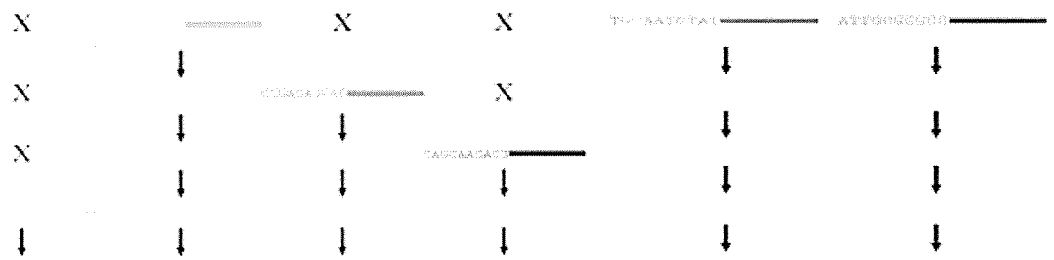
a) obtaining segments of genomic material



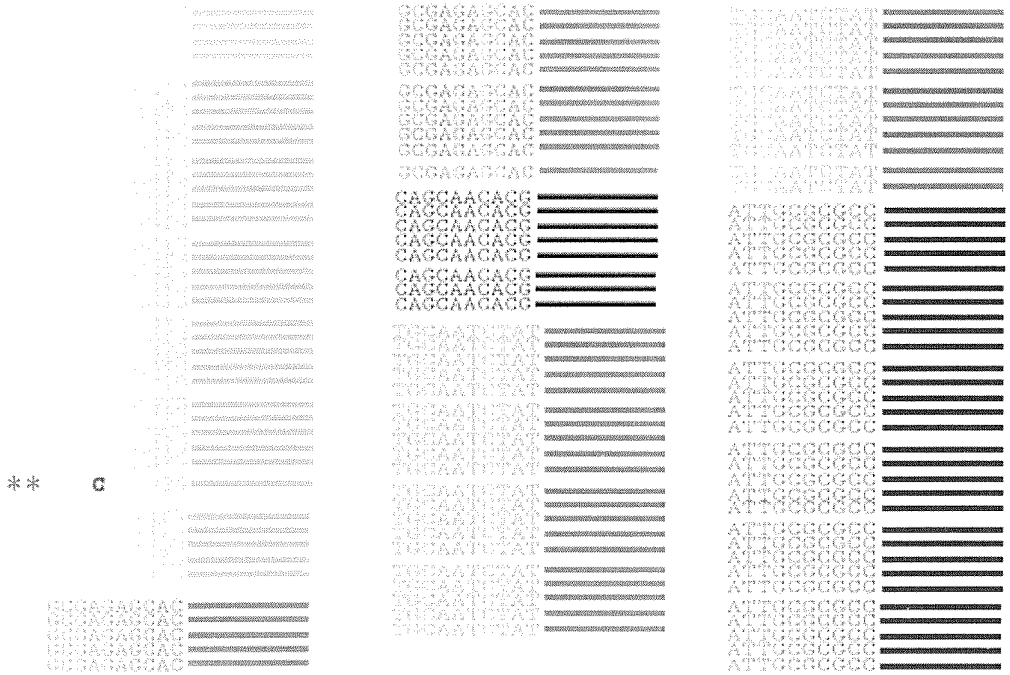
b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic material from step (a) and a substantially unique tag



c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR) *



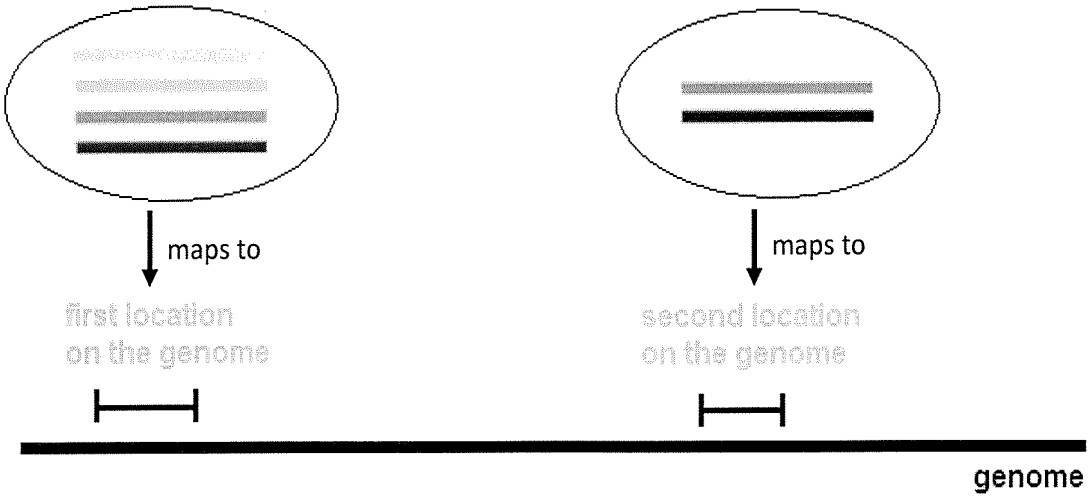
d) Generating tag associated sequence reads by sequencing the product of step (c)



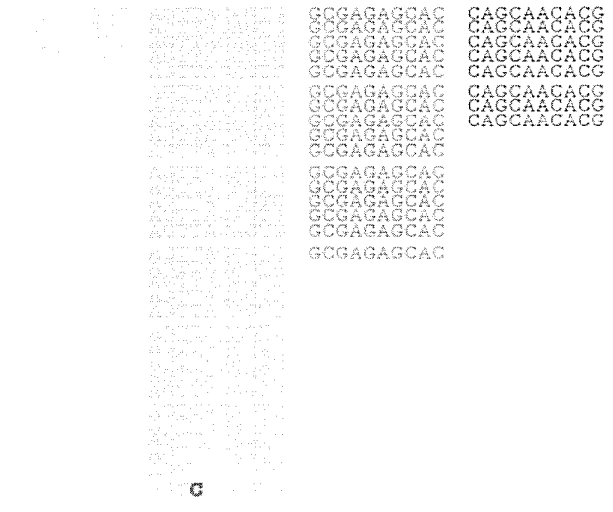

*Amplification distortion, refers to non-uniform amplification.

** A 1-base mismatch, highlighted in red, represents the 1/1000 frequency of sequence error in PCR

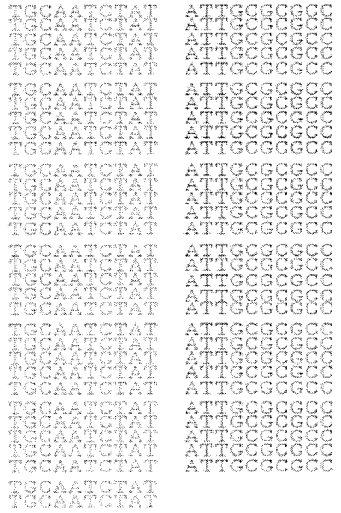
e) Assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read...




f) Counting the number of tagged nucleic acid molecules having a substantially unique tag that have been assigned to the same location on the genome,...



First location
Total unique tag counts = 4



Second location
Total unique tag counts = 2

Table 1. Average error rates of thermostable DNA polymerases during PCR[a]

| DNA polymerase | No. of PCRs | Target (ng) | Template doublings[b] | *lacI*⁻ plaques[c] (% ± SD) | Error rate[d] (×10⁻⁶ ± SD) |
|---|---|---|---|---|---|
| *Pfu* | 10 | 24 | 9.7 | 0.42 ± 0.08 | 1.3 ± 0.2 |
| | 2 | 2 | 12.7 | 0.30 ± 0.06[e,f] | 0.7 ± 0.1[e] |
| | 2 | 0.2 | 16.0 | 0.47 ± 0.03[e,f] | 0.8 ± 0.02[e] |
| | 2 | 0.02 | 19.4 | 0.66 ± 0.03[e,f] | 1.0 ± 0.04[e] |
| *Deep Vent* | 4 | 24 | 9.7–10 | 0.9 ± 0.1 | 2.7 ± 0.2 |
| *Vent* | 6 | 24 | 8.7–10 | 0.9 ± 0.3 | 2.8 ± 0.9 |
| *Taq* | 11 | 24 | 8.7–11 | 2.7 ± 1.1 | 8.0 ± 3.9 |
| *UITma* | 2 | 24 | 9.7 | 18.8 ± 0.8[e] | 55 ± 2[e] |

[a]PCRs were performed in each manufacturers' recommended buffer (all pH 8.8) in the presence of 200 μM each dNTP and 2 mM $MgSO_4$ (*Pfu*, *Deep Vent* and *Vent*), 2 mM $MgCl_2$ (*UITma*) or 1.5 mM $MgCl_2$ (*Taq*).

[b]Template doublings (*d*) were determined using the equation $2^d$ = (amount of PCR product)/( amount of starting target). 24 ng *lacI* target corresponds to 50 ng *lacIOZα* plasmid template. The range of *d* obtained is indicated.

[c]Mutant frequencies (*mf*) were determined by dividing the total number of blue plaques (*lacI*⁻ mutants) on X-gal plates by the total number of plaques containing a functional *lacZα* sequence (blue plaques on X-gal plus IPTG plates).

[d]Error rates were calculated using the equation $ER = mf/(bp \times d)$, where *mf* is the mutation frequency, *bp* is the number of detectable sites in *lacI* (=349; 15) and *d* is the number of template doublings.

[e]± indicates range of duplicate measurements.

[f]Mutation frequencies for *Pfu* amplification of 0.02–2 ng target were normalized such that the mean mutation frequency for *Pfu* amplification of 24 ng target (assay internal control) was 0.42%.

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit C



IPR2025-01353; IPR2025-01355
CSHL EX2002

Docket No. 5981/81503-A/GJG/MW

## PROPOSED AMENDMENTS TO THE CLAIMS

### Listing of Claims

1. (Currently Amended) A method for obtaining from genomic material genomic copy number information unaffected by amplification distortion, comprising:

a) obtaining segments of the genomic material;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic material from step (a) and a tag;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic material by mapping the subsequence of each tag associated sequence read corresponding to [[a]]one segment of the segments of the genomic material to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information unaffected by amplification distortion.

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Previously Presented) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising summing the counts obtained in step (f) for locations which comprise a region of the genome, [illegible] obtained from a reference sample for the same region of the genome, [illegible] thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Previously Presented) A method for obtaining from mRNA transcripts mRNA copy number information unaffected by amplification distortion, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information unaffected by amplification distortion.

10. (Cancelled)

11. (Previously Presented) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:


IPR2025-01353; IPR2025-01355
CSHL EX2002

a) adding a polynucleotide tail to the ends of the segments of the genomic material to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Previously Presented) The method of claim 1, wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic material to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands


IPR2025-01353; IPR2025-01355
CSHL EX2002

added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein adding [[a]]the polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to step (d).

25. (Cancelled)

26. (Currently Amended) The method of claim 1, wherein the tag further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag

prior to PCR amplification or prior to said sequencing of the product of step (c).

28. (Currently Amended) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 13278333 |
| **Filing Date:** | 21-Oct-2011 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | John P. White/Jason Keith |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |

Filed as Small Entity

**Filing Fees for Utility under 35 USC 111(a)**

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| **Extension-of-Time:** | | | | |

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| Extension - 1 month with $0 paid | 2251 | 1 | 100 | 100 |
| **Miscellaneous:** | | | | |
| | | | **Total in USD ($)** | **100** |


IPR2025-01353; IPR2025-01355
CSHL EX2002

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 22542789 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Jason Keith |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 04-JUN-2015 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 17:54:33 |
| **Application Type:** | Utility under 35 USC 111(a) |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | Deposit Account |
| Payment was successfully received in RAM | $100 |
| RAM confirmation Number | 16800 |
| Deposit Account | 033125 |
| Authorized User | GERSHIK, GARY J |

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

Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)



IPR2025-01353; IPR2025-01355
CSHL EX2002

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | | 150604_5981_81503_A_Response_MW.pdf | 983910<br>25d2d2e68729952520d2d8d0ca2b9bb00f9451ed | yes | 32 |

**Multipart Description/PDF files in .zip description**

| Document Description | Start | End |
|---|---|---|
| Response After Final Action | 1 | 1 |
| Claims | 2 | 9 |
| Applicant Arguments/Remarks Made in an Amendment | 10 | 10 |
| Applicant summary of interview with examiner | 11 | 13 |
| Applicant Arguments/Remarks Made in an Amendment | 14 | 32 |

**Warnings:**

**Information:**

| | | | | | |
|---|---|---|---|---|---|
| 2 | Fee Worksheet (SB06) | fee-info.pdf | 30752<br>4868d4e95153fc92b22a6e92992598b6d2ca347f | no | 2 |

**Warnings:**

**Information:**

| Total Files Size (in bytes): | 1014662 |
|---|---|

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

**New Applications Under 35 U.S.C. 111**

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

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

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

**New International Application Filed with the USPTO as a Receiving Office**

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

PTO/SB/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<br>Substitute for Form PTO-875 | Application or Docket Number<br>13/278,333 | Filing Date<br>10/21/2011 | ☐ 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 **06/04/2015**

| | (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)) | * 20 | Minus | ** 20 | = | X $ = | |
| Independent (37 CFR 1.16(h)) | * 2 | Minus | ***3 | = | X $ = | |
| ☐ Application Size Fee (37 CFR 1.16(s)) | | | | | | |
| ☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | |
| | | | | | TOTAL ADD'L FEE | |

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
/KATRINA . TURNER/

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.*



IPR2025-01353; IPR2025-01355
CSHL EX2002

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 06/11/2015

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER |
|---|
| LU, FRANK WEI MIN |

| ART UNIT | PAPER NUMBER |
|---|---|
| 1634 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 06/11/2015 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Applicant-Initiated Interview Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 13/278,333 | HICKS ET AL. | |
| | Examiner | Art Unit | |
| | FRANK LU | 1634 | |

All participants (applicant, applicant's representative, PTO personnel):

(1) *FRANK LU*. (3)_____.

(2) *Mr. Gary Gershik (Reg. No. 39,992)*. (4)_____.

Date of Interview: *21 April 2015*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☒112 ☐102 ☐103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *1-7, 9, 12, 13, 24, 26-28, 70, and 71*.

Identification of prior art discussed: _____.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

*Mr. Gershik and the examiner discussed the rejections and possible amendments. The examiner indicated that, although the specification defined "amplification distortion" as non-uniform amplification of template nucleic acid molecules, it is unclear what means amplification distortion in the claims and how the method recited in claim 1 can obtain genomic copy number information unaffected by amplification distortion and how the method recited in claim 9 can obtain mRNA copy number information unaffected by amplification distortion. If amplification distortion is well known in the art as argued by applicant, applicant is required to provide some documents to support his position.*

**Applicant recordation instructions:** The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

**Examiner recordation instructions**: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

| /FRANK LU/<br>Primary Examiner, Art Unit 1634 | April 21, 2015 |
|---|---|


IPR2025-01353, IPR2025-01355
CSHL EX2002

# Summary of Record of Interview Requirements

**Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record**
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

**Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews**
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

---

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

**Examiner to Check for Accuracy**

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

## DETAILED ACTION

### *Advisory Action*

1. The proposed amendments filed on June 4, 2015 have been fully considered but will not be entered because they raise new issues that would require further consideration and/or search.

***Response to Arguments***

I. In page 14, first to third paragraphs of applicant's remarks, applicant argues the proposed amendments have overcome all objections.

After carefully considering applicant's argument and the proposed amendments, the examiner agrees to withdraw all objections if the proposed amendments have been entered.

II. In page 14, fourth paragraph bridging to page 15, last paragraph of applicant's remarks, applicant argues the proposed amendments have overcome the rejections under 35 U.S.C. 112(a) and 112 (b).

The above argument has been fully considered but it is not persuasive toward the withdrawal of the rejections because applicant's argument with respect to claims 1-7, 9, 12, 13, 24, 26-28, 70, and 71 is moot since applicant has amended claims 1, 7, 9, 11, 12, 24, 26, and 70. The proposed amendments in claims 1, 7, 9, 11, 12, 24, 26, and 70 raise new issues that would require further consideration and/or search. For example, the phrase "each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic nucleic acids from step (a) and a substantially unique tag having a unique nucleotide sequence, so as to permit differentiation from another tag having a unique nucleotide sequence, and which tag is part of another tagged nucleic acid molecule" in claim 1 of the proposed amendments is a new

limitation and has not been found in claim 1 of the amendment filed on July 30, 2014 (the amendment used for the final rejection).

2. Papers related to this application may be submitted to Group 1600 by facsimile transmission. Papers should be faxed to Group 1600 via the PTO Fax Center. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993)(See 37 CAR § 1.6(d)). The CM Fax Center number is (571)273-8300.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Frank Lu, Ph.D., whose telephone number is (571)272-0746. The examiner can normally be reached on Monday-Friday from 9 A.M. to 5 P.M.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dave Nguyen, can be reached on (571)272-0731.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Frank W Lu /
Primary Examiner, Art Unit 1634
June 9, 2015

| **Advisory Action Before the Filing of an Appeal Brief** | **Application No.** 13/278,333 | **Applicant(s)** HICKS ET AL. | |
|---|---|---|---|
| | **Examiner** FRANK LU | **Art Unit** 1634 | **AIA (First Inventor to File) Status** No |

***--The MAILING DATE of this communication appears on the cover sheet with the correspondence address --***

THE REPLY FILED 04 June 2015 FAILS TO PLACE THIS APPLICATION IN CONDITION FOR ALLOWANCE.

NO NOTICE OF APPEAL FILED

1. ☒ The reply was filed after a final rejection. No Notice of Appeal has been filed. To avoid abandonment of this application, applicant must timely file one of the following replies: (1) an amendment, affidavit, or other evidence, which places the application in condition for allowance; (2) a Notice of Appeal (with appeal fee) in compliance with 37 CFR 41.31; or (3) a Request for Continued Examination (RCE) in compliance with 37 CFR 1.114 if this is a utility or plant application. Note that RCEs are not permitted in design applications. The reply must be filed within one of the following time periods:
   a) ☒ The period for reply expires *4* months from the mailing date of the final rejection.
   b) ☐ The period for reply expires on: (1) the mailing date of this Advisory Action; or (2) the date set forth in the final rejection, whichever is later. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of the final rejection.
   c) ☐ A prior Advisory Action was mailed more than 3 months after the mailing date of the final rejection in response to a first after-final reply filed within 2 months of the mailing date of the final rejection. The current period for reply expires months from the mailing date of *the prior Advisory Action* or SIX MONTHS from the mailing date of the final rejection, whichever is earlier.
   *Examiner Note*: If box 1 is checked, check either box (a), (b) or (c). ONLY CHECK BOX (b) WHEN THIS ADVISORY ACTION IS THE FIRST RESPONSE TO APPLICANT'S FIRST AFTER-FINAL REPLY WHICH WAS FILED WITHIN TWO MONTHS OF THE FINAL REJECTION. ONLY CHECK BOX (c) IN THE LIMITED SITUATION SET FORTH UNDER BOX (c). See MPEP 706.07(f).

Extensions of time may be obtained under 37 CFR 1.136(a). The date on which the petition under 37 CFR 1.136(a) and the appropriate extension fee have been filed is the date for purposes of determining the period of extension and the corresponding amount of the fee. The appropriate extension fee under 37 CFR 1.17(a) is calculated from: (1) the expiration date of the shortened statutory period for reply originally set in the final Office action; or (2) as set forth in (b) or (c) above, if checked. Any reply received by the Office later than three months after the mailing date of the final rejection, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

NOTICE OF APPEAL

2. ☐ The Notice of Appeal was filed on _____. A brief in compliance with 37 CFR 41.37 must be filed within two months of the date of filing the Notice of Appeal (37 CFR 41.37(a)), or any extension thereof (37 CFR 41.37(e)), to avoid dismissal of the appeal. Since a Notice of Appeal has been filed, any reply must be filed within the time period set forth in 37 CFR 41.37(a).

AMENDMENTS

3. ☒ The proposed amendments filed after a final rejection, but prior to the date of filing a brief, will not be entered because
   a) ☒ They raise new issues that would require further consideration and/or search (see NOTE below);
   b) ☐ They raise the issue of new matter (see NOTE below);
   c) ☐ They are not deemed to place the application in better form for appeal by materially reducing or simplifying the issues for appeal; and/or
   d) ☐ They present additional claims without canceling a corresponding number of finally rejected claims.
   NOTE: *see atatched office action*. (See 37 CFR 1.116 and 41.33(a)).
4. ☐ The amendments are not in compliance with 37 CFR 1.121. See attached Notice of Non-Compliant Amendment (PTOL-324).
5. ☐ Applicant's reply has overcome the following rejection(s): _____.
6. ☐ Newly proposed or amended claim(s) _____ would be allowable if submitted in a separate, timely filed amendment canceling the non-allowable claim(s).
7. ☒ For purposes of appeal, the proposed amendment(s): (a) ☒ will not be entered, or (b) ☐ will be entered, and an explanation of how the new or amended claims would be rejected is provided below or appended.

AFFIDAVIT OR OTHER EVIDENCE

8. ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.
9. ☐ The affidavit or other evidence filed after final action, but before or on the date of filing a Notice of Appeal will not be entered because applicant failed to provide a showing of good and sufficient reasons why the affidavit or other evidence is necessary and was not earlier presented. See 37 CFR 1.116(e).
10. ☐ The affidavit or other evidence filed after the date of filing the Notice of Appeal, but prior to the date of filing a brief, will not be entered because the affidavit or other evidence failed to overcome all rejections under appeal and/or appellant fails to provide a showing of good and sufficient reasons why it is necessary and was not earlier presented. See 37 CFR 41.33(d)(1).
11. ☐ The affidavit or other evidence is entered. An explanation of the status of the claims after entry is below or attached.

REQUEST FOR RECONSIDERATION/OTHER

12. ☐ The request for reconsideration has been considered but does NOT place the application in condition for allowance because: _____.
13. ☐ Note the attached Information *Disclosure Statement*(s). (PTO/SB/08) Paper No(s). _____
14. ☒ Other: Interview Summary .

STATUS OF CLAIMS

15. The status of the claim(s) is (or will be) as follows:
    Claim(s) allowed: .
    Claim(s) objected to: .
    Claim(s) rejected: 1-7,9,12,13,24,26-28,70 and 71.
    Claim(s) withdrawn from consideration: 11,68 and 69.

/FRANK LU/
Primary Examiner, Art Unit 1634


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Applicant-Initiated Interview Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 13/278,333 | HICKS ET AL. | |
| | Examiner | Art Unit | |
| | FRANK LU | 1634 | |

All participants (applicant, applicant's representative, PTO personnel):

(1) *FRANK LU*. (3)_____.

(2) *Mr. Gary Gershik (Reg. No. 39,992)*. (4)_____.

Date of Interview: *21 April 2015*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☒112 ☐102 ☐103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *1-7, 9, 12, 13, 24, 26-28, 70, and 71*.

Identification of prior art discussed: _____.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

*Mr. Gershik and the examiner discussed the rejections and possible amendments. The examiner indicated that, although the specification defined "amplification distortion" as non-uniform amplification of template nucleic acid molecules, it is unclear what means amplification distortion in the claims and how the method recited in claim 1 can obtain genomic copy number information unaffected by amplification distortion and how the method recited in claim 9 can obtain mRNA copy number information unaffected by amplification distortion. If amplification distortion is well known in the art as argued by applicant, applicant is required to provide some documents to support his position*.

**Applicant recordation instructions:** The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

**Examiner recordation instructions**: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

| /FRANK LU/<br>Primary Examiner, Art Unit 1634 | April 21, 2015 |
|---|---|


IPR2025-01353, IPR2025-01355
CSHL EX2002

# Summary of Record of Interview Requirements

**Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record**
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

**Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews**
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

---

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

**Examiner to Check for Accuracy**

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

DO NOT ENTER: /F.L./

30 Rockefeller Plaza
20th Floor
New York, New York 10112
June 4, 2015

BY EFS
Commissioner for Patents
Alexandria, VA

**AMENDMENT UNDER 37 C.F.R. § 1.116 IN RESPONSE TO FEBRUARY 4, 2015 FINAL OFFICE ACTION, PETITION FOR ONE-MONTH EXTENSION OF TIME, AND SUMMARY OF APRIL 21, 2015 EXAMINER INTERVIEW**

This Amendment is a submission under 37 C.F.R. § 1.116 and responds to the issues raised in a February 4, 2015 Final Office Action issued by the U.S. Patent and Trademark Office in connection with the above-identified application. A response to the February 4, 2015 Final Office Action was due May 4, 2015. Applicants hereby petition for a one-month extension of time. The fee for a one-month extension of time for a small entity is ONE HUNDRED DOLLARS ($100.00), and authorization is hereby given to charge this amount of fee to Deposit Account No. 03-3125. With a one-month extension of time, a response is due June 4, 2015. Accordingly, this response is being timely filed.

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

**Remarks** begin on page **10** of this paper.

A **Summary of April 21, 2015 Examiner Interview** begins on page **11** of this paper.

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| Applicant | : | Cold Spring Harbor Laboratory |
| Inventors | : | James Hicks, Nicholas Navin, Jennifer Troge, Zihua Wang, and Michael Wigler |
| U.S. Serial No. | : | 13/278,333 Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 Art Unit: 1634 |
| Confirmation No. | : | 4469 |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
August 4, 2015

**VIA EFS-WEB**
Commissioner for Patents
Alexandria, VA 22313-1450

**AMENDMENT UNDER 37 C.F.R. § 1.116 ACCOMPANYING A REQUEST FOR CONSIDERATION UNDER THE AFTER FINAL CONSIDERATION PILOT PROGRAM 2.0 AND SUBMISSION OF A DECLARATION UNDER 37 C.F.R. § 1.131 IN RESPONSE TO THE EXAMINER, REQUEST TO UPDATE THE NAME OF THE APPLICANT PURSUANT TO 37 C.F.R. § 1.46(c), AND REQUEST FOR CORRECTED FILING RECEIPT**

This Amendment under 37 C.F.R. § 1.116 accompanying a Request for Consideration Under the After Final Consideration Pilot Program 2.0 and submission of a Declaration under 37 C.F.R. § 1.131 are being submitted at the request of the Examiner.

Please amend the subject application as follows:

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

**Remarks** begin on page **10** of this paper.

A **Summary of July 28, 2015 Examiner Interview** begins on page **11** of this paper.

A **Summary of July 30, 2015 Examiner Interview** begins on page **12** of this paper.


IPR2025-01353; IPR2025-01355
CSHL EX2002

**AMENDMENTS TO THE CLAIMS**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims**

1. (Currently Amended) A method for obtaining from genomic ~~material~~ nucleic acids genomic copy number information ~~unaffected by amplification distortion~~, comprising:

a) obtaining segments of the genomic ~~material~~ nucleic acids;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic ~~material~~ nucleic acids from step (a) and a substantially unique tag having a unique nucleotide sequence, so as to permit differentiation from another tag having a unique nucleotide sequence, and which tag is part of another tagged nucleic acid molecule;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic ~~material~~ nucleic acids by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (d) corresponding


IPR2025-01353; IPR2025-01355
CSHL EX2002

to one segment of the segments of the genomic ~~material~~ nucleic acids to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information ~~unaffected by amplification distortion~~.

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Previously Presented) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is

comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising i) summing the counts obtained in step (f) for locations which comprise a region of the genome~~, and comparing the sum to a sum~~; ii) summing the counts obtained from a reference sample for the same region of the genome[[,]]; and iii) comparing the sums

obtained in step (i) and (ii), thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Currently Amended) A method for obtaining from mRNA transcripts mRNA copy number information ~~unaffected by amplification distortion~~, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (c) corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information ~~unaffected by amplification distortion~~.

10. (Cancelled)

11. (Currently Amended) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic ~~material~~ nucleic acids to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;


IPR2025-01353; IPR2025-01355
CSHL EX2002

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Currently Amended) The method of claim 1 step (b), wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic ~~material~~ nucleic acids to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide

sequences than the primers of step (b), and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein adding the polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to ~~sequencing in~~ step (d).

25. (Cancelled)

26. (Currently Amended) The method of claim 1, wherein the substantially unique tags ~~sequences~~ further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to said sequencing of the product of step (c).

28. (Previously Presented) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.


IPR2025-01353; IPR2025-01355
CSHL EX2002

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.


IPR2025-01353; IPR2025-01355
CSHL EX2002

**REMARKS**

Claims 1-7, 9, 11-13, 24, 26-28 and 68-71 are pending in the subject application. Applicant has amended claims 1, 7, 9, 11, 12, 24, 26, and 70 for clarity at the Examiner's request. Accordingly, upon entry of this Amendment claims 1-7, 9, 11-13, 24, 26-28 and 68-71 will be pending and under examination in the subject application.

The claim amendments herein are appropriate for entry under 37 C.F.R. § 1.116 because the amendments (1) comply with requirements of form expressly set forth in the February 4, 2015 Final Office Action, and (2) do not raise new issues that would require further consideration and/or search. Specifically, claims 1, 7, 9, 11, 12, 24, 26, and 70 were amended for clarity in response to suggestions by Examiner Lu during an April 21, 2015 Examiner Interview and a July 28, 2015 Examiner Interview, and do not introduce any new features which would require additional search. A Summary of April 21, 2015 Examiner Interview was filed in applicant's June 4, 2015 Response to February 4, 2015 Office Action. A Summary of the July 28, 2015 Examiner Interview and Summary of a July 30, 2015 Examiner Interview are below. Therefore, the amendments are proper and must be entered.


IPR2025-01353; IPR2025-01355
CSHL EX2002

**Summary of July 28, 2015 Examiner Interview**

On July 28, 2015, a telephonic Examiner Interview was conducted in connection with the subject application. Present at the interview were Examiner Frank Wei Min Lu of the U.S. Patent and Trademark Office, and applicant's undersigned representative. Applicant acknowledges with appreciation the courtesy that Examiner Lu extended during the July 28, 2015 Interview.

During the July 28, 2015 Telephonic Examiner Interview, the Examiner indicated removal of the term "amplification distortion" from the preamble of claims 1 and 9 is likely to overcome the Examiner's remaining clarity issues. In addition, the Examiner for the first time during the July 28, 2015 Interview cited U.S. Patent No. 8,741,606 ("the '606 Patent") and requested applicant to explain whether the disclosure of this patent presents any patentability issues for applicant's claims.

**Summary of July 30, 2015 Examiner Interview**

On July 30, 2015, a telephonic Examiner Interview was conducted in connection with the subject application. Present at the interview were Examiner Frank Wei Min Lu of the U.S. Patent and Trademark Office, and applicant's undersigned representative. Applicant acknowledges with appreciation the courtesy that Examiner Lu extended during the July 30, 2015 Interview.

On July 29, 2015, applicant sent an Agenda to Examiner Lu by facsimile. A copy of the July 29, 2015 Agenda is attached hereto as **Exhibit A.**

During the July 30, 2015 Interview, applicant's representative explained that because the '606 Patent published after the October 22, 2010 earliest filing date of applicant's application, the '606 Patent cannot be used in any obviousness rejection under 35 U.S.C. § 103.

During the July 30, 2015 Interview, applicant's representative also explained that the content of the earliest priority document of the '606 Patent, i.e. U.S. Provisional Patent Application No. 61/385,001, filed September 21, 2010 (hereinafter "Casbon et al. Provisional"), did not anticipate Applicant's claims.

Specifically, applicant's representative explained that the Casbon et al. Provisional does not disclose a method for "obtaining from genomic nucleic acids genomic copy number information", let alone each and every step of the method as recited in proposed amended independent claim 1 (from which claims 2-7, 11-13, 24, 27-28, 31, and 68-71 depend) comprising:

i) "obtaining segments of the genomic nucleic acids";
ii) "tagging the segments with substantially unique tags to generate

tagged nucleic acid molecules"; and

iii) "assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic nucleic acids by mapping the subsequence of each tag associated sequence read";

as recited in applicant's independent claim 1 (emphasis added).

Furthermore, applicant's representative explained that the Casbon et al. Provisional does not disclose a method for "obtaining from mRNA transcripts mRNA copy number information", let alone each and every step of the method as recited in proposed amended independent claim 9 comprising:

i) "subjecting mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement"; and

ii) "assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read",

as recited in applicant's independent claim 9.

Following the July 30, 2015 Interview, Examiner Lu indicated during a July 30, 2015 telephone voice message to applicant's undersigned representative that applicant should submit a declaration establishing prior invention to avoid the need for substantive review of the content of the Casbon et al. Provisional. This Amendment, including the accompanying Declaration under 37 CFR § 1.131, is being filed in response to the July 30, 2015 telephone voice message from the Examiner.

Applicant attaches hereto as **Exhibit B**, a Declaration Under 37 CFR § 1.131 showing that the inventors conceived of an embodiment of the


IPR2025-01353; IPR2025-01355
CSHL EX2002

claimed invention prior to the filing date of the Casbon et al. Provisional, i.e., prior to September 21, 2010.

**Request to Update the Name of the Applicant**

Applicants hereby request under 37 C.F.R. § 1.46(c) that the applicant of the subject application be changed to COLD SPRING HARBOR LABORATORY ("CSHL"). CSHL is the assignee of record of the entire right, title, and interest in the subject application.

37 C.F.R. § 1.46(c) requires that a request to change the applicant include an Application Data Sheet under 37 C.F.R. § 1.76 specifying the applicant in the applicant information section and comply with 37 C.F.R. §§ 3.71 and 3.73.

Accordingly, a corrected Application Data Sheet (ADS) identifying CSHL as the applicant of the above-identified application accompanies this Request.

CSHL is the assignee of the entire right, title, and interest in the above-identified application pursuant to 37 C.F.R. § 3.73(c). As assignee of the above-identified application, CSHL is entitled to conduct prosecution of the subject application as applicant pursuant to 37 C.F.R. § 3.71.

**Request for Corrected Filing Receipt**

Applicant respectfully requests that the United States Patent and Trademark Office issue a corrected filing receipt identifying "COLD SPRING HARBOR LABORATORY" as applicant.


IPR2025-01353; IPR2025-01355
CSHL EX2002

**Concluding Remarks**

Applicant believes the remarks presented hereinabove to be fully responsive to all of the grounds of rejection raised by the Examiner. In view of these remarks, applicant respectfully submits that all of the claims in the present application are now in order for allowance. Notice to this effect is respectfully requested.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicant's undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this Amendment. However, if any additional fee is required, authorization is hereby given to charge the amount of any such additional fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik

Gary J. Gershik
Registration No. 39.992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

Minerva Y. Wong 8/4/2015
Minerva Y. Wong / Date



IPR2025-01353; IPR2025-01355
CSHL EX2002

**AGENDA FOR TELEPHONIC EXAMINER INTERVIEW**

During a July 28, 2015 Telephonic Examiner Interview, the Examiner indicated he would consider applicants' proposal to remove the term "amplification distortion" from the preamble of claims 1 and 9, if applicants submit an analysis of subject matter disclosed in U.S. Patent No. 8,741,606.

Applicants claim benefit of U.S. Provisional Patent Application Nos. 61/510,579, filed July 22, 2011, and 61/406,067, filed October 22, 2010. Proposed amended independent claims 1 and 9 (from which all other pending claims depend) are supported by U.S. Provisional Patent Application No. 61/406,067, filed October 22, 2010.

U.S. Patent No. 8,741,606, claims benefit of U.S. Provisional Patent Application Nos. 61/432,119, filed January 12, 2011, and 61/385,001, filed September 21, 2010. Accordingly, only U.S. Provisional Patent Application No. 61/385,001 (hereinafter "Casbon et al"), filed September 21, 2010, can be considered as prior art. However, the subject matter of Casbon et al was not available to the public as of October 22, 2010, therefore Casbon et al cannot be used in any obviousness rejection under 35 U.S.C. § 103. A copy of Casbon et al is attached hereto as **Exhibit A.** A copy of proposed draft claim amendments are attached hereto as **Exhibit B.**

## I Casbon et al does not disclose each and every element of the claims.

Casbon et al does not anticipate the claimed invention at least because it does not disclose each and every element of the claims. Pursuant to M.P.E.P. § 2131, "'A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference.' *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987)" and "'The identical invention must be shown in as


IPR2025-01353; IPR2025-01355
CSHL EX2002

complete detail as is contained in the ... claim.' *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1236, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989). The elements must be arranged as required by the claim . . . ."

**II Claim 1 is not anticipated by Casbon et al.**

Casbon et al discloses that "a tag or combination of tags associated with a polynucleotide whose identity (e.g., the tag DNA sequence) can be used to differentiate polynucleotides in a sample" (page 7, lines 18-20). Casbon et al further discloses that "evaluating the number of different DBR sequences present in a sequencing run allows the establishment of the minimum number of different starting polynucleotides that have been sequenced" (page 20, lines 25-27).

However, Casbon et al does not disclose a method for "obtaining from genomic nucleic acids genomic copy number information", let alone each and every step of the method as recited in proposed amended independent claim 1 (from which claims 2-7, 11-13, 24, 27-28, 31, and 68-71 depend) comprising:

i) "obtaining segments of the genomic nucleic acids";
ii) "tagging the segments with substantially unique tags to generate tagged nucleic acid molecules"; and
iii) "assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic nucleic acids by mapping the subsequence of each tag associated sequence read";

as recited in independent claim 1 (emphasis added).

Accordingly, Casbon et al does not anticipate independent claim 1.

**III Claim 9 is not anticipated by Casbon et al.**

Casbon et al does not disclose a method for "obtaining from mRNA transcripts mRNA copy number information", let alone each and every step of the method as recited in proposed amended independent claim 9 comprising:


IPR2025-01353; IPR2025-01355
CSHL EX2002

i) "subjecting mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement"; and

ii) "assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read",

as recited in independent claim 9.

Accordingly, Casbon et al does not anticipate independent claim 9.


IPR2025-01353; IPR2025-01355
CSHL EX2002

# INCREASING CONFIDENCE OF ALLELE CALLS WITH MOLECULAR COUNTING

## BACKGROUND

Genotyping is an important technique in genetic research for mapping a genome and localizing genes that are linked to inherited characteristics (e.g., genetic diseases). The genotype of a subject generally includes determining alleles for one or more genomic locus based on sequencing data obtained from the subject's DNA. Diploid genomes (e.g., human genomes) may be classified as, for example, homozygous or heterozygous at a genomic locus depending on the number of different alleles they possess for that locus, where heterozygous individuals have two different alleles for a locus and homozygous individuals have two copies of the same allele for the locus.

In genotyping analysis of diploid genomes, the coverage (number of sequencing reads) for a particular genomic locus is used to establish the confidence of an allele call. However, confidence in allele calling is significantly reduced when bias is introduced during sample preparation, e.g., when the starting sample is in limiting amounts and/or when one or more amplification reactions are employed to prepare the sample for sequencing. Thus, in samples having limited amounts of DNA, one may see high coverage (i.e., a high number of sequencing reads) for an allele on one chromosome over the allele on a different chromosome due to amplification bias (e.g., amplification from only a few, or even one, polynucleotide molecule). In this case, coverage alone may be misleading when measuring confidence in an allele call.

The present invention finds use in increasing the confidence in allele calling as well as in other applications based on nucleic acid sequence analysis.

## SUMMARY OF THE INVENTION

Aspects of the present invention include methods and compositions for establishing the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. In these aspects of the invention, a degenerate base region (DBR) is attached to the starting polynucleotide molecules that



Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Examiner Interview Agenda Exhibit A



IPR2020-01353, IPR2020-01355
CSHL EX2002

are subsequently sequenced (e.g., after certain process steps are performed, e.g., amplification and/or enrichment). The number of different DBR sequences present in a sequencing run establishes the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. DBRs can be used to improve the analysis of many different nucleic acid sequencing applications. For example, DBRs enable the determination of a statistical value for an allele call in genotyping assays that cannot be derived from the read number alone.

## BRIEF DESCRIPTION OF THE DRAWINGS

The invention is best understood from the following detailed description when read in conjunction with the accompanying drawings. Included in the drawings are the following figures:

Figure 1 shows the allele ratio for each MID in samples prepared from the indicated amount of starting material (top of each panel; in nanograms).

Figure 2 shows the fraction of DBR sequences for each MID associated with each allele at a synthetic polymorphic position. Samples were prepared from the indicated amount of starting material (top of each panel; in nanograms).

## DEFINITIONS

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Still, certain elements are defined for the sake of clarity and ease of reference.

Terms and symbols of nucleic acid chemistry, biochemistry, genetics, and molecular biology used herein follow those of standard treatises and texts in the field, e.g. Kornberg and Baker, DNA Replication, Second Edition (W.H. Freeman, New York, 1992); Lehninger, Biochemistry, Second Edition (Worth Publishers, New York, 1975); Strachan and Read, Human Molecular Genetics, Second Edition (Wiley-Liss, New York, 1999); Eckstein, editor, Oligonucleotides and Analogs: A Practical Approach (Oxford

University Press, New York, 1991); Gait, editor, Oligonucleotide Synthesis: A Practical Approach (IRL Press, Oxford, 1984); and the like.

"Amplicon" means the product of a polynucleotide amplification reaction. That is, it is a population of polynucleotides, usually double stranded, that are replicated from one or more starting sequences. The one or more starting sequences may be one or more copies of the same sequence, or it may be a mixture of different sequences. Amplicons may be produced by a variety of amplification reactions whose products are multiple replicates of one or more target nucleic acids. Generally, amplification reactions producing amplicons are "template-driven" in that base pairing of reactants, either nucleotides or oligonucleotides, have complements in a template polynucleotide that are required for the creation of reaction products. In one aspect, template-driven reactions are primer extensions with a nucleic acid polymerase or oligonucleotide ligations with a nucleic acid ligase. Such reactions include, but are not limited to, polymerase chain reactions (PCRs), linear polymerase reactions, nucleic acid sequence-based amplification (NASBAs), rolling circle amplifications, and the like, disclosed in the following references that are incorporated herein by reference: Mullis et al, U.S. patents 4,683,195; 4,965,188; 4,683,202; 4,800,159 (PCR); Gelfand et al, U.S. patent 5,210,015 (real-time PCR with "TAQMAN™" probes); Wittwer et al, U.S. patent 6,174,670; Kacian et al, U.S. patent 5,399,491 ("NASBA"); Lizardi, U.S. patent 5,854,033; Aono et al, Japanese patent publ. JP 4-262799 (rolling circle amplification); and the like. In one aspect, amplicons of the invention are produced by PCRs. An amplification reaction may be a "real-time" amplification if a detection chemistry is available that permits a reaction product to be measured as the amplification reaction progresses, e.g. "real-time PCR" described below, or "real-time NASBA" as described in Leone et al, Nucleic Acids Research, 26: 2150-2155 (1998), and like references. As used herein, the term "amplifying" means performing an amplification reaction. A "reaction mixture" means a solution containing all the necessary reactants for performing a reaction, which may include, but not be limited to, buffering agents to maintain pH at a selected level during a reaction, salts, co-factors, scavengers, and the like.

The term "assessing" includes any form of measurement, and includes determining if an element is present or not. The terms "determining", "measuring",

"evaluating", "assessing" and "assaying" are used interchangeably and includes quantitative and qualitative determinations. Assessing may be relative or absolute. "Assessing the presence of" includes determining the amount of something present, and/or determining whether it is present or absent. As used herein, the terms "determining," "measuring," and "assessing," and "assaying" are used interchangeably and include both quantitative and qualitative determinations.

Polynucleotides that are "asymmetrically tagged" have left and right adapter domains that are not identical. This process is referred to generically as attaching adapters asymmetrically or asymmetrically tagging a polynucleotide, e.g., a polynucleotide fragment. Production of polynucleotides having asymmetric adapter termini may be achieved in any convenient manner. Exemplary asymmetric adapters are described in: U.S. Patents 5,712,126 and 6,372,434; U.S. Patent Publications 2007/0128624 and 2007/0172839; and PCT publication WO/2009/032167; all of which are incorporated by reference herein in their entirety. In certain embodiments, the asymmetric adapters employed are those described in U.S. Patent Application Ser. No. 12/432,080, filed on April 29, 2009, incorporated herein by reference in its entirety.

As one example, a user of the subject invention may use an asymmetric adapter to tag polynucleotides. An "asymmetric adapter" is one that, when ligated to both ends of a double stranded nucleic acid fragment, will lead to the production of primer extension or amplification products that have non-identical sequences flanking the genomic insert of interest. The ligation is usually followed by subsequent processing steps so as to generate the non-identical terminal adapter sequences. For example, replication of an asymmetric adapter attached fragment(s) results in polynucleotide products in which there is at least one nucleic acid sequence difference, or nucleotide/nucleoside modification, between the terminal adapter sequences. Attaching adapters asymmetrically to polynucleotides (e.g., polynucleotide fragments) results in polynucleotides that have one or more adapter sequences on one end (e.g., one or more region or domain, e.g., a primer binding site) that are either not present or have a different nucleic acid sequence as compared to the adapter sequence on the other end. It is noted that an adapter that is termed an "asymmetric adapter" is not necessarily itself structurally asymmetric, nor does the mere act of attaching an asymmetric adapter to a polynucleotide

fragment render it immediately asymmetric. Rather, an asymmetric adapter-attached polynucleotide, which has an identical asymmetric adapter at each end, produces replication products (or isolated single stranded polynucleotides) that are asymmetric with respect to the adapter sequences on opposite ends (e.g., after at least one round of amplification/primer extension).

Any convenient asymmetric adapter, or process for attaching adapters asymmetrically, may be employed in practicing the present invention. Exemplary asymmetric adapters are described in: U.S. Patents 5,712,126 and 6,372,434; U.S. Patent Publications 2007/0128624 and 2007/0172839; and PCT publication WO/2009/032167; all of which are incorporated by reference herein in their entirety. In certain embodiments, the asymmetric adapters employed are those described in U.S. Patent Application Ser. No. 12/432,080, filed on April 29, 2009, incorporated herein by reference in its entirety.

"Complementary" or "substantially complementary" refers to the hybridization or base pairing or the formation of a duplex between nucleotides or nucleic acids, such as, for instance, between the two strands of a double stranded DNA molecule or between an oligonucleotide primer and a primer binding site on a single stranded nucleic acid. Complementary nucleotides are, generally, A and T (or A and U), or C and G. Two single stranded RNA or DNA molecules are said to be substantially complementary when the nucleotides of one strand, optimally aligned and compared and with appropriate nucleotide insertions or deletions, pair with at least about 80% of the nucleotides of the other strand, usually at least about 90% to 95%, and more preferably from about 98 to 100%. Alternatively, substantial complementarity exists when an RNA or DNA strand will hybridize under selective hybridization conditions to its complement. Typically, selective hybridization will occur when there is at least about 65% complementary over a stretch of at least 14 to 25 nucleotides, preferably at least about 75%, more preferably at least about 90% complementary. See, M. Kanehisa Nucleic Acids Res. 12:203 (1984), incorporated herein by reference.

"Duplex" means at least two oligonucleotides and/or polynucleotides that are fully or partially complementary undergo Watson-Crick type base pairing among all or most of their nucleotides so that a stable complex is formed. The terms "annealing" and

"hybridization" are used interchangeably to mean the formation of a stable duplex. "Perfectly matched" in reference to a duplex means that the poly- or oligonucleotide strands making up the duplex form a double stranded structure with one another such that every nucleotide in each strand undergoes Watson-Crick base pairing with a nucleotide in the other strand. A stable duplex can include Watson-Crick base pairing and/or non-Watson-Crick base pairing between the strands of the duplex (where base pairing means the forming hydrogen bonds). In certain embodiments, a non-Watson-Crick base pair includes a nucleoside analog, such as deoxyinosine, 2, 6-diaminopurine, PNAs, LNA's and the like. In certain embodiments, a non-Watson-Crick base pair includes a "wobble base", such as deoxyinosine, 8-oxo-dA, 8-oxo-dG and the like, where by "wobble base" is meant a nucleic acid base that can base pair with a first nucleotide base in a complementary nucleic acid strand but that, when employed as a template strand for nucleic acid synthesis, leads to the incorporation of a second, different nucleotide base into the synthesizing strand (wobble bases are described in further detail below). A "mismatch" in a duplex between two oligonucleotides or polynucleotides means that a pair of nucleotides in the duplex fails to undergo Watson-Crick bonding.

"Genetic locus," "locus," or "locus of interest" in reference to a genome or target polynucleotide, means a contiguous sub-region or segment of the genome or target polynucleotide. As used herein, genetic locus, locus, or locus of interest may refer to the position of a nucleotide, a gene or a portion of a gene in a genome, including mitochondrial DNA or other non-chromosomal DNA (e.g., bacterial plasmid), or it may refer to any contiguous portion of genomic sequence whether or not it is within, or associated with, a gene. A genetic locus, locus, or locus of interest can be from a single nucleotide to a segment of a few hundred or a few thousand nucleotides in length or more. In general, a locus of interest will have a reference sequence associated with it (see description of "reference sequence" below).

"Kit" refers to any delivery system for delivering materials or reagents for carrying out a method of the invention. In the context of reaction assays, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (e.g., probes, enzymes, etc. in the appropriate containers) and/or supporting materials (e.g., buffers, written instructions for performing the assay etc.) from one

location to another. For example, kits include one or more enclosures (e.g., boxes) containing the relevant reaction reagents and/or supporting materials. Such contents may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second container contains probes.

"Ligation" means to form a covalent bond or linkage between the termini of two or more nucleic acids, e.g. oligonucleotides and/or polynucleotides. The nature of the bond or linkage may vary widely and the ligation may be carried out enzymatically or chemically. As used herein, ligations are usually carried out enzymatically to form a phosphodiester linkage between a 5' carbon of a terminal nucleotide of one oligonucleotide with 3' carbon of another oligonucleotide. A variety of template-driven ligation reactions are described in the following references, which are incorporated by reference: Whiteley et al, U.S. patent 4,883,750; Letsinger et al, U.S. patent 5,476,930; Fung et al, U.S. patent 5,593,826; Kool, U.S. patent 5,426,180; Landegren et al, U.S. patent 5,871,921; Xu and Kool, Nucleic Acids Research, 27: 875-881 (1999); Higgins et al, Methods in Enzymology, 68: 50-71 (1979); Engler et al, The Enzymes, 15: 3-29 (1982); and Namsaraev, U.S. patent publication 2004/0110213.

"Multiplex Identifier" (MID) as used herein refers to a tag or combination of tags associated with a polynucleotide whose identity (e.g., the tag DNA sequence) can be used to differentiate polynucleotides in a sample. In certain embodiments, the MID on a polynucleotide is used to identify the source from which the polynucleotide is derived. For example, a nucleic acid sample may be a pool of polynucleotides derived from different sources, (e.g., polynucleotides derived from different individuals, different tissues or cells, or polynucleotides isolated at different times points), where the polynucleotides from each different source are tagged with a unique MID. As such, a MID provides a correlation between a polynucleotide and its source. In certain embodiments, MIDs are employed to uniquely tag each individual polynucleotide in a sample. Identification of the number of unique MIDs in a sample can provide a readout of how many individual polynucleotides are present in the sample (or from how many original polynucleotides a manipulated polynucleotide sample was derived; see, e.g., U.S. Patent No. 7,537,897, issued on May 26, 2009, incorporated herein by reference in its

entirety). MIDs are typically comprised of nucleotide bases and can range in length from 2 to 100 nucleotide bases or more and may include multiple subunits, where each different MID has a distinct identity and/or order of subunits. Exemplary nucleic acid tags that find use as MIDs are described in U.S. Patent 7,544,473, issued on June 6, 2009, and titled "Nucleic Acid Analysis Using Sequence Tokens", as well as U.S. Patent 7,393,665, issued on July 1, 2008, and titled "Methods and Compositions for Tagging and Identifying Polynucleotides", both of which are incorporated herein by reference in their entirety for their description of nucleic acid tags and their use in identifying polynucleotides. In certain embodiments, a set of MIDs employed to tag a plurality of samples need not have any particular common property (e.g., Tm, length, base composition, etc.), as the methods described herein can accommodate a wide variety of unique MID sets. It is emphasized here that MIDs need only be unique within a given experiment. Thus, the same MID may be used to tag a different sample being processed in a different experiment. In addition, in certain experiments, a user may use the same MID to tag a subset of different samples within the same experiment. For example, all samples derived from individuals having a specific phenotype may be tagged with the same MID, e.g., all samples derived from control (or wildtype) subjects can be tagged with a first MID while subjects having a disease condition can be tagged with a second MID (different than the first MID). As another example, it may be desirable to tag different samples derived from the same source with different MIDs (e.g., samples derived over time or derived from different sites within a tissue). Further, MIDs can be generated in a variety of different ways, e.g., by a combinatorial tagging approach in which one MID is attached by ligation and a second MID is attached by primer extension. Thus, MIDs can be designed and implemented in a variety of different ways to track polynucleotide fragments during processing and analysis, and thus no limitation in this regard is intended.

"Nucleoside" as used herein includes the natural nucleosides, including 2'-deoxy and 2'-hydroxyl forms, e.g. as described in Kornberg and Baker, DNA Replication, 2nd Ed. (Freeman, San Francisco, 1992). "Analogs" in reference to nucleosides includes synthetic nucleosides having modified base moieties and/or modified sugar moieties, e.g. described by Scheit, Nucleotide Analogs (John Wiley, New York, 1980); Uhlman and

Peyman, Chemical Reviews, 90: 543-584 (1990), or the like, with the proviso that they are capable of specific hybridization. Such analogs include synthetic nucleosides designed to enhance binding properties, reduce complexity, increase specificity, and the like. Polynucleotides comprising analogs with enhanced hybridization or nuclease resistance properties are described in Uhlman and Peyman (cited above); Crooke et al, Exp. Opin. Ther. Patents, 6: 855-870 (1996); Mesmaeker et al, Current Opinion in Structual Biology, 5: 343-355 (1995); and the like. Exemplary types of polynucleotides that are capable of enhancing duplex stability include oligonucleotide N3'→P5' phosphoramidates (referred to herein as "amidates"), peptide nucleic acids (referred to herein as "PNAs"), oligo-2'-O-alkylribonucleotides, polynucleotides containing C-5 propynylpyrimidines, locked nucleic acids ("LNAs"), and like compounds. Such oligonucleotides are either available commercially or may be synthesized using methods described in the literature.

"Polymerase chain reaction," or "PCR," means a reaction for the in vitro amplification of specific DNA sequences by the simultaneous primer extension of complementary strands of DNA. In other words, PCR is a reaction for making multiple copies or replicates of a target nucleic acid flanked by primer binding sites, such reaction comprising one or more repetitions of the following steps: (i) denaturing the target nucleic acid, (ii) annealing primers to the primer binding sites, and (iii) extending the primers by a nucleic acid polymerase in the presence of nucleoside triphosphates. Usually, the reaction is cycled through different temperatures optimized for each step in a thermal cycler instrument. Particular temperatures, durations at each step, and rates of change between steps depend on many factors well-known to those of ordinary skill in the art, e.g. exemplified by the references: McPherson et al, editors, PCR: A Practical Approach and PCR2: A Practical Approach (IRL Press, Oxford, 1991 and 1995, respectively). For example, in a conventional PCR using Taq DNA polymerase, a double stranded target nucleic acid may be denatured at a temperature >90°C, primers annealed at a temperature in the range 50-75°C, and primers extended at a temperature in the range 72-78°C. The term "PCR" encompasses derivative forms of the reaction, including but not limited to, RT-PCR, real-time PCR, nested PCR, quantitative PCR, multiplexed PCR, and the like. Reaction volumes range from a few nanoliters, e.g. 2 nL, to a few hundred

μL, e.g. 200 μL. "Reverse transcription PCR," or "RT-PCR," means a PCR that is preceded by a reverse transcription reaction that converts a target RNA to a complementary single stranded DNA, which is then amplified, e.g. Tecott et al, U.S. patent 5,168,038, which patent is incorporated herein by reference. "Real-time PCR" means a PCR for which the amount of reaction product, i.e. amplicon, is monitored as the reaction proceeds. There are many forms of real-time PCR that differ mainly in the detection chemistries used for monitoring the reaction product, e.g. Gelfand et al, U.S. patent 5,210,015 ("TAQMAN™"); Wittwer et al, U.S. patents 6,174,670 and 6,569,627 (intercalating dyes); Tyagi et al, U.S. patent 5,925,517 (molecular beacons); which patents are incorporated herein by reference. Detection chemistries for real-time PCR are reviewed in Mackay et al, Nucleic Acids Research, 30: 1292-1305 (2002), which is also incorporated herein by reference. "Nested PCR" means a two-stage PCR wherein the amplicon of a first PCR becomes the sample for a second PCR using a new set of primers, at least one of which binds to an interior location of the first amplicon. As used herein, "initial primers" in reference to a nested amplification reaction mean the primers used to generate a first amplicon, and "secondary primers" mean the one or more primers used to generate a second, or nested, amplicon. "Multiplexed PCR" means a PCR wherein multiple target sequences (or a single target sequence and one or more reference sequences) are simultaneously carried out in the same reaction mixture, e.g. Bernard et al, Anal. Biochem., 273: 221-228 (1999)(two-color real-time PCR). Usually, distinct sets of primers are employed for each sequence being amplified.

"Polynucleotide" or "oligonucleotide" is used interchangeably and each means a linear polymer of nucleotide monomers. Monomers making up polynucleotides and oligonucleotides are capable of specifically binding to a natural polynucleotide by way of a regular pattern of monomer-to-monomer interactions, such as Watson-Crick type of base pairing, base stacking, Hoogsteen or reverse Hoogsteen types of base pairing, wobble base pairing, or the like. As described in detail below, by "wobble base" is meant a nucleic acid base that can base pair with a first nucleotide base in a complementary nucleic acid strand but that, when employed as a template strand for nucleic acid synthesis, leads to the incorporation of a second, different nucleotide base into the synthesizing strand. Such monomers and their internucleosidic linkages may be naturally

occurring or may be analogs thereof, e.g. naturally occurring or non-naturally occurring analogs. Non-naturally occurring analogs may include peptide nucleic acids (PNAs, e.g., as described in U.S. Patent 5,539,082, incorporated herein by reference), locked nucleic acids (LNAs, e.g., as described in U.S. Patent 6,670,461, incorporated herein by reference), phosphorothioate internucleosidic linkages, bases containing linking groups permitting the attachment of labels, such as fluorophores, or haptens, and the like. Whenever the use of an oligonucleotide or polynucleotide requires enzymatic processing, such as extension by a polymerase, ligation by a ligase, or the like, one of ordinary skill would understand that oligonucleotides or polynucleotides in those instances would not contain certain analogs of internucleosidic linkages, sugar moieties, or bases at any or some positions. Polynucleotides typically range in size from a few monomeric units, e.g. 5-40, when they are usually referred to as "oligonucleotides," to several thousand monomeric units. Whenever a polynucleotide or oligonucleotide is represented by a sequence of letters (upper or lower case), such as "ATGCCTG," it will be understood that the nucleotides are in 5'→3' order from left to right and that "A" denotes deoxyadenosine, "C" denotes deoxycytidine, "G" denotes deoxyguanosine, and "T" denotes thymidine, "I" denotes deoxyinosine, "U" denotes uridine, unless otherwise indicated or obvious from context. Unless otherwise noted the terminology and atom numbering conventions will follow those disclosed in Strachan and Read, Human Molecular Genetics 2 (Wiley-Liss, New York, 1999). Usually polynucleotides comprise the four natural nucleosides (e.g. deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine for DNA or their ribose counterparts for RNA) linked by phosphodiester linkages; however, they may also comprise non-natural nucleotide analogs, e.g. including modified bases, sugars, or internucleosidic linkages. It is clear to those skilled in the art that where an enzyme has specific oligonucleotide or polynucleotide substrate requirements for activity, e.g. single stranded DNA, RNA/DNA duplex, or the like, then selection of appropriate composition for the oligonucleotide or polynucleotide substrates is well within the knowledge of one of ordinary skill, especially with guidance from treatises, such as Sambrook et al, Molecular Cloning, Second Edition (Cold Spring Harbor Laboratory, New York, 1989), and like references.

"Primer" means an oligonucleotide, either natural or synthetic, that is capable, upon forming a duplex with a polynucleotide template, of acting as a point of initiation of nucleic acid synthesis and being extended from its 3' end along the template so that an extended duplex is formed. The sequence of nucleotides added during the extension process is determined by the sequence of the template polynucleotide. Usually primers are extended by a DNA polymerase. Primers are generally of a length compatible with its use in synthesis of primer extension products, and are usually are in the range of between 8 to 100 nucleotides in length, such as 10 to 75, 15 to 60, 15 to 40, 18 to 30, 20 to 40, 21 to 50, 22 to 45, 25 to 40, and so on, more typically in the range of between 18-40, 20-35, 21-30 nucleotides long, and any length between the stated ranges. Typical primers can be in the range of between 10-50 nucleotides long, such as 15-45, 18-40, 20-30, 21-25 and so on, and any length between the stated ranges. In some embodiments, the primers are usually not more than about 10, 12, 15, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, or 70 nucleotides in length.

Primers are usually single-stranded for maximum efficiency in amplification, but may alternatively be double-stranded. If double-stranded, the primer is usually first treated to separate its strands before being used to prepare extension products. This denaturation step is typically affected by heat, but may alternatively be carried out using alkali, followed by neutralization. Thus, a "primer" is complementary to a template, and complexes by hydrogen bonding or hybridization with the template to give a primer/template complex for initiation of synthesis by a polymerase, which is extended by the addition of covalently bonded bases linked at its 3' end complementary to the template in the process of DNA synthesis.

A "primer pair" as used herein refers to first and second primers having nucleic acid sequence suitable for nucleic acid-based amplification of a target nucleic acid. Such primer pairs generally include a first primer having a sequence that is the same or similar to that of a first portion of a target nucleic acid, and a second primer having a sequence that is complementary to a second portion of a target nucleic acid to provide for amplification of the target nucleic acid or a fragment thereof. Reference to "first" and "second" primers herein is arbitrary, unless specifically indicated otherwise. For example, the first primer can be designed as a "forward primer" (which initiates nucleic

acid synthesis from a 5' end of the target nucleic acid) or as a "reverse primer" (which initiates nucleic acid synthesis from a 5' end of the extension product produced from synthesis initiated from the forward primer). Likewise, the second primer can be designed as a forward primer or a reverse primer.

"Readout" means a parameter, or parameters, which are measured and/or detected that can be converted to a number or value. In some contexts, readout may refer to an actual numerical representation of such collected or recorded data. For example, a readout of fluorescent intensity signals from a microarray is the address and fluorescence intensity of a signal being generated at each hybridization site of the microarray; thus, such a readout may be registered or stored in various ways, for example, as an image of the microarray, as a table of numbers, or the like.

"Reflex site", "reflex sequence" and equivalents are used to indicate one or more sequences present in a polynucleotide that are employed to move a domain intra-molecularly from its initial location to a different location in the polynucleotide. The use of reflex sequences is described in detail in U.S. provisional applications 61/235,595 and 61/288,792, filed on August 20, 2009 and December 21, 2009, respectively, and entitled "Compositions and Methods for Intramolecular Nucleic Acid Rearrangement Using Reflex Sequences", both of which are incorporated herein by reference. In certain embodiments, a reflex sequence is chosen so as to be distinct from other sequences in the polynucleotide (i.e., with little sequence homology to other sequences likely to be present in the polynucleotide, e.g., genomic or sub-genomic sequences to be processed). As such, a reflex sequence should be selected so as to not hybridize to any sequence except its complement under the conditions employed in the reflex processes. The reflex sequence may be a synthetic or artificially generated sequence (e.g., added to a polynucleotide in an adapter domain) or a sequence present normally in a polynucleotide being assayed (e.g., a sequence present within a region of interest in a polynucleotide being assayed). In the reflex system, a complement to the reflex sequence is present (e.g., inserted in an adapter domain) on the same strand of the polynucleotide as the reflex sequence (e.g., the same strand of a double-stranded polynucleotide or on the same single stranded polynucleotide), where the complement is placed in a particular location so as to facilitate an intramolecular binding and polymerization event on such particular

strand. Reflex sequences employed in the reflex process described herein can thus have a wide range of lengths and sequences. Reflex sequences may range from 5 to 200 nucleotide bases in length.

"Solid support", "support", and "solid phase support" are used interchangeably and refer to a material or group of materials having a rigid or semi-rigid surface or surfaces. In many embodiments, at least one surface of the solid support will be substantially flat, although in some embodiments it may be desirable to physically separate synthesis regions for different compounds with, for example, wells, raised regions, pins, etched trenches, or the like. According to other embodiments, the solid support(s) will take the form of beads, resins, gels, microspheres, or other geometric configurations. Microarrays usually comprise at least one planar solid phase support, such as a glass microscope slide.

"Specific" or "specificity" in reference to the binding of one molecule to another molecule, such as a labeled target sequence for a probe, means the recognition, contact, and formation of a stable complex between the two molecules, together with substantially less recognition, contact, or complex formation of that molecule with other molecules. In one aspect, "specific" in reference to the binding of a first molecule to a second molecule means that to the extent the first molecule recognizes and forms a complex with another molecule in a reaction or sample, it forms the largest number of the complexes with the second molecule. Preferably, this largest number is at least fifty percent. Generally, molecules involved in a specific binding event have areas on their surfaces or in cavities giving rise to specific recognition between the molecules binding to each other. Examples of specific binding include antibody-antigen interactions, enzyme-substrate interactions, formation of duplexes or triplexes among polynucleotides and/or oligonucleotides, biotin-avidin or biotin-streptavidin interactions, receptor-ligand interactions, and the like. As used herein, "contact" in reference to specificity or specific binding means two molecules are close enough that weak noncovalent chemical interactions, such as Van der Waal forces, hydrogen bonding, base-stacking interactions, ionic and hydrophobic interactions, and the like, dominate the interaction of the molecules.

As used herein, the term "$T_m$" is used in reference to the "melting temperature." The melting temperature is the temperature (e.g., as measured in °C) at which a



IPR2025-01353; IPR2025-01355
CSHL EX2002

population of double-stranded nucleic acid molecules becomes half dissociated into single strands. Several equations for calculating the $T_m$ of nucleic acids are known in the art (see e.g., Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization (1985). Other references (e.g., Allawi, H.T. & SantaLucia, J., Jr., Biochemistry 36, 10581-94 (1997)) include alternative methods of computation which take structural and environmental, as well as sequence characteristics into account for the calculation of $T_m$.

"Sample" means a quantity of material from a biological, environmental, medical, or patient source in which detection, measurement, or labeling of target nucleic acids is sought. On the one hand it is meant to include a specimen or culture (e.g., microbiological cultures). On the other hand, it is meant to include both biological and environmental samples. A sample may include a specimen of synthetic origin. Biological samples may be animal, including human, fluid, solid (e.g., stool) or tissue, as well as liquid and solid food and feed products and ingredients such as dairy items, vegetables, meat and meat by-products, and waste. Biological samples may include materials taken from a patient including, but not limited to cultures, blood, saliva, cerebral spinal fluid, pleural fluid, milk, lymph, sputum, semen, needle aspirates, and the like. Biological samples may be obtained from all of the various families of domestic animals, as well as feral or wild animals, including, but not limited to, such animals as ungulates, bear, fish, rodents, etc. Environmental samples include environmental material such as surface matter, soil, water and industrial samples, as well as samples obtained from food and dairy processing instruments, apparatus, equipment, utensils, disposable and non-disposable items. These examples are not to be construed as limiting the sample types applicable to the present invention.

The terms "upstream" and "downstream" in describing nucleic acid molecule orientation and/or polymerization are used herein as understood by one of skill in the art. As such, "downstream" generally means proceeding in the 5' to 3' direction, i.e., the direction in which a nucleotide polymerase normally extends a sequence, and "upstream" generally means the converse. For example, a first primer that hybridizes "upstream" of a second primer on the same target nucleic acid molecule is located on the 5' side of the





IPR2025-01353; IPR2025-01355
CSHL EX2002

second primer (and thus nucleic acid polymerization from the first primer proceeds towards the second primer).

It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely", "only" and the like in connection with the recitation of claim elements, or the use of a "negative" limitation.

## DETAILED DESCRIPTION OF THE INVENTION

Before the present invention is described, it is to be understood that this invention is not limited to particular embodiments described, 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, since the scope of the present invention will be limited only by the appended claims.

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in that stated range is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included or excluded in the range, and each range where either, neither or both limits are included in the smaller ranges is also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, some potential and preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. It is

understood that the present disclosure supersedes any disclosure of an incorporated publication to the extent there is a contradiction.

It must be noted that as used herein and in 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 nucleic acid" includes a plurality of such nucleic acids and reference to "the compound" includes reference to one or more compounds and equivalents thereof known to those skilled in the art, and so forth.

The practice of the present invention may employ, unless otherwise indicated, conventional techniques and descriptions of organic chemistry, polymer technology, molecular biology (including recombinant techniques), cell biology, biochemistry, and immunology, which are within the skill of the art. Such conventional techniques include polymer array synthesis, hybridization, ligation, and detection of hybridization using a label. Specific illustrations of suitable techniques can be had by reference to the example herein below. However, other equivalent conventional procedures can, of course, also be used. Such conventional techniques and descriptions can be found in standard laboratory manuals such as *Genome Analysis: A Laboratory Manual Series* (Vols. I-IV), *Using Antibodies: A Laboratory Manual, Cells: A Laboratory Manual, PCR Primer: A Laboratory Manual, and Molecular Cloning: A Laboratory Manual* (all from Cold Spring Harbor Laboratory Press), Stryer, L. (1995) *Biochemistry* (4th Ed.) Freeman, New York, Gait, "*Oligonucleotide Synthesis: A Practical Approach*" 1984, IRL Press, London, Nelson and Cox (2000), Lehninger, A., *Principles of Biochemistry* 3rd Ed., W. H. Freeman Pub., New York, N.Y. and Berg et al. (2002) *Biochemistry,* 5th Ed., W. H. Freeman Pub., New York, N.Y., all of which are herein incorporated in their entirety by reference for all purposes.

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

As summarized above, aspects of the present invention are drawn to the use of degenerate nucleotide bases (e.g., in a degenerate base region, or DBR) added to polynucleotides undergoing sequence analysis that find use in establishing the number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. Including a DBR in polynucleotides undergoing sequencing analysis finds use in a variety of genetic analyses, including increasing the confidence of allele calling by providing a mechanism to determine a statistical value for an allele call, a value that cannot be derived from the read number alone. The DBR may be added to a polynucleotide in any convenient manner, including as part of an adapter (or pool of adapters) attached to the polynucleotides being sequenced, e.g., the DBR can be in an adapter that also includes a sequencing primer site.

Nucleic Acids

The present invention (as described in detail below) can be employed for the manipulation and analysis of nucleic acid sequences of interest (or polynucleotides) from virtually any nucleic acid source, including but not limited to genomic DNA, complementary DNA (cDNA), RNA (e.g., messenger RNA, ribosomal RNA, short interfering RNA, microRNA, etc.), plasmid DNA, mitochondrial DNA, synthetic DNA, etc. Furthermore, any organism, organic material or nucleic acid-containing substance can be used as a source of nucleic acids to be processed in accordance with the present invention including, but not limited to, plants, animals (e.g., reptiles, mammals, insects, worms, fish, etc.), tissue samples, bacteria, fungi (e.g., yeast), phage, viruses, cadaveric tissue, archaeological/ancient samples, etc. In certain embodiments, the nucleic acids in the nucleic acid sample are derived from a mammal, where in certain embodiments the mammal is a human.

In certain embodiments, the nucleic acid sequences are enriched. By enriched is meant that the nucleic acids (e.g., in a polynucleotide sample) are subjected to a process that reduces the complexity of the nucleic acids, generally by increasing the relative concentration of particular nucleic acid species in the sample (e.g., having a specific locus of interest, including a specific nucleic acid sequence, lacking a locus or sequence,



IPR2025-01353; IPR2025-01355
CSHL EX2002

being within a specific size range, etc.). There are a wide variety of ways to enrich nucleic acids having a specific characteristic(s) or sequence, and as such any convenient method to accomplish this may be employed. The enrichment (or complexity reduction) can take place at any of a number of steps in the process, and will be determined by the desires of the user. For example, enrichment can take place in individual parental samples (e.g., untagged nucleic acids prior to adaptor ligation) or in multiplexed samples (e.g., nucleic acids tagged with adapter sequences having MID; MID are described in further detail below).

In certain embodiments, nucleic acids in the nucleic acid sample are amplified prior to analysis. In certain of these embodiments, the amplification reaction also serves to enrich a starting nucleic acid sample for a sequence or locus of interest. For example, a starting nucleic acid sample can be subjected to a polymerase chain reaction (PCR) that amplifies one or more region of interest. In certain embodiments, the amplification reaction is an exponential amplification reaction, whereas in certain other embodiments, the amplification reaction is a linear amplification reaction. Any convenient method for performing amplification reactions on a starting nucleic acid sample can be used in practicing the subject invention. In certain embodiments, the nucleic acid polymerase employed in the amplification reaction is a polymerase that has proofreading capability (e.g., phi29 DNA Polymerase, Thermococcus litoralis DNA polymerase, Pyrococcus furiosus DNA polymerase, etc.).

In certain embodiments, the nucleic acid sample being analyzed is derived from a single source (e.g., a single organism, virus, tissue, cell, subject, etc.), whereas in other embodiments, the nucleic acid sample is a pool of nucleic acids extracted from a plurality of sources (e.g., a pool of nucleic acids from a plurality of organisms, tissues, cells, subjects, etc.), where by "plurality" is meant two or more. As such, in certain embodiments, a nucleic acid sample can contain nucleic acids from 2 or more sources, 3 or more sources, 5 or more sources, 10 or more sources, 50 or more sources, 100 or more sources, 500 or more sources, 1000 or more sources, 5000 or more sources, up to and including about 10,000 or more sources.

In certain embodiments, nucleic acid fragments that are to be pooled with nucleic acid fragments derived from a plurality of sources (e.g., a plurality of organisms, tissues,

cells, subjects, etc.), where by "plurality" is meant two or more. In such embodiments, the nucleic acids derived from each source include a multiplex identifier (MID) such that the source from which each tagged nucleic acid fragment was derived can be determined. In such embodiments, each nucleic acid sample source is correlated with a unique MID, where by unique MID is meant that each different MID employed can be differentiated from every other MID employed by virtue of at least one characteristic, e.g., the nucleic acid sequence of the MID. Any type of MID can be used, including but not limited to those described in co-pending U.S. Patent Application Serial Number 11/656,746, filed on January 22, 2007, and titled "Nucleic Acid Analysis Using Sequence Tokens", as well as U.S. Patent 7,393,665, issued on July 1, 2008, and titled "Methods and Compositions for Tagging and Identifying Polynucleotides", both of which are incorporated herein by reference in their entirety for their description of nucleic acid tags and their use in identifying polynucleotides. In certain embodiments, a set of MIDs employed to tag a plurality of samples need not have any particular common property (e.g., $T_m$, length, base composition, etc.), as the asymmetric tagging methods (and many tag readout methods, including but not limited to sequencing of the tag or measuring the length of the tag) can accommodate a wide variety of unique MID sets.

Degenerate Base Region (DBR)

Aspects of the present invention include methods and compositions for establishing the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. In these aspects of the invention, a degenerate base region (DBR) is attached to the starting polynucleotide molecules that are subsequently sequenced (e.g., after certain process steps are performed, e.g., amplification and/or enrichment, e.g., PCR). As detailed below, evaluating the number of different DBR sequences present in a sequencing run allows the establishment of the minimum number of different starting polynucleotides that have been sequenced for a particular polynucleotide (or region of interest; ROI). This number can be used to give a statistical measure of allele calls, thus increasing the confidence in making such allele determinations (e.g., when calling homozygous alleles).

DNA sequencing typically includes a step of attaching an adapter to the polynucleotides in a sample to be sequenced, where the adaptor contains a sequencing primer site (e.g., by ligation). As used herein, a "sequencing primer site" is a region of a polynucleotide that is either identical to or complementary to the sequence of a sequencing primer (when in a single stranded form) or a double stranded region formed between a sequencing primer sequence and its complement. The specific orientation of a sequencing primer site can be inferred by those of ordinary skill in the art from the structural features of the specific polynucleotide containing the sequencing primer site.

In addition to the sequencing primer site, a degenerate base region (DBR) is also attached to the polynucleotides, either as part of the adapter containing the sequencing primer site or independently (e.g., in a second adapter attached to the polynucleotide). Any convenient method for attaching or adding a DBR to the polynucleotides may be employed. A DBR is a region that can have a variable base composition or sequence (which may be considered a "random") as compared to other tagged polynucleotides in the sample. The number of different DBRs in a population of polynucleotides in a sample will be dependent on the number of bases in the DBR as well as the potential number of different bases that can be present at each position. Thus, a population of polynucleotides having attached DBRs with two base positions, where each position can be any one of A, C, G and T, will have potentially 16 different DBRs (AA, AC, AG, etc.). DBR may thus include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more bases, including 15 or more, 20 or more, etc. In certain embodiments the DBR is from 3 to 10 bases in length. Moreover, each position in a DBR may have a different base composition. For example, a 4 base DBR may have any of the following compositions: NNNN; NRSN; SWSW; BDHV (see Table 1 below for IUPAC nucleotide code). As such, no limitation with regard to the length or base composition of a DBR is intended.

| IUPAC nucleotide code | Base |
|---|---|
| A | Adenine |
| C | Cytosine |
| G | Guanine |
| T (or U) | Thymine (or Uracil) |





IPR2025-01353; IPR2025-01355
CSHL EX2002

| | |
|---|---|
| R | A or G |
| Y | C or T |
| S | G or C |
| W | A or T |
| K | G or T |
| M | A or C |
| B | C or G or T |
| D | A or G or T |
| H | A or C or T |
| V | A or C or G |
| N | any base |

It is noted here that a DBR may be a single region (i.e., having all nucleotide bases adjacent to one another) or may be present in different location on a polynucleotide (i.e., the bases of the DBR are separated by non-DBR sequences, also called a split DBR). For example, a DBR may have one or more bases in a first adapter at a first location on a polynucleotide and one or more bases in a second adapter at a second location on the same polynucleotide (e.g., the DBR may have bases present at both ends of an asymmetrically tagged polynucleotide, i.e., a polynucleotide having asymmetric adapters). No limitation in this regard is intended.

In certain embodiments in which the DBR is present within an adapter population that includes other functional domains (e.g., sequencing primer site, MID, reflex sequence), the functional domains in the adapter population will be identical to each other whereas the DBR will vary. In other words, unlike the other domains in an adapter population, the DBR has a variable (or random) base composition. By "adapter population", "population of adapters", and the like, is meant a sample of adapter molecules that are designed to be attached to the polynucleotides in a sample.

Generating adapters having a DBR may be achieved in any convenient manner, e.g., using DNA synthesis methods well known in the art (see citations in definitions section above).

Once attached to the polynucleotides in the parent sample, the polynucleotides may be subjected to further processing and ultimately sequenced. Processing steps that

may be performed include any process steps that are desired by a user, e.g., enrichment, amplification, and the like. In the sequencing step, the sequence of the DBR as well as a portion of the polynucleotide (e.g., containing a region of interest) is obtained. Once the sequences are obtained, the number of different DBRs attached to a polynucleotide of interest is determined. This number represents the minimum number of different polynucleotides of interest from the starting parent sample that are represented in the sequencing results.

Consider, for example, a two base DBR having base composition NN (where N is any deoxynucleotide base, i.e., A, G, C or T) employed in sequencing a locus to make an allele call for a particular subject sample (i.e., whether a subject is homozygous or heterozygous at the locus). Although there may be some biases from oligonucleotide synthesis, one can expect there to be 16 different DBRs in the population of adaptors with roughly equal probability (as described above). When a potential homozygous allele call is identified, determining the number of DBRs present in the sequencing run can be used to find out the minimum number of polynucleotide molecules that were actually sequenced (and thus the minimum number that were amplified during the processing steps).

For diploid genomes, allele calling (in the ideal or theoretical case) can be modeled by the binomial distribution. Given that the two alleles copies (X and Y) differ at some site, the probability of observing all Xs or all Ys is given by the formula $( ½ )^c$, where c is the number of observations (reads) of the site. If we observe X ten times at a site (and no Y), we can say that the sample is likely to be homozygous for type X. The probability of an error in this call is therefore $( ½ )^{10}$ (just under one in a thousand).

Our experiments show that low amounts of DNA in the early stage of sample preparation can result in a high coverage of reads all corresponding to one allele, and that this can occur many more times than should be expected according to the binomial distribution. This is due to amplification of a few molecules of DNA (or even a single molecule) that results in a large number of reads derived from a genetic locus on a single chromosome (i.e., only one of the two diploid chromosomes actually present in the sample of interest). The result of this is that the error as a function of coverage deviates wildly from the predicted binomial error.

Using a DBR as described herein will increase the confidence in making allele calls from samples having limited amounts of DNA. For example, if 16 sequencing reads of one allele from a genetic locus all contain the DBR sequence GA, then it is likely that all these reads are from the same parent polynucleotide molecule (and thus a homozygous allele call is not justified). However, if the 16 sequencing reads each have a different DBR sequence, a homozygous call can be made with more confidence, as each read came from a different parent polynucleotide molecule.

It is noted here that in many embodiments, it is not possible to conclude that polynucleotides having identical DBR sequences are derived from the same parent polynucleotide molecule, as multiple identical DBRs may be present in the DBR-attached polynucleotides. For example, if an adapter population that contains a DBR of two N bases is used to tag a sample containing more that 16 polynucleotides, a subset of the tagged polynucleotides will have identical DBRs, and thus it will not be possible to determine that their sequences were derived from different parent polynucleotide molecules.

One exemplary way to determine more accurately the actual number of starting or parent molecules would be to increase the degeneracy of DBRs (i.e., to increase the number of unique sequences in the DBR used to label the particular sample of interest) so that every single molecule is likely to have a different DBR. In any event we can, in exemplary methods, either use the number of observed DBRs or else the probability distribution of the numbers of reads likely to produce that numbers of observed DBRs.

When calculating estimates of whether a particular allele call is a heterozygote or homozygote, one can create/employ an appropriate function $L(r,v)$ that returns the likelihood of a genotype given $r$ reference and $v$ variant reads. When employing DBRs as described herein, a modified function for calculating estimates $L(r', v')$ may be used, where $r'$ is the number of unique DBRs for the reference read and $v'$ is the number of unique DBRs for the variant read. Any convenient function for making allele calls may be employed and modified to employ data regarding DBR reads as described herein.

It is noted here that aspects of the invention can be used to increase the confidence in calling copy number variations in a polynucleotide sample, e.g., a genomic sample. Copy number variations can include genomic rearrangements such as deletions,

duplications, inversions, and translocations or whole chromosome aneuploidies such as monosomies, disomies, trisomies, tetrasomies, and pentasomies. Consider, for example, a duplication event where parents have genotypes AC and CC at a given SNP, and the proband has genotype ACC. In the parent with genotype AC, given sufficient depth of sequencing coverage (i.e. sufficient sequencing reads), the number of DBRs associated with the C-allele and the A-allele are expected to be similar. In the proband, given sufficient depth of sequencing coverage, the number of DBRs associated with the C-allele is expected to be 2-times the number of DBRs associated with the A-allele, which provides evidence for duplication event encompassing the C-allele. The use of DBRs, rather than the number of sequencing reads, provides more confidence in calling a copy number variation since each DBR came from a different parent polynucleotide molecule.

Kits and Systems

Also provided by the subject invention are kits and systems for practicing the subject methods, i.e., for using DBRs to determine the minimum number of different starting polynucleotides that have been sequenced for a particular polynucleotide. As such, systems and kits may include polynucleotides containing DBRs (e.g., adapters) as well as any other functional domains of interest as described herein (e.g., sequencing primer sites, MIDs, reflex sequences, etc.). Systems and kits may also include and regents for performing any steps in attaching adapters to polynucleotides in a parents sample, preparing a parent sample for adapter/DBR attachment, and/or reagents for performing sequencing reactions (e.g., ligases, restriction enzymes, nucleotides, polymerases, primers, sequencing primers, dNTPs, ddNTPs, exonucleases, etc.). The various components of the systems and kits may be present in separate containers or certain compatible components may be precombined into a single container, as desired.

The subject systems and kits may also include one or more other reagents for preparing or processing a nucleic acid sample according to the subject methods. The reagents may include one or more matrices, solvents, sample preparation reagents, buffers, desalting reagents, enzymatic reagents, denaturing reagents, where calibration standards such as positive and negative controls may be provided as well. As such, the kits may include one or more containers such as vials or bottles, with each container

containing a separate component for carrying out a sample processing or preparing step according to the present invention.

In addition to above-mentioned components, the subject kits typically further include instructions for using the components of the kit to practice the subject methods, e.g., to employ DBR as described above. The instructions for practicing the subject methods are generally recorded on a suitable recording medium. For example, the instructions may be printed on a substrate, such as paper or plastic, etc. As such, the instructions may be present in the kits as a package insert, in the labeling of the container of the kit or components thereof (i.e., associated with the packaging or sub-packaging) etc. In other embodiments, the instructions are present as an electronic storage data file present on a suitable computer readable storage medium, e.g. CD-ROM, diskette, etc. In yet other embodiments, the actual instructions are not present in the kit, but means for obtaining the instructions from a remote source, e.g. via the internet, are provided. An example of this embodiment is a kit that includes a web address where the instructions can be viewed and/or from which the instructions can be downloaded. As with the instructions, this means for obtaining the instructions is recorded on a suitable substrate.

In addition to the subject database, programming and instructions, the kits may also include one or more control samples and reagents, e.g., two or more control samples for use in testing the kit.

## EXAMPLES

*Methods*

Two identical samples of mouse genomic DNA were labeled with adapters. One sample used an adapter that has a redundantly synthesized region consisting of 7 bases (RYBDHVB), each of which could be one of two (for R and Y positions) or three (for B, D, H or V positions) bases (or a total of 972 different sequences) followed by the bases AC**A**; the second sample used an adapter that has a redundantly synthesized region consisting of 7 bases (RYBDHVB), each of which could be one of two (for R and Y positions) or three (for B, D, H or V positions) bases (or a total of 972 different sequences) followed by the bases AC**G**. Note the bolded underlined bases correspond to



IPR2025-01353; IPR2025-01355
CSHL EX2002

a synthetic polymorphic site. In these adapters, the sequence RYB serves as the DBR region and the DHVB serves as the MID. Thus, there were 12 possible DBR (2 x 2 x 3) codes and 81 different MIDs present (3 x 3 x 3 x 3).

The two samples were then mixed together in equal amounts to create, in effect, a perfect 50/50 heterozygote of **A** and **G** three bases downstream from the MID (i.e., the DHVB sequence). Different amounts of the mixture (100ng, 300ng, 600ng, 2500ng, 5000ng, and 10,000ng) were subjected to hybridization pull-down reactions followed by amplification through 10 cycles of PCR with TiA and TiB primers. The capture probes employed were 5'-biotinylated 60-mer reverse phase cartridge purified oligonucleotides (BioSearch). After amplification with TiA and TiB, a secondary PCR reaction with TiA and a sequence-specific primer tailed with TiB (5'-CCTATCCCCTGTGTGCCTTGGCAGTCTCAGGGACACCCAGCCAAGACAGC-3') was used to amplify a specific fragment. The PCR fragment generated from each sample in the hybridization pull-down/PCR was sent for 454 Ti Shot sequencing to determine the DBR, MID and A/G allele.

The amplicon sequence from the hybridization pull down/PCR is shown below. The DBR region is underlined, the MID is in bold, and the allele (R, which corresponds to A or G) is bold underline.

CCATCTCATCCCTGCGTGTCTCCGACTCAGRYB**DHVBAC**R**T**AGAATGTGCAT
GGATCGTATGAGCACCTGTGGGCAGGGCAAGTGGCAGATGCCTTAGTGGATC
TCACTGGAAGCCTGGCAGAAAGGTGGAGCTTGAAGGATGTAACGAAAGCCA
GCGGCCAGCAGGACAGACCCAGTGGTGGGGAGCACAGAACTTGTCGGCAGC
TACTCCACCTGAAGGACCGGTGTCTAATCAGCTGCTCTGTGCTTAGCCCCAGA
GCAGGTACAGCTATGGCTACAACTCCCTCCACCATTAGCTTGTTACAGAGAA
GGAAATCGGTCCTTGAGAGGCTGTCTTGGCTGGGTGTCCCTGAGACTGCCAA
GGCACACAGGGGATAGG

*Results*

Figure 1 shows the allele ratio for each MID in the sample. The numbers at the top of each of the 6 panels show the input mass of genomic DNA used (in nanograms).



IPR2025-01353; IPR2025-01355
CSHL EX2002

The abscissa shows, for any particular MID, the fraction of A sequencing reads, or calls, (i.e., the number of reads of one of the polymorphic bases at the synthetic SNP position) to the total number of calls, e.g., the ratio of (A calls)/(A calls + G calls). This is called the allele ratio. The ordinate is the number of MIDs observed with a particular allele ratio (the total number of MIDs, as noted above, is 81). Because the input DNA was known to be at a 50/50 A/G ratio, the allele ratio for each sample should be 0.5.

At low input mass, the allele ratio is distorted – either too high or two low – because the molecules on the input to the first PCR step were limiting and so one of the two alleles was preferentially observed. As mass into the first step increases, both alleles are observed and at closer and closer to the expected 0.5 ratio. This analysis shows that there was considerable allele drop out at 100ngs, 300ngs and 600ngs, while little to no allele drop out occurred at the higher input levels.

Figure 2 shows the fraction of DBR sequences for each MID associated with each allele. Since the input material is nominally 50/50 A and G at the synthetic polymorphic position, we would expect each of the 81 effective MIDs to be associated with 50% A reads and 50% G reads (as noted above). Furthermore, since the 12 DBRs are random and are associated with each of the 81 different MIDs, one would expect, with sufficient input DNA copies, that all 12 DBRs would be observed for the A allele and all 12 for the G allele. Thus, the fraction of DBRs observed for a particular base in the ideal case would be 12/24 or 0.5 for each allele. To the extent that insufficient numbers of molecules go into the processing steps, there may be less than 12 DBRs for each MID allele, and so the ratio can stray away from the ideal.

In Figure 2, the abscissa shows, for any particular MID, the proportion of the DBRs actually observed for each allele divided by the total number of DBRs actually observed. The ordinate is the number of MIDs observed with a particular proportion or DBRs associated with an allele. There are, and should be, a total of 81 different MIDs seen. We observe that at low input mass, one frequently can see a distorted proportion – either too high or two low. This is likely due to limiting number of molecules input into the first PCR step, and so one of the two alleles was preferentially observed. As mass into the first step increases, both alleles are observed more frequently and associated with



IPR2025-01353; IPR2025-01355
CSHL EX2002

each allele is an observation of more and more of the DBRs, and thus a proportion closer to the 0.5.

A comparison of the date in Figures 1 and 2 shows an additional feature of employing DBRs as described herein. For a true heterozygote, one starts to see sooner (i.e., at lower mass) in the DBR analysis (Figure 2) that the distribution clusters around the expected 0.5 ratio. This makes sense because, for example, observation of 6 of the 12 DRBs for the A allele and 4 of the 12 DBRs of the G allele would result in a proportion of (6 / [6+4] ) = 0.60 which, in fact is fairly close to the expected 0.5. The net effect is that the use of DBRs gives much greater confidence in the presence of a true heterozygote or true homozygote.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

Accordingly, the preceding merely illustrates the principles of the invention. It will be appreciated that those skilled in the art will be able to devise various arrangements which, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope. Furthermore, all examples and conditional language recited herein are principally intended to aid the reader in understanding the principles of the invention and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents and equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure. The scope of the present invention, therefore, is not intended to be limited to the exemplary embodiments shown and

described herein. Rather, the scope and spirit of present invention is embodied by the appended claims.

## CLAIMS

That which is claimed is:

1. A method of determining the minimum number of polynucleotide molecules sequenced from a sample, the method comprising:

attaching an adapter to polynucleotide molecules in a sample, wherein the adapter comprises a degenerate base region (DBR) comprising at least one nucleotide base selected from: R, Y, S, W, K, M, B, D, H, V, and N;

sequencing a plurality of the adapter-attached polynucleotides, wherein the sequence of the DBR and at least a portion of the polynucleotide is obtained for each of the plurality of adapter-attached polynucleotides; and

determining the number of distinct DBR sequences present in the plurality of sequenced adapter-attached polynucleotides to determine the minimum number of polynucleotides from the sample that were sequenced in the sequencing step.

2. The method of claim 1, wherein the adapter further comprises a sequencing primer site.

3. The method of claim 1, wherein the adapter further comprises one or more of: an MID, a reflex sequence, and a promoter site.

4. The method of claim 1, wherein the attaching step further comprises:

attaching adapters to multiple different samples, wherein the adapter employed for each different sample comprises a unique MID; and

pooling the multiple different adapter attached samples prior to the sequencing step.



IPR2025-01353; IPR2025-01355
CSHL EX2002

5. The method of claim 1, wherein the DBR comprises at least 2 nucleotide bases, wherein each of the at least 2 nucleotide bases are selected from: R, Y, S, W, K, M, B, D, H, V, and N.

6. The method of claim 5, wherein the DBR comprises from 3 to 20 nucleotide bases, wherein each of the 3 to 10 nucleotide bases is selected from: R, Y, S, W, K, M, B, D, H, V, and N.

7. The method of claim 1, wherein the determined minimum number of polynucleotides from the sample that were sequenced in the sequencing step is employed in a method for allele calling.

8. The method of claim 1, wherein the method further comprises amplifying the adapter-attached polynucleotides prior to the sequencing step.

9. The method of claim 1, wherein the sequencing step comprises performing a next generation sequencing process.

10. The method of claim 1, wherein the sample is genomic DNA sample.

11. The method of claim 4, wherein each of the multiple different samples is a genomic DNA sample derived from a different subject.

12. The method of claim 11, wherein the subject is a human.

13. The method of claim 1, further wherein the sample is enriched for a region of interest prior to the attaching step.

14. The method of claim 1, further comprising enriching the adapter-attached polynucleotides for a region of interest.



IPR2025-01353; IPR2025-01355
CSHL EX2002

15. The method of claim 5, wherein the adapter is an asymmetric adapter, wherein a first adapter domain is present at a first end of the polynucleotide and a second adapter domain is present at a second end of the polynucleotide, wherein the DBR is a split DBR comprising one or more of the at least 2 nucleotide bases in the first adapter domain and one or more of the at least 2 nucleotide bases in the second adapter domain.

# INCREASING CONFIDENCE OF ALLELE CALLS WITH MOLECULAR COUNTING

## ABSTRACT OF THE DISCLOSURE

Aspects of the present invention include methods and compositions for establishing the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. In these aspects of the invention, a degenerate base region (DBR) is attached to the starting polynucleotide molecules that are subsequently sequenced (e.g., after certain process steps are performed, e.g., amplification and/or enrichment). The number of different DBR sequences present in a sequencing run establishes a minimum number of different starting polynucleotides that have been sequenced. DBRs can be used to enhance numerous different nucleic acid sequence analysis applications, including allowing higher confidence allele call determinations in genotyping applications.

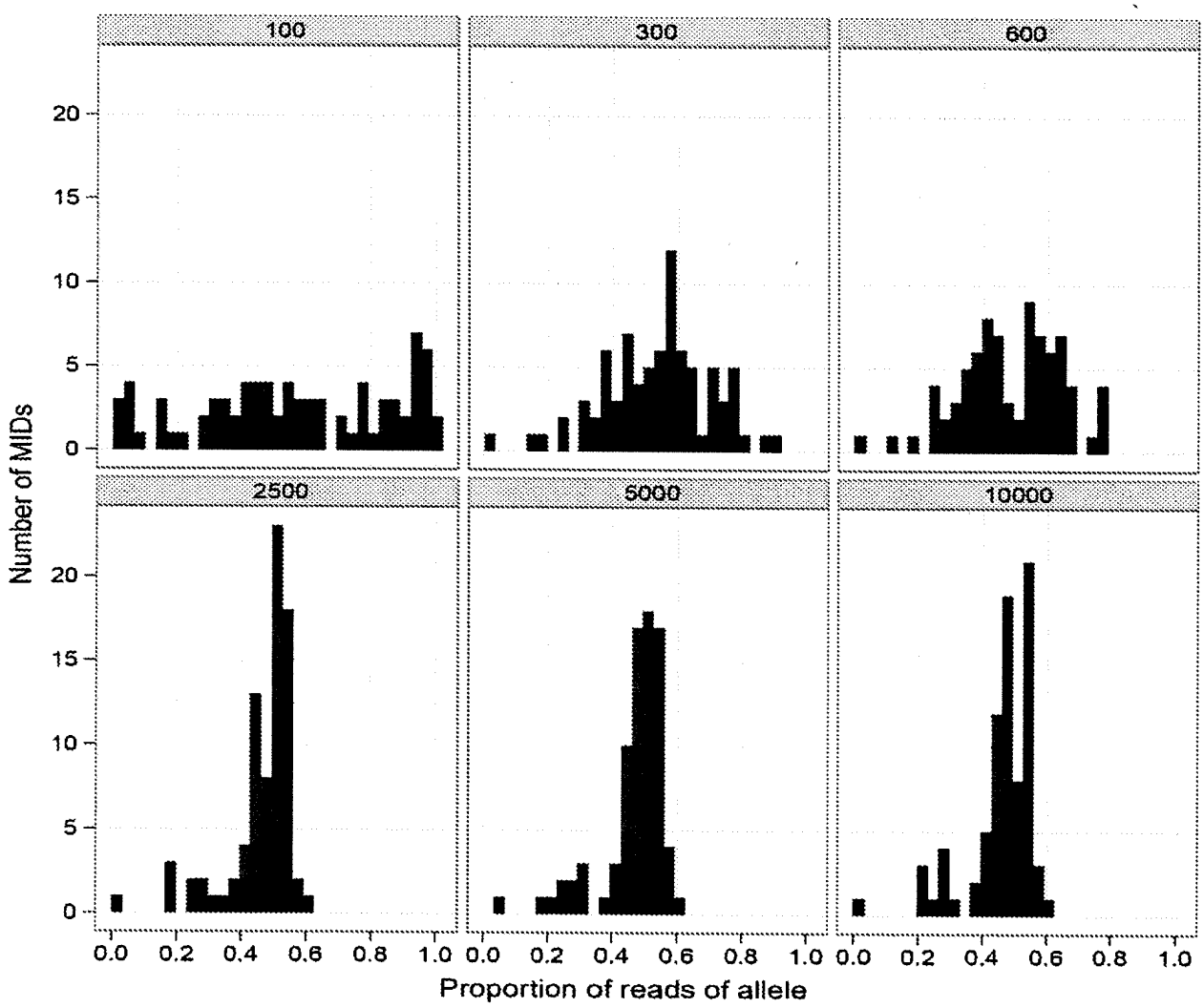

100
300
600
2500
5000
10000
Number of MIDs
0
5
10
15
20
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of reads of allele

Fig. 1

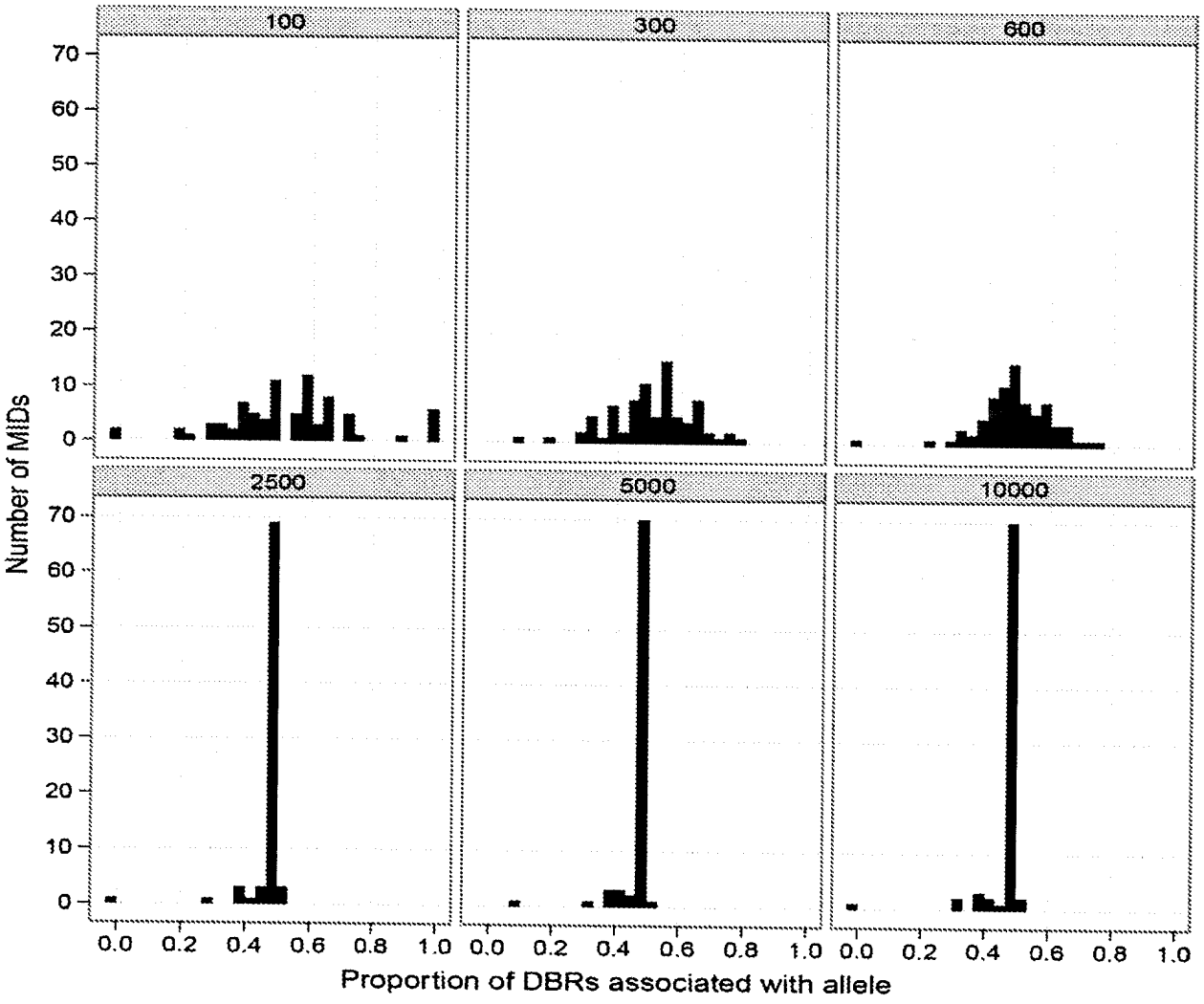



Fig. 2



IPR2025-01353; IPR2025-01355
CSHL EX2002

Blue: Amendments submitted in a June 4, 2015 Response to February 4, 2015 Final Office Action (not entered).

Red: Proposed additional amendments.

**Listing of Claims**

1. (Currently Amended) A method for obtaining from genomic ~~material~~ nucleic acids genomic copy number information ~~unaffected by amplification distortion~~, comprising:

a) obtaining segments of the genomic ~~material~~ nucleic acids;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic ~~material~~ nucleic acids from step (a) and a substantially unique tag having a unique nucleotide sequence, so as to permit differentiation from another tag having a unique nucleotide sequence, and which tag is part of another tagged nucleic acid molecule;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic ~~material~~ nucleic acids by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (d) corresponding to one segment of the segments of the


IPR2025-01353; IPR2025-01355
CSHL EX2002

genomic ~~material~~ nucleic acids to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information ~~unaffected by amplification distortion~~.

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Previously Presented) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is


IPR2025-01353; IPR2025-01355
CSHL EX2002

comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising i) summing the counts obtained in step (f) for locations which comprise a region of the genome~~, and comparing the sum to a sum~~; ii) summing the counts obtained from a reference sample for the same region of the genome[[,]]; and iii) comparing the sums obtained in step (i) and (ii), thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Currently Amended) A method for obtaining from mRNA transcripts mRNA copy number information ~~unaffected by amplification distortion~~, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag

associated sequence read of said tag associated sequence reads of step (c) corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information ~~unaffected by amplification distortion~~.

10. (Cancelled)

11. (Currently Amended) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic ~~material~~ nucleic acids to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Currently Amended) The method of claim 1 step (b), wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic ~~material~~ nucleic acids to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein adding the polynucleotide tail comprises the use of a terminal transferase.


IPR2025-01353; IPR2025-01355
CSHL EX2002

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to ~~sequencing in~~ step (d).

25. (Cancelled)

26. (Currently Amended) The method of claim 1, wherein the substantially unique tags ~~sequences~~ further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to said sequencing of the product of step (c).

28. (Previously Presented) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture

has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.


IPR2025-01353; IPR2025-01355
CSHL EX2002

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicant | : | Cold Spring Harbor Laboratory | |
| Inventors | : | James Hicks, Nicholas Navin, Jennifer Troge, Zihua Wang, and Michael Wigler | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

**VIA EFS-WEB**
Commissioner for Patents
Alexandria, VA 22313-1450

Sir:

**DECLARATION UNDER 37 C.F.R. § 1.131**

Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, New York 11724, U.S.A., is the owner of (assignee of record of the entire right, title and interest in) the above-identified application by virtue of an assignment of U.S. Serial No. 13/278,333, filed October 21, 2011, which assignment was recorded with the U.S. Patent and Trademark Office on January 16, 2012 at Reel No. 027537, Frame 0950-09. A copy of the recorded assignment is attached hereto as **Exhibit 1.**

Cold Spring Harbor Laboratory hereby declares that:

1. The inventors James Hicks, Nicholas Navin, Jennifer Troge, Zihua Wang, and Michael Wigler, jointly conceived of the invention claimed in claims 1-7, 9, 11-13, 24, 26-28 and 68-71 as proposed in the concurrently filed Amendment in the subject application, the text of which claims is attached hereto as **Exhibit 2.**

2. During a July 28, 2015 Interview, the Examiner has identified U.S. Patent No. 8,741,606 as a potential reference against the claims pending in the subject application. U.S. Patent No. 8,741,606 is a continuation of U.S. Patent No. 8,685,678, which issued from U.S. Serial No. 13/852,887, filed on March 28, 2013, which is a continuation of U.S. Serial No. 13/237,124, filed on September 20, 2011, which claimed benefit of U.S. Provisional Patent Application Nos. 61/432,119, filed January 12, 2011, and 61/385,001, filed

Applicant: Cold Spring Harbor Laboratory
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Exhibit B


IPR2025-01353; IPR2025-01355
CSHL EX2002

September 21, 2010.

3. Applicant claims benefit of U.S. Provisional Patent Application No. 61/406,067, filed October 22, 2010. Applicant's proposed amended independent claims 1 and 9 (from which all other pending claims depend) are supported by U.S. Provisional Patent Application No. 61/406,067, filed October 22, 2010. Accordingly, only the content that was present in U.S. Provisional Patent Application No. 61/385,001, filed September 21, 2010 (hereinafter "Casbon et al. Provisional"), could potentially be cited against the Applicant's claims. A copy of Casbon et al. Provisional is attached hereto as **Exhibit 3**.

4. While Applicant maintains that its claims are patentable over the disclosure of Casbon et al. Provisional, to advance prosecution this Declaration is being submitted to antedate the Casbon et al. Provisional.

5. The inventors conceived embodiments of the claimed invention before September 21, 2010, i.e., the filing date of the Casbon et al. Provisional. As evidence, Applicant attaches as **Exhibit 4** hereto, a scheme representing an embodiment of the method recited in the proposed claims. This scheme was sent to outside patent counsel and subsequently included in a draft provisional application which was sent to Cold Spring Harbor Laboratory before September 21, 2010. This scheme shows the conception of an embodiment of the claimed methodology in the United States before September 21, 2010. This scheme is the same as Figure 8 of U.S. Provisional Application No. 61/406,067, filed October 22, 2010, and also the same as Figure 8 of the subject application.

6. Also prior to September 21, 2010, the inventors described a method for obtaining genomic copy number information using a methodology represented by the scheme attached hereto as Exhibit 4. As evidence, Applicant attaches as **Exhibit 5** hereto, a portion of a draft provisional application circulated between Cold Spring


IPR2025-01353; IPR2025-01355
CSHL EX2002

Harbor Laboratory and its outside patent counsel before September 21, 2010. This text provides further evidence that the inventors conceived of an embodiment of the claimed method in the United States before September 21, 2010. This text in substantially similar form appears on page 48, lines 14-33 of the subject application as filed.

7. The evidence discussed above shows that the inventors conceived of an embodiment of the claimed invention prior to September 21, 2010. Furthermore, the embodiment was included in Applicant's subsequently filed U.S. Provisional Application No. 61/406,067, filed October 22, 2010.

I have reviewed the assignments attached hereto and certify that, to the best of my knowledge and belief, Cold Spring Harbor Laboratory is the assignee of all right, title and interest in and to the subject application. I further certify that I am authorized to sign this Declaration on behalf of Cold Spring Harbor Laboratory.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made herein on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment of not more than five (5) years, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: 8/4/2015 Cold Spring Harbor Laboratory

By: _____
Vladimir Drozdoff, Ph.D., J.D.
Patent and Business Development Counsel
Office of Technology Transfer
Cold Spring Harbor Laboratory

5981/81503-A

GJG/REB

**UNITED STATES PATENT AND TRADEMARK OFFICE**

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

JANUARY 18, 2012

PTAS

COOPER & DUNHAM LLP
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

**501787086**

UNITED STATES PATENT AND TRADEMARK OFFICE
NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT RECORDATION BRANCH OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE ASSIGNMENT RECORDATION BRANCH AT 571-272-3350. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, MAIL STOP: ASSIGNMENT RECORDATION BRANCH, P.O. BOX 1450, ALEXANDRIA, VA 22313.

RECORDATION DATE: 01/16/2012

REEL/FRAME: 027537/0950
NUMBER OF PAGES: 6

BRIEF: ASSIGNMENT OF ASSIGNORS INTEREST (SEE DOCUMENT FOR DETAILS).

DOCKET NUMBER: 5981/81503-A/GJG/GC

ASSIGNOR:
HICKS, JAMES — DOC DATE: 11/16/2011

ASSIGNOR:
NAVIN, NICHOLAS — DOC DATE: 11/14/2011

ASSIGNOR:
TROGE, JENNIFER — DOC DATE: 11/16/2011

ASSIGNOR:
WANG, ZIHUA — DOC DATE: 11/16/2011

ASSIGNOR:
WIGLER, MICHAEL — DOC DATE: 11/16/2011

ASSIGNEE:
COLD SPRING HARBOR LABORATORY
ONE BUNGTOWN ROAD
COLD SPRING HARBOR, NEW YORK 11724

Applicant: Cold Spring Harbor Labora
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Declaration Exhibit 1

P.O. Box 1450, Alexandria, Virginia 22313-1450 - WWW.USPTO.GOV

APPLICATION NUMBER: 13278333 FILING DATE: 10/21/2011
PATENT NUMBER: ISSUE DATE:
TITLE: VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

ASSIGNMENT RECORDATION BRANCH
PUBLIC RECORDS DIVISION


IPR2025-01353; IPR2025-01355
CSHL EX2002

**Listing of Claims**

1. (Currently Amended) A method for obtaining from genomic ~~material~~ nucleic acids genomic copy number information ~~unaffected by amplification distortion~~, comprising:

a) obtaining segments of the genomic ~~material~~ nucleic acids;

b) tagging the segments with substantially unique tags to generate tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic ~~material~~ nucleic acids from step (a) and a substantially unique tag having a unique nucleotide sequence, so as to permit differentiation from another tag having a unique nucleotide sequence, and which tag is part of another tagged nucleic acid molecule;

c) subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction (PCR);

d) generating tag associated sequence reads by sequencing the product of step (c);

e) assigning each of the tagged nucleic acid molecules to a location on a genome associated with the genomic ~~material~~ nucleic acids by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (d) corresponding to one segment of the segments of the genomic ~~material~~ nucleic acids to a location on the genome; and

f) counting the number of tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome,

thereby obtaining genomic copy number information ~~unaffected by amplification distortion~~.

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising more than one location on the genome, by counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the highest count obtained in step (f) for the locations within the region.

3. (Previously Presented) The method of claim 1, further comprising comparing a count obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

a) summing the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

b) summing the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region is comprised of a number of locations which is comparable to the number of locations of the first region;

c) comparing the value obtained in step (a) to the value obtained in step (b),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (b) further comprises

i) summing the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region is comprised of a number of locations which is comparable to the number of locations of the first region; and

ii) obtaining an average of the sum of the counts obtained in step (f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region.

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising i) summing the counts obtained in step (f) for locations which comprise a region of the genome~~, and comparing the sum to a sum~~; ii) summing the counts obtained from a reference sample for the same region of the genome[[,]]; and iii) comparing the sums obtained in step (i) and (ii), thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Currently Amended) A method for obtaining from mRNA transcripts mRNA copy number information ~~unaffected by amplification distortion~~, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to the first order derivative strands; and

iii) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique primers, thereby generating tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to polymerase chain reaction (PCR) amplification;

c) generating tag associated sequence reads by sequencing the product of step (b);

d) assigning each tagged nucleic acid molecule to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read <u>of said tag associated sequence reads of step (c)</u> corresponding to a mRNA transcript to a location on the cDNA library; and

e) counting the number of tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining mRNA copy number information ~~unaffected by amplification distortion~~.

10. (Cancelled)

11. (Currently Amended) The method of claim 1, wherein tagging the segments to generate tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of the genomic ~~material~~ nucleic acids to generate zero-th order derivative strands;

b) subjecting the zero-th order derivative strands of step (a) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement, thereby generating first order derivative strands;

c) adding a polynucleotide tail to the first order derivative strands;

d) subjecting the first order derivative strands to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement, thereby generating second order derivative strands,

wherein the primers of at least one of steps (b) and (d) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

12. (Currently Amended) The method of claim 1 step (b), wherein tagging the segments to generate the tagged nucleic acid molecules comprises:

a) adding a polynucleotide tail to the ends of the segments of genomic ~~material~~ nucleic acids to generate zero-th order derivative strands,

b) subjecting the zero-th order derivative strands of step (a) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (a) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

c) subjecting the product of step (b) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (a) under conditions which promote the formation of only one complement, wherein the primers of step (c) have different nucleotide sequences than the primers of step (b), and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (b) and (c) comprise substantially unique primers, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein adding the polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Currently Amended) The method of claim 1, wherein i) the tagged nucleic acid molecules are subject to hybrid capture prior to PCR

or ii) amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to ~~sequencing in~~ step (d).

25. (Cancelled)

26. (Currently Amended) The method of claim 1, wherein the substantially unique tags ~~sequences~~ further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to PCR amplification or prior to said sequencing of the product of step (c).

28. (Previously Presented) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads according to the sample tag.

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to PCR in step(c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on an array.


IPR2025-01353; IPR2025-01355
CSHL EX2002

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to sequencing in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.


IPR2025-01353; IPR2025-01355
CSHL EX2002

# INCREASING CONFIDENCE OF ALLELE CALLS WITH MOLECULAR COUNTING

## BACKGROUND

Genotyping is an important technique in genetic research for mapping a genome and localizing genes that are linked to inherited characteristics (e.g., genetic diseases). The genotype of a subject generally includes determining alleles for one or more genomic locus based on sequencing data obtained from the subject's DNA. Diploid genomes (e.g., human genomes) may be classified as, for example, homozygous or heterozygous at a genomic locus depending on the number of different alleles they possess for that locus, where heterozygous individuals have two different alleles for a locus and homozygous individuals have two copies of the same allele for the locus.

In genotyping analysis of diploid genomes, the coverage (number of sequencing reads) for a particular genomic locus is used to establish the confidence of an allele call. However, confidence in allele calling is significantly reduced when bias is introduced during sample preparation, e.g., when the starting sample is in limiting amounts and/or when one or more amplification reactions are employed to prepare the sample for sequencing. Thus, in samples having limited amounts of DNA, one may see high coverage (i.e., a high number of sequencing reads) for an allele on one chromosome over the allele on a different chromosome due to amplification bias (e.g., amplification from only a few, or even one, polynucleotide molecule). In this case, coverage alone may be misleading when measuring confidence in an allele call.

The present invention finds use in increasing the confidence in allele calling as well as in other applications based on nucleic acid sequence analysis.

## SUMMARY OF THE INVENTION

Aspects of the present invention include methods and compositions for establishing the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. In these aspects of the invention, a degenerate base region (DBR) is attached to the starting polynucleotide molecules that



Applicant: Cold Spring Harbor Laborator
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Declaration Exhibit 3



IPR2025-01353; IPR2025-01355
CSHL EX2002

are subsequently sequenced (e.g., after certain process steps are performed, e.g., amplification and/or enrichment). The number of different DBR sequences present in a sequencing run establishes the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. DBRs can be used to improve the analysis of many different nucleic acid sequencing applications. For example, DBRs enable the determination of a statistical value for an allele call in genotyping assays that cannot be derived from the read number alone.

## BRIEF DESCRIPTION OF THE DRAWINGS

The invention is best understood from the following detailed description when read in conjunction with the accompanying drawings. Included in the drawings are the following figures:

Figure 1 shows the allele ratio for each MID in samples prepared from the indicated amount of starting material (top of each panel; in nanograms).

Figure 2 shows the fraction of DBR sequences for each MID associated with each allele at a synthetic polymorphic position. Samples were prepared from the indicated amount of starting material (top of each panel; in nanograms).

## DEFINITIONS

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Still, certain elements are defined for the sake of clarity and ease of reference.

Terms and symbols of nucleic acid chemistry, biochemistry, genetics, and molecular biology used herein follow those of standard treatises and texts in the field, e.g. Kornberg and Baker, DNA Replication, Second Edition (W.H. Freeman, New York, 1992); Lehninger, Biochemistry, Second Edition (Worth Publishers, New York, 1975); Strachan and Read, Human Molecular Genetics, Second Edition (Wiley-Liss, New York, 1999); Eckstein, editor, Oligonucleotides and Analogs: A Practical Approach (Oxford



IPR2025-01353; IPR2025-01355
CSHL EX2002

University Press, New York, 1991); Gait, editor, Oligonucleotide Synthesis: A Practical Approach (IRL Press, Oxford, 1984); and the like.

"Amplicon" means the product of a polynucleotide amplification reaction. That is, it is a population of polynucleotides, usually double stranded, that are replicated from one or more starting sequences. The one or more starting sequences may be one or more copies of the same sequence, or it may be a mixture of different sequences. Amplicons may be produced by a variety of amplification reactions whose products are multiple replicates of one or more target nucleic acids. Generally, amplification reactions producing amplicons are "template-driven" in that base pairing of reactants, either nucleotides or oligonucleotides, have complements in a template polynucleotide that are required for the creation of reaction products. In one aspect, template-driven reactions are primer extensions with a nucleic acid polymerase or oligonucleotide ligations with a nucleic acid ligase. Such reactions include, but are not limited to, polymerase chain reactions (PCRs), linear polymerase reactions, nucleic acid sequence-based amplification (NASBAs), rolling circle amplifications, and the like, disclosed in the following references that are incorporated herein by reference: Mullis et al, U.S. patents 4,683,195; 4,965,188; 4,683,202; 4,800,159 (PCR); Gelfand et al, U.S. patent 5,210,015 (real-time PCR with "TAQMAN™" probes); Wittwer et al, U.S. patent 6,174,670; Kacian et al, U.S. patent 5,399,491 ("NASBA"); Lizardi, U.S. patent 5,854,033; Aono et al, Japanese patent publ. JP 4-262799 (rolling circle amplification); and the like. In one aspect, amplicons of the invention are produced by PCRs. An amplification reaction may be a "real-time" amplification if a detection chemistry is available that permits a reaction product to be measured as the amplification reaction progresses, e.g. "real-time PCR" described below, or "real-time NASBA" as described in Leone et al, Nucleic Acids Research, 26: 2150-2155 (1998), and like references. As used herein, the term "amplifying" means performing an amplification reaction. A "reaction mixture" means a solution containing all the necessary reactants for performing a reaction, which may include, but not be limited to, buffering agents to maintain pH at a selected level during a reaction, salts, co-factors, scavengers, and the like.

The term "assessing" includes any form of measurement, and includes determining if an element is present or not. The terms "determining", "measuring",

"evaluating", "assessing" and "assaying" are used interchangeably and includes quantitative and qualitative determinations. Assessing may be relative or absolute. "Assessing the presence of" includes determining the amount of something present, and/or determining whether it is present or absent. As used herein, the terms "determining," "measuring," and "assessing," and "assaying" are used interchangeably and include both quantitative and qualitative determinations.

Polynucleotides that are "asymmetrically tagged" have left and right adapter domains that are not identical. This process is referred to generically as attaching adapters asymmetrically or asymmetrically tagging a polynucleotide, e.g., a polynucleotide fragment. Production of polynucleotides having asymmetric adapter termini may be achieved in any convenient manner. Exemplary asymmetric adapters are described in: U.S. Patents 5,712,126 and 6,372,434; U.S. Patent Publications 2007/0128624 and 2007/0172839; and PCT publication WO/2009/032167; all of which are incorporated by reference herein in their entirety. In certain embodiments, the asymmetric adapters employed are those described in U.S. Patent Application Ser. No. 12/432,080, filed on April 29, 2009, incorporated herein by reference in its entirety.

As one example, a user of the subject invention may use an asymmetric adapter to tag polynucleotides. An "asymmetric adapter" is one that, when ligated to both ends of a double stranded nucleic acid fragment, will lead to the production of primer extension or amplification products that have non-identical sequences flanking the genomic insert of interest. The ligation is usually followed by subsequent processing steps so as to generate the non-identical terminal adapter sequences. For example, replication of an asymmetric adapter attached fragment(s) results in polynucleotide products in which there is at least one nucleic acid sequence difference, or nucleotide/nucleoside modification, between the terminal adapter sequences. Attaching adapters asymmetrically to polynucleotides (e.g., polynucleotide fragments) results in polynucleotides that have one or more adapter sequences on one end (e.g., one or more region or domain, e.g., a primer binding site) that are either not present or have a different nucleic acid sequence as compared to the adapter sequence on the other end. It is noted that an adapter that is termed an "asymmetric adapter" is not necessarily itself structurally asymmetric, nor does the mere act of attaching an asymmetric adapter to a polynucleotide

fragment render it immediately asymmetric. Rather, an asymmetric adapter-attached polynucleotide, which has an identical asymmetric adapter at each end, produces replication products (or isolated single stranded polynucleotides) that are asymmetric with respect to the adapter sequences on opposite ends (e.g., after at least one round of amplification/primer extension).

Any convenient asymmetric adapter, or process for attaching adapters asymmetrically, may be employed in practicing the present invention. Exemplary asymmetric adapters are described in: U.S. Patents 5,712,126 and 6,372,434; U.S. Patent Publications 2007/0128624 and 2007/0172839; and PCT publication WO/2009/032167; all of which are incorporated by reference herein in their entirety. In certain embodiments, the asymmetric adapters employed are those described in U.S. Patent Application Ser. No. 12/432,080, filed on April 29, 2009, incorporated herein by reference in its entirety.

"Complementary" or "substantially complementary" refers to the hybridization or base pairing or the formation of a duplex between nucleotides or nucleic acids, such as, for instance, between the two strands of a double stranded DNA molecule or between an oligonucleotide primer and a primer binding site on a single stranded nucleic acid. Complementary nucleotides are, generally, A and T (or A and U), or C and G. Two single stranded RNA or DNA molecules are said to be substantially complementary when the nucleotides of one strand, optimally aligned and compared and with appropriate nucleotide insertions or deletions, pair with at least about 80% of the nucleotides of the other strand, usually at least about 90% to 95%, and more preferably from about 98 to 100%. Alternatively, substantial complementarity exists when an RNA or DNA strand will hybridize under selective hybridization conditions to its complement. Typically, selective hybridization will occur when there is at least about 65% complementary over a stretch of at least 14 to 25 nucleotides, preferably at least about 75%, more preferably at least about 90% complementary. See, M. Kanehisa Nucleic Acids Res. 12:203 (1984), incorporated herein by reference.

"Duplex" means at least two oligonucleotides and/or polynucleotides that are fully or partially complementary undergo Watson-Crick type base pairing among all or most of their nucleotides so that a stable complex is formed. The terms "annealing" and



IPR2025-01353; IPR2025-01355
CSHL EX2002

"hybridization" are used interchangeably to mean the formation of a stable duplex. "Perfectly matched" in reference to a duplex means that the poly- or oligonucleotide strands making up the duplex form a double stranded structure with one another such that every nucleotide in each strand undergoes Watson-Crick base pairing with a nucleotide in the other strand. A stable duplex can include Watson-Crick base pairing and/or non-Watson-Crick base pairing between the strands of the duplex (where base pairing means the forming hydrogen bonds). In certain embodiments, a non-Watson-Crick base pair includes a nucleoside analog, such as deoxyinosine, 2, 6-diaminopurine, PNAs, LNA's and the like. In certain embodiments, a non-Watson-Crick base pair includes a "wobble base", such as deoxyinosine, 8-oxo-dA, 8-oxo-dG and the like, where by "wobble base" is meant a nucleic acid base that can base pair with a first nucleotide base in a complementary nucleic acid strand but that, when employed as a template strand for nucleic acid synthesis, leads to the incorporation of a second, different nucleotide base into the synthesizing strand (wobble bases are described in further detail below). A "mismatch" in a duplex between two oligonucleotides or polynucleotides means that a pair of nucleotides in the duplex fails to undergo Watson-Crick bonding.

"Genetic locus," "locus," or "locus of interest" in reference to a genome or target polynucleotide, means a contiguous sub-region or segment of the genome or target polynucleotide. As used herein, genetic locus, locus, or locus of interest may refer to the position of a nucleotide, a gene or a portion of a gene in a genome, including mitochondrial DNA or other non-chromosomal DNA (e.g., bacterial plasmid), or it may refer to any contiguous portion of genomic sequence whether or not it is within, or associated with, a gene. A genetic locus, locus, or locus of interest can be from a single nucleotide to a segment of a few hundred or a few thousand nucleotides in length or more. In general, a locus of interest will have a reference sequence associated with it (see description of "reference sequence" below).

"Kit" refers to any delivery system for delivering materials or reagents for carrying out a method of the invention. In the context of reaction assays, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (e.g., probes, enzymes, etc. in the appropriate containers) and/or supporting materials (e.g., buffers, written instructions for performing the assay etc.) from one

location to another. For example, kits include one or more enclosures (e.g., boxes) containing the relevant reaction reagents and/or supporting materials. Such contents may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second container contains probes.

"Ligation" means to form a covalent bond or linkage between the termini of two or more nucleic acids, e.g. oligonucleotides and/or polynucleotides. The nature of the bond or linkage may vary widely and the ligation may be carried out enzymatically or chemically. As used herein, ligations are usually carried out enzymatically to form a phosphodiester linkage between a 5' carbon of a terminal nucleotide of one oligonucleotide with 3' carbon of another oligonucleotide. A variety of template-driven ligation reactions are described in the following references, which are incorporated by reference: Whiteley et al, U.S. patent 4,883,750; Letsinger et al, U.S. patent 5,476,930; Fung et al, U.S. patent 5,593,826; Kool, U.S. patent 5,426,180; Landegren et al, U.S. patent 5,871,921; Xu and Kool, Nucleic Acids Research, 27: 875-881 (1999); Higgins et al, Methods in Enzymology, 68: 50-71 (1979); Engler et al, The Enzymes, 15: 3-29 (1982); and Namsaraev, U.S. patent publication 2004/0110213.

"Multiplex Identifier" (MID) as used herein refers to a tag or combination of tags associated with a polynucleotide whose identity (e.g., the tag DNA sequence) can be used to differentiate polynucleotides in a sample. In certain embodiments, the MID on a polynucleotide is used to identify the source from which the polynucleotide is derived. For example, a nucleic acid sample may be a pool of polynucleotides derived from different sources, (e.g., polynucleotides derived from different individuals, different tissues or cells, or polynucleotides isolated at different times points), where the polynucleotides from each different source are tagged with a unique MID. As such, a MID provides a correlation between a polynucleotide and its source. In certain embodiments, MIDs are employed to uniquely tag each individual polynucleotide in a sample. Identification of the number of unique MIDs in a sample can provide a readout of how many individual polynucleotides are present in the sample (or from how many original polynucleotides a manipulated polynucleotide sample was derived; see, e.g., U.S. Patent No. 7,537,897, issued on May 26, 2009, incorporated herein by reference in its

entirety). MIDs are typically comprised of nucleotide bases and can range in length from 2 to 100 nucleotide bases or more and may include multiple subunits, where each different MID has a distinct identity and/or order of subunits. Exemplary nucleic acid tags that find use as MIDs are described in U.S. Patent 7,544,473, issued on June 6, 2009, and titled "Nucleic Acid Analysis Using Sequence Tokens", as well as U.S. Patent 7,393,665, issued on July 1, 2008, and titled "Methods and Compositions for Tagging and Identifying Polynucleotides", both of which are incorporated herein by reference in their entirety for their description of nucleic acid tags and their use in identifying polynucleotides. In certain embodiments, a set of MIDs employed to tag a plurality of samples need not have any particular common property (e.g., Tm, length, base composition, etc.), as the methods described herein can accommodate a wide variety of unique MID sets. It is emphasized here that MIDs need only be unique within a given experiment. Thus, the same MID may be used to tag a different sample being processed in a different experiment. In addition, in certain experiments, a user may use the same MID to tag a subset of different samples within the same experiment. For example, all samples derived from individuals having a specific phenotype may be tagged with the same MID, e.g., all samples derived from control (or wildtype) subjects can be tagged with a first MID while subjects having a disease condition can be tagged with a second MID (different than the first MID). As another example, it may be desirable to tag different samples derived from the same source with different MIDs (e.g., samples derived over time or derived from different sites within a tissue). Further, MIDs can be generated in a variety of different ways, e.g., by a combinatorial tagging approach in which one MID is attached by ligation and a second MID is attached by primer extension. Thus, MIDs can be designed and implemented in a variety of different ways to track polynucleotide fragments during processing and analysis, and thus no limitation in this regard is intended.

"Nucleoside" as used herein includes the natural nucleosides, including 2'-deoxy and 2'-hydroxyl forms, e.g. as described in Kornberg and Baker, DNA Replication, 2nd Ed. (Freeman, San Francisco, 1992). "Analogs" in reference to nucleosides includes synthetic nucleosides having modified base moieties and/or modified sugar moieties, e.g. described by Scheit, Nucleotide Analogs (John Wiley, New York, 1980); Uhlman and

Peyman, Chemical Reviews, 90: 543-584 (1990), or the like, with the proviso that they are capable of specific hybridization. Such analogs include synthetic nucleosides designed to enhance binding properties, reduce complexity, increase specificity, and the like. Polynucleotides comprising analogs with enhanced hybridization or nuclease resistance properties are described in Uhlman and Peyman (cited above); Crooke et al, Exp. Opin. Ther. Patents, 6: 855-870 (1996); Mesmaeker et al, Current Opinion in Structual Biology, 5: 343-355 (1995); and the like. Exemplary types of polynucleotides that are capable of enhancing duplex stability include oligonucleotide N3'→P5' phosphoramidates (referred to herein as "amidates"), peptide nucleic acids (referred to herein as "PNAs"), oligo-2'-O-alkylribonucleotides, polynucleotides containing C-5 propynylpyrimidines, locked nucleic acids ("LNAs"), and like compounds. Such oligonucleotides are either available commercially or may be synthesized using methods described in the literature.

"Polymerase chain reaction," or "PCR," means a reaction for the in vitro amplification of specific DNA sequences by the simultaneous primer extension of complementary strands of DNA. In other words, PCR is a reaction for making multiple copies or replicates of a target nucleic acid flanked by primer binding sites, such reaction comprising one or more repetitions of the following steps: (i) denaturing the target nucleic acid, (ii) annealing primers to the primer binding sites, and (iii) extending the primers by a nucleic acid polymerase in the presence of nucleoside triphosphates. Usually, the reaction is cycled through different temperatures optimized for each step in a thermal cycler instrument. Particular temperatures, durations at each step, and rates of change between steps depend on many factors well-known to those of ordinary skill in the art, e.g. exemplified by the references: McPherson et al, editors, PCR: A Practical Approach and PCR2: A Practical Approach (IRL Press, Oxford, 1991 and 1995, respectively). For example, in a conventional PCR using Taq DNA polymerase, a double stranded target nucleic acid may be denatured at a temperature >90°C, primers annealed at a temperature in the range 50-75°C, and primers extended at a temperature in the range 72-78°C. The term "PCR" encompasses derivative forms of the reaction, including but not limited to, RT-PCR, real-time PCR, nested PCR, quantitative PCR, multiplexed PCR, and the like. Reaction volumes range from a few nanoliters, e.g. 2 nL, to a few hundred

μL, e.g. 200 μL. "Reverse transcription PCR," or "RT-PCR," means a PCR that is preceded by a reverse transcription reaction that converts a target RNA to a complementary single stranded DNA, which is then amplified, e.g. Tecott et al, U.S. patent 5,168,038, which patent is incorporated herein by reference. "Real-time PCR" means a PCR for which the amount of reaction product, i.e. amplicon, is monitored as the reaction proceeds. There are many forms of real-time PCR that differ mainly in the detection chemistries used for monitoring the reaction product, e.g. Gelfand et al, U.S. patent 5,210,015 ("TAQMAN™"); Wittwer et al, U.S. patents 6,174,670 and 6,569,627 (intercalating dyes); Tyagi et al, U.S. patent 5,925,517 (molecular beacons); which patents are incorporated herein by reference. Detection chemistries for real-time PCR are reviewed in Mackay et al, Nucleic Acids Research, 30: 1292-1305 (2002), which is also incorporated herein by reference. "Nested PCR" means a two-stage PCR wherein the amplicon of a first PCR becomes the sample for a second PCR using a new set of primers, at least one of which binds to an interior location of the first amplicon. As used herein, "initial primers" in reference to a nested amplification reaction mean the primers used to generate a first amplicon, and "secondary primers" mean the one or more primers used to generate a second, or nested, amplicon. "Multiplexed PCR" means a PCR wherein multiple target sequences (or a single target sequence and one or more reference sequences) are simultaneously carried out in the same reaction mixture, e.g. Bernard et al, Anal. Biochem., 273: 221-228 (1999)(two-color real-time PCR). Usually, distinct sets of primers are employed for each sequence being amplified.

"Polynucleotide" or "oligonucleotide" is used interchangeably and each means a linear polymer of nucleotide monomers. Monomers making up polynucleotides and oligonucleotides are capable of specifically binding to a natural polynucleotide by way of a regular pattern of monomer-to-monomer interactions, such as Watson-Crick type of base pairing, base stacking, Hoogsteen or reverse Hoogsteen types of base pairing, wobble base pairing, or the like. As described in detail below, by "wobble base" is meant a nucleic acid base that can base pair with a first nucleotide base in a complementary nucleic acid strand but that, when employed as a template strand for nucleic acid synthesis, leads to the incorporation of a second, different nucleotide base into the synthesizing strand. Such monomers and their internucleosidic linkages may be naturally





IPR2025-01353; IPR2025-01355
CSHL EX2002

occurring or may be analogs thereof, e.g. naturally occurring or non-naturally occurring analogs. Non-naturally occurring analogs may include peptide nucleic acids (PNAs, e.g., as described in U.S. Patent 5,539,082, incorporated herein by reference), locked nucleic acids (LNAs, e.g., as described in U.S. Patent 6,670,461, incorporated herein by reference), phosphorothioate internucleosidic linkages, bases containing linking groups permitting the attachment of labels, such as fluorophores, or haptens, and the like. Whenever the use of an oligonucleotide or polynucleotide requires enzymatic processing, such as extension by a polymerase, ligation by a ligase, or the like, one of ordinary skill would understand that oligonucleotides or polynucleotides in those instances would not contain certain analogs of internucleosidic linkages, sugar moieties, or bases at any or some positions. Polynucleotides typically range in size from a few monomeric units, e.g. 5-40, when they are usually referred to as "oligonucleotides," to several thousand monomeric units. Whenever a polynucleotide or oligonucleotide is represented by a sequence of letters (upper or lower case), such as "ATGCCTG," it will be understood that the nucleotides are in 5'→3' order from left to right and that "A" denotes deoxyadenosine, "C" denotes deoxycytidine, "G" denotes deoxyguanosine, and "T" denotes thymidine, "I" denotes deoxyinosine, "U" denotes uridine, unless otherwise indicated or obvious from context. Unless otherwise noted the terminology and atom numbering conventions will follow those disclosed in Strachan and Read, Human Molecular Genetics 2 (Wiley-Liss, New York, 1999). Usually polynucleotides comprise the four natural nucleosides (e.g. deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine for DNA or their ribose counterparts for RNA) linked by phosphodiester linkages; however, they may also comprise non-natural nucleotide analogs, e.g. including modified bases, sugars, or internucleosidic linkages. It is clear to those skilled in the art that where an enzyme has specific oligonucleotide or polynucleotide substrate requirements for activity, e.g. single stranded DNA, RNA/DNA duplex, or the like, then selection of appropriate composition for the oligonucleotide or polynucleotide substrates is well within the knowledge of one of ordinary skill, especially with guidance from treatises, such as Sambrook et al, Molecular Cloning, Second Edition (Cold Spring Harbor Laboratory, New York, 1989), and like references.



IPR2025-01353; IPR2025-01355
CSHL EX2002

"Primer" means an oligonucleotide, either natural or synthetic, that is capable, upon forming a duplex with a polynucleotide template, of acting as a point of initiation of nucleic acid synthesis and being extended from its 3' end along the template so that an extended duplex is formed. The sequence of nucleotides added during the extension process is determined by the sequence of the template polynucleotide. Usually primers are extended by a DNA polymerase. Primers are generally of a length compatible with its use in synthesis of primer extension products, and are usually are in the range of between 8 to 100 nucleotides in length, such as 10 to 75, 15 to 60, 15 to 40, 18 to 30, 20 to 40, 21 to 50, 22 to 45, 25 to 40, and so on, more typically in the range of between 18-40, 20-35, 21-30 nucleotides long, and any length between the stated ranges. Typical primers can be in the range of between 10-50 nucleotides long, such as 15-45, 18-40, 20-30, 21-25 and so on, and any length between the stated ranges. In some embodiments, the primers are usually not more than about 10, 12, 15, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, or 70 nucleotides in length.

Primers are usually single-stranded for maximum efficiency in amplification, but may alternatively be double-stranded. If double-stranded, the primer is usually first treated to separate its strands before being used to prepare extension products. This denaturation step is typically affected by heat, but may alternatively be carried out using alkali, followed by neutralization. Thus, a "primer" is complementary to a template, and complexes by hydrogen bonding or hybridization with the template to give a primer/template complex for initiation of synthesis by a polymerase, which is extended by the addition of covalently bonded bases linked at its 3' end complementary to the template in the process of DNA synthesis.

A "primer pair" as used herein refers to first and second primers having nucleic acid sequence suitable for nucleic acid-based amplification of a target nucleic acid. Such primer pairs generally include a first primer having a sequence that is the same or similar to that of a first portion of a target nucleic acid, and a second primer having a sequence that is complementary to a second portion of a target nucleic acid to provide for amplification of the target nucleic acid or a fragment thereof. Reference to "first" and "second" primers herein is arbitrary, unless specifically indicated otherwise. For example, the first primer can be designed as a "forward primer" (which initiates nucleic

acid synthesis from a 5' end of the target nucleic acid) or as a "reverse primer" (which initiates nucleic acid synthesis from a 5' end of the extension product produced from synthesis initiated from the forward primer). Likewise, the second primer can be designed as a forward primer or a reverse primer.

"Readout" means a parameter, or parameters, which are measured and/or detected that can be converted to a number or value. In some contexts, readout may refer to an actual numerical representation of such collected or recorded data. For example, a readout of fluorescent intensity signals from a microarray is the address and fluorescence intensity of a signal being generated at each hybridization site of the microarray; thus, such a readout may be registered or stored in various ways, for example, as an image of the microarray, as a table of numbers, or the like.

"Reflex site", "reflex sequence" and equivalents are used to indicate one or more sequences present in a polynucleotide that are employed to move a domain intra-molecularly from its initial location to a different location in the polynucleotide. The use of reflex sequences is described in detail in U.S. provisional applications 61/235,595 and 61/288,792, filed on August 20, 2009 and December 21, 2009, respectively, and entitled "Compositions and Methods for Intramolecular Nucleic Acid Rearrangement Using Reflex Sequences", both of which are incorporated herein by reference. In certain embodiments, a reflex sequence is chosen so as to be distinct from other sequences in the polynucleotide (i.e., with little sequence homology to other sequences likely to be present in the polynucleotide, e.g., genomic or sub-genomic sequences to be processed). As such, a reflex sequence should be selected so as to not hybridize to any sequence except its complement under the conditions employed in the reflex processes. The reflex sequence may be a synthetic or artificially generated sequence (e.g., added to a polynucleotide in an adapter domain) or a sequence present normally in a polynucleotide being assayed (e.g., a sequence present within a region of interest in a polynucleotide being assayed). In the reflex system, a complement to the reflex sequence is present (e.g., inserted in an adapter domain) on the same strand of the polynucleotide as the reflex sequence (e.g., the same strand of a double-stranded polynucleotide or on the same single stranded polynucleotide), where the complement is placed in a particular location so as to facilitate an intramolecular binding and polymerization event on such particular

strand. Reflex sequences employed in the reflex process described herein can thus have a wide range of lengths and sequences. Reflex sequences may range from 5 to 200 nucleotide bases in length.

"Solid support", "support", and "solid phase support" are used interchangeably and refer to a material or group of materials having a rigid or semi-rigid surface or surfaces. In many embodiments, at least one surface of the solid support will be substantially flat, although in some embodiments it may be desirable to physically separate synthesis regions for different compounds with, for example, wells, raised regions, pins, etched trenches, or the like. According to other embodiments, the solid support(s) will take the form of beads, resins, gels, microspheres, or other geometric configurations. Microarrays usually comprise at least one planar solid phase support, such as a glass microscope slide.

"Specific" or "specificity" in reference to the binding of one molecule to another molecule, such as a labeled target sequence for a probe, means the recognition, contact, and formation of a stable complex between the two molecules, together with substantially less recognition, contact, or complex formation of that molecule with other molecules. In one aspect, "specific" in reference to the binding of a first molecule to a second molecule means that to the extent the first molecule recognizes and forms a complex with another molecule in a reaction or sample, it forms the largest number of the complexes with the second molecule. Preferably, this largest number is at least fifty percent. Generally, molecules involved in a specific binding event have areas on their surfaces or in cavities giving rise to specific recognition between the molecules binding to each other. Examples of specific binding include antibody-antigen interactions, enzyme-substrate interactions, formation of duplexes or triplexes among polynucleotides and/or oligonucleotides, biotin-avidin or biotin-streptavidin interactions, receptor-ligand interactions, and the like. As used herein, "contact" in reference to specificity or specific binding means two molecules are close enough that weak noncovalent chemical interactions, such as Van der Waal forces, hydrogen bonding, base-stacking interactions, ionic and hydrophobic interactions, and the like, dominate the interaction of the molecules.

As used herein, the term "$T_m$" is used in reference to the "melting temperature." The melting temperature is the temperature (e.g., as measured in °C) at which a

population of double-stranded nucleic acid molecules becomes half dissociated into single strands. Several equations for calculating the $T_m$ of nucleic acids are known in the art (see e.g., Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization (1985). Other references (e.g., Allawi, H.T. & SantaLucia, J., Jr., Biochemistry 36, 10581-94 (1997)) include alternative methods of computation which take structural and environmental, as well as sequence characteristics into account for the calculation of $T_m$.

"Sample" means a quantity of material from a biological, environmental, medical, or patient source in which detection, measurement, or labeling of target nucleic acids is sought. On the one hand it is meant to include a specimen or culture (e.g., microbiological cultures). On the other hand, it is meant to include both biological and environmental samples. A sample may include a specimen of synthetic origin. Biological samples may be animal, including human, fluid, solid (e.g., stool) or tissue, as well as liquid and solid food and feed products and ingredients such as dairy items, vegetables, meat and meat by-products, and waste. Biological samples may include materials taken from a patient including, but not limited to cultures, blood, saliva, cerebral spinal fluid, pleural fluid, milk, lymph, sputum, semen, needle aspirates, and the like. Biological samples may be obtained from all of the various families of domestic animals, as well as feral or wild animals, including, but not limited to, such animals as ungulates, bear, fish, rodents, etc. Environmental samples include environmental material such as surface matter, soil, water and industrial samples, as well as samples obtained from food and dairy processing instruments, apparatus, equipment, utensils, disposable and non-disposable items. These examples are not to be construed as limiting the sample types applicable to the present invention.

The terms "upstream" and "downstream" in describing nucleic acid molecule orientation and/or polymerization are used herein as understood by one of skill in the art. As such, "downstream" generally means proceeding in the 5' to 3' direction, i.e., the direction in which a nucleotide polymerase normally extends a sequence, and "upstream" generally means the converse. For example, a first primer that hybridizes "upstream" of a second primer on the same target nucleic acid molecule is located on the 5' side of the





IPR2025-01353; IPR2025-01355
CSHL EX2002

second primer (and thus nucleic acid polymerization from the first primer proceeds towards the second primer).

It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely", "only" and the like in connection with the recitation of claim elements, or the use of a "negative" limitation.

## DETAILED DESCRIPTION OF THE INVENTION

Before the present invention is described, it is to be understood that this invention is not limited to particular embodiments described, 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, since the scope of the present invention will be limited only by the appended claims.

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in that stated range is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included or excluded in the range, and each range where either, neither or both limits are included in the smaller ranges is also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, some potential and preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. It is

understood that the present disclosure supersedes any disclosure of an incorporated publication to the extent there is a contradiction.

It must be noted that as used herein and in 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 nucleic acid" includes a plurality of such nucleic acids and reference to "the compound" includes reference to one or more compounds and equivalents thereof known to those skilled in the art, and so forth.

The practice of the present invention may employ, unless otherwise indicated, conventional techniques and descriptions of organic chemistry, polymer technology, molecular biology (including recombinant techniques), cell biology, biochemistry, and immunology, which are within the skill of the art. Such conventional techniques include polymer array synthesis, hybridization, ligation, and detection of hybridization using a label. Specific illustrations of suitable techniques can be had by reference to the example herein below. However, other equivalent conventional procedures can, of course, also be used. Such conventional techniques and descriptions can be found in standard laboratory manuals such as *Genome Analysis: A Laboratory Manual Series* (Vols. I-IV), *Using Antibodies: A Laboratory Manual, Cells: A Laboratory Manual, PCR Primer: A Laboratory Manual, and Molecular Cloning: A Laboratory Manual* (all from Cold Spring Harbor Laboratory Press), Stryer, L. (1995) *Biochemistry* (4th Ed.) Freeman, New York, Gait, "*Oligonucleotide Synthesis: A Practical Approach*" 1984, IRL Press, London, Nelson and Cox (2000), Lehninger, A., *Principles of Biochemistry* 3rd Ed., W. H. Freeman Pub., New York, N.Y. and Berg et al. (2002) *Biochemistry,* 5th Ed., W. H. Freeman Pub., New York, N.Y., all of which are herein incorporated in their entirety by reference for all purposes.

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.





IPR2025-01353; IPR2025-01355
CSHL EX2002

As summarized above, aspects of the present invention are drawn to the use of degenerate nucleotide bases (e.g., in a degenerate base region, or DBR) added to polynucleotides undergoing sequence analysis that find use in establishing the number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. Including a DBR in polynucleotides undergoing sequencing analysis finds use in a variety of genetic analyses, including increasing the confidence of allele calling by providing a mechanism to determine a statistical value for an allele call, a value that cannot be derived from the read number alone. The DBR may be added to a polynucleotide in any convenient manner, including as part of an adapter (or pool of adapters) attached to the polynucleotides being sequenced, e.g., the DBR can be in an adapter that also includes a sequencing primer site.

Nucleic Acids

The present invention (as described in detail below) can be employed for the manipulation and analysis of nucleic acid sequences of interest (or polynucleotides) from virtually any nucleic acid source, including but not limited to genomic DNA, complementary DNA (cDNA), RNA (e.g., messenger RNA, ribosomal RNA, short interfering RNA, microRNA, etc.), plasmid DNA, mitochondrial DNA, synthetic DNA, etc. Furthermore, any organism, organic material or nucleic acid-containing substance can be used as a source of nucleic acids to be processed in accordance with the present invention including, but not limited to, plants, animals (e.g., reptiles, mammals, insects, worms, fish, etc.), tissue samples, bacteria, fungi (e.g., yeast), phage, viruses, cadaveric tissue, archaeological/ancient samples, etc. In certain embodiments, the nucleic acids in the nucleic acid sample are derived from a mammal, where in certain embodiments the mammal is a human.

In certain embodiments, the nucleic acid sequences are enriched. By enriched is meant that the nucleic acids (e.g., in a polynucleotide sample) are subjected to a process that reduces the complexity of the nucleic acids, generally by increasing the relative concentration of particular nucleic acid species in the sample (e.g., having a specific locus of interest, including a specific nucleic acid sequence, lacking a locus or sequence,



IPR2025-01353; IPR2025-01355
CSHL EX2002

being within a specific size range, etc.). There are a wide variety of ways to enrich nucleic acids having a specific characteristic(s) or sequence, and as such any convenient method to accomplish this may be employed. The enrichment (or complexity reduction) can take place at any of a number of steps in the process, and will be determined by the desires of the user. For example, enrichment can take place in individual parental samples (e.g., untagged nucleic acids prior to adaptor ligation) or in multiplexed samples (e.g., nucleic acids tagged with adapter sequences having MID; MID are described in further detail below).

In certain embodiments, nucleic acids in the nucleic acid sample are amplified prior to analysis. In certain of these embodiments, the amplification reaction also serves to enrich a starting nucleic acid sample for a sequence or locus of interest. For example, a starting nucleic acid sample can be subjected to a polymerase chain reaction (PCR) that amplifies one or more region of interest. In certain embodiments, the amplification reaction is an exponential amplification reaction, whereas in certain other embodiments, the amplification reaction is a linear amplification reaction. Any convenient method for performing amplification reactions on a starting nucleic acid sample can be used in practicing the subject invention. In certain embodiments, the nucleic acid polymerase employed in the amplification reaction is a polymerase that has proofreading capability (e.g., phi29 DNA Polymerase, Thermococcus litoralis DNA polymerase, Pyrococcus furiosus DNA polymerase, etc.).

In certain embodiments, the nucleic acid sample being analyzed is derived from a single source (e.g., a single organism, virus, tissue, cell, subject, etc.), whereas in other embodiments, the nucleic acid sample is a pool of nucleic acids extracted from a plurality of sources (e.g., a pool of nucleic acids from a plurality of organisms, tissues, cells, subjects, etc.), where by "plurality" is meant two or more. As such, in certain embodiments, a nucleic acid sample can contain nucleic acids from 2 or more sources, 3 or more sources, 5 or more sources, 10 or more sources, 50 or more sources, 100 or more sources, 500 or more sources, 1000 or more sources, 5000 or more sources, up to and including about 10,000 or more sources.

In certain embodiments, nucleic acid fragments that are to be pooled with nucleic acid fragments derived from a plurality of sources (e.g., a plurality of organisms, tissues,



IPR2025-01353; IPR2025-01355
CSHL EX2002

cells, subjects, etc.), where by "plurality" is meant two or more. In such embodiments, the nucleic acids derived from each source include a multiplex identifier (MID) such that the source from which each tagged nucleic acid fragment was derived can be determined. In such embodiments, each nucleic acid sample source is correlated with a unique MID, where by unique MID is meant that each different MID employed can be differentiated from every other MID employed by virtue of at least one characteristic, e.g., the nucleic acid sequence of the MID. Any type of MID can be used, including but not limited to those described in co-pending U.S. Patent Application Serial Number 11/656,746, filed on January 22, 2007, and titled "Nucleic Acid Analysis Using Sequence Tokens", as well as U.S. Patent 7,393,665, issued on July 1, 2008, and titled "Methods and Compositions for Tagging and Identifying Polynucleotides", both of which are incorporated herein by reference in their entirety for their description of nucleic acid tags and their use in identifying polynucleotides. In certain embodiments, a set of MIDs employed to tag a plurality of samples need not have any particular common property (e.g., $T_m$, length, base composition, etc.), as the asymmetric tagging methods (and many tag readout methods, including but not limited to sequencing of the tag or measuring the length of the tag) can accommodate a wide variety of unique MID sets.

Degenerate Base Region (DBR)

Aspects of the present invention include methods and compositions for establishing the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. In these aspects of the invention, a degenerate base region (DBR) is attached to the starting polynucleotide molecules that are subsequently sequenced (e.g., after certain process steps are performed, e.g., amplification and/or enrichment, e.g., PCR). As detailed below, evaluating the number of different DBR sequences present in a sequencing run allows the establishment of the minimum number of different starting polynucleotides that have been sequenced for a particular polynucleotide (or region of interest; ROI). This number can be used to give a statistical measure of allele calls, thus increasing the confidence in making such allele determinations (e.g., when calling homozygous alleles).



IPR2025-01353; IPR2025-01355
CSHL EX2002

DNA sequencing typically includes a step of attaching an adapter to the polynucleotides in a sample to be sequenced, where the adaptor contains a sequencing primer site (e.g., by ligation). As used herein, a "sequencing primer site" is a region of a polynucleotide that is either identical to or complementary to the sequence of a sequencing primer (when in a single stranded form) or a double stranded region formed between a sequencing primer sequence and its complement. The specific orientation of a sequencing primer site can be inferred by those of ordinary skill in the art from the structural features of the specific polynucleotide containing the sequencing primer site.

In addition to the sequencing primer site, a degenerate base region (DBR) is also attached to the polynucleotides, either as part of the adapter containing the sequencing primer site or independently (e.g., in a second adapter attached to the polynucleotide). Any convenient method for attaching or adding a DBR to the polynucleotides may be employed. A DBR is a region that can have a variable base composition or sequence (which may be considered a "random") as compared to other tagged polynucleotides in the sample. The number of different DBRs in a population of polynucleotides in a sample will be dependent on the number of bases in the DBR as well as the potential number of different bases that can be present at each position. Thus, a population of polynucleotides having attached DBRs with two base positions, where each position can be any one of A, C, G and T, will have potentially 16 different DBRs (AA, AC, AG, etc.). DBR may thus include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more bases, including 15 or more, 20 or more, etc. In certain embodiments the DBR is from 3 to 10 bases in length. Moreover, each position in a DBR may have a different base composition. For example, a 4 base DBR may have any of the following compositions: NNNN; NRSN; SWSW; BDHV (see Table 1 below for IUPAC nucleotide code). As such, no limitation with regard to the length or base composition of a DBR is intended.

| IUPAC nucleotide code | Base |
|---|---|
| A | Adenine |
| C | Cytosine |
| G | Guanine |
| T (or U) | Thymine (or Uracil) |





IPR2025-01353; IPR2025-01355
CSHL EX2002

| | |
|---|---|
| R | A or G |
| Y | C or T |
| S | G or C |
| W | A or T |
| K | G or T |
| M | A or C |
| B | C or G or T |
| D | A or G or T |
| H | A or C or T |
| V | A or C or G |
| N | any base |

It is noted here that a DBR may be a single region (i.e., having all nucleotide bases adjacent to one another) or may be present in different location on a polynucleotide (i.e., the bases of the DBR are separated by non-DBR sequences, also called a split DBR). For example, a DBR may have one or more bases in a first adapter at a first location on a polynucleotide and one or more bases in a second adapter at a second location on the same polynucleotide (e.g., the DBR may have bases present at both ends of an asymmetrically tagged polynucleotide, i.e., a polynucleotide having asymmetric adapters). No limitation in this regard is intended.

In certain embodiments in which the DBR is present within an adapter population that includes other functional domains (e.g., sequencing primer site, MID, reflex sequence), the functional domains in the adapter population will be identical to each other whereas the DBR will vary. In other words, unlike the other domains in an adapter population, the DBR has a variable (or random) base composition. By "adapter population", "population of adapters", and the like, is meant a sample of adapter molecules that are designed to be attached to the polynucleotides in a sample.

Generating adapters having a DBR may be achieved in any convenient manner, e.g., using DNA synthesis methods well known in the art (see citations in definitions section above).

Once attached to the polynucleotides in the parent sample, the polynucleotides may be subjected to further processing and ultimately sequenced. Processing steps that



IPR2025-01353; IPR2025-01355
CSHL EX2002

may be performed include any process steps that are desired by a user, e.g., enrichment, amplification, and the like. In the sequencing step, the sequence of the DBR as well as a portion of the polynucleotide (e.g., containing a region of interest) is obtained. Once the sequences are obtained, the number of different DBRs attached to a polynucleotide of interest is determined. This number represents the minimum number of different polynucleotides of interest from the starting parent sample that are represented in the sequencing results.

Consider, for example, a two base DBR having base composition NN (where N is any deoxynucleotide base, i.e., A, G, C or T) employed in sequencing a locus to make an allele call for a particular subject sample (i.e., whether a subject is homozygous or heterozygous at the locus). Although there may be some biases from oligonucleotide synthesis, one can expect there to be 16 different DBRs in the population of adaptors with roughly equal probability (as described above). When a potential homozygous allele call is identified, determining the number of DBRs present in the sequencing run can be used to find out the minimum number of polynucleotide molecules that were actually sequenced (and thus the minimum number that were amplified during the processing steps).

For diploid genomes, allele calling (in the ideal or theoretical case) can be modeled by the binomial distribution. Given that the two alleles copies (X and Y) differ at some site, the probability of observing all Xs or all Ys is given by the formula $( \frac{1}{2} )^c$, where c is the number of observations (reads) of the site. If we observe X ten times at a site (and no Y), we can say that the sample is likely to be homozygous for type X. The probability of an error in this call is therefore $( \frac{1}{2} )^{10}$ (just under one in a thousand).

Our experiments show that low amounts of DNA in the early stage of sample preparation can result in a high coverage of reads all corresponding to one allele, and that this can occur many more times than should be expected according to the binomial distribution. This is due to amplification of a few molecules of DNA (or even a single molecule) that results in a large number of reads derived from a genetic locus on a single chromosome (i.e., only one of the two diploid chromosomes actually present in the sample of interest). The result of this is that the error as a function of coverage deviates wildly from the predicted binomial error.



IPR2025-01353; IPR2025-01355
CSHL EX2002

Using a DBR as described herein will increase the confidence in making allele calls from samples having limited amounts of DNA. For example, if 16 sequencing reads of one allele from a genetic locus all contain the DBR sequence GA, then it is likely that all these reads are from the same parent polynucleotide molecule (and thus a homozygous allele call is not justified). However, if the 16 sequencing reads each have a different DBR sequence, a homozygous call can be made with more confidence, as each read came from a different parent polynucleotide molecule.

It is noted here that in many embodiments, it is not possible to conclude that polynucleotides having identical DBR sequences are derived from the same parent polynucleotide molecule, as multiple identical DBRs may be present in the DBR-attached polynucleotides. For example, if an adapter population that contains a DBR of two N bases is used to tag a sample containing more that 16 polynucleotides, a subset of the tagged polynucleotides will have identical DBRs, and thus it will not be possible to determine that their sequences were derived from different parent polynucleotide molecules.

One exemplary way to determine more accurately the actual number of starting or parent molecules would be to increase the degeneracy of DBRs (i.e., to increase the number of unique sequences in the DBR used to label the particular sample of interest) so that every single molecule is likely to have a different DBR. In any event we can, in exemplary methods, either use the number of observed DBRs or else the probability distribution of the numbers of reads likely to produce that numbers of observed DBRs.

When calculating estimates of whether a particular allele call is a heterozygote or homozygote, one can create/employ an appropriate function $L(r,v)$ that returns the likelihood of a genotype given $r$ reference and $v$ variant reads. When employing DBRs as described herein, a modified function for calculating estimates $L(r', v')$ may be used, where $r'$ is the number of unique DBRs for the reference read and $v'$ is the number of unique DBRs for the variant read. Any convenient function for making allele calls may be employed and modified to employ data regarding DBR reads as described herein.

It is noted here that aspects of the invention can be used to increase the confidence in calling copy number variations in a polynucleotide sample, e.g., a genomic sample. Copy number variations can include genomic rearrangements such as deletions,



IPR2025-01353; IPR2025-01355
CSHL EX2002

duplications, inversions, and translocations or whole chromosome aneuploidies such as monosomies, disomies, trisomies, tetrasomies, and pentasomies. Consider, for example, a duplication event where parents have genotypes AC and CC at a given SNP, and the proband has genotype ACC. In the parent with genotype AC, given sufficient depth of sequencing coverage (i.e. sufficient sequencing reads), the number of DBRs associated with the C-allele and the A-allele are expected to be similar. In the proband, given sufficient depth of sequencing coverage, the number of DBRs associated with the C-allele is expected to be 2-times the number of DBRs associated with the A-allele, which provides evidence for duplication event encompassing the C-allele. The use of DBRs, rather than the number of sequencing reads, provides more confidence in calling a copy number variation since each DBR came from a different parent polynucleotide molecule.

Kits and Systems

Also provided by the subject invention are kits and systems for practicing the subject methods, i.e., for using DBRs to determine the minimum number of different starting polynucleotides that have been sequenced for a particular polynucleotide. As such, systems and kits may include polynucleotides containing DBRs (e.g., adapters) as well as any other functional domains of interest as described herein (e.g., sequencing primer sites, MIDs, reflex sequences, etc.). Systems and kits may also include and regents for performing any steps in attaching adapters to polynucleotides in a parents sample, preparing a parent sample for adapter/DBR attachment, and/or reagents for performing sequencing reactions (e.g., ligases, restriction enzymes, nucleotides, polymerases, primers, sequencing primers, dNTPs, ddNTPs, exonucleases, etc.). The various components of the systems and kits may be present in separate containers or certain compatible components may be precombined into a single container, as desired.

The subject systems and kits may also include one or more other reagents for preparing or processing a nucleic acid sample according to the subject methods. The reagents may include one or more matrices, solvents, sample preparation reagents, buffers, desalting reagents, enzymatic reagents, denaturing reagents, where calibration standards such as positive and negative controls may be provided as well. As such, the kits may include one or more containers such as vials or bottles, with each container



IPR2025-01353; IPR2025-01355
CSHL EX2002

containing a separate component for carrying out a sample processing or preparing step according to the present invention.

In addition to above-mentioned components, the subject kits typically further include instructions for using the components of the kit to practice the subject methods, e.g., to employ DBR as described above. The instructions for practicing the subject methods are generally recorded on a suitable recording medium. For example, the instructions may be printed on a substrate, such as paper or plastic, etc. As such, the instructions may be present in the kits as a package insert, in the labeling of the container of the kit or components thereof (i.e., associated with the packaging or sub-packaging) etc. In other embodiments, the instructions are present as an electronic storage data file present on a suitable computer readable storage medium, e.g. CD-ROM, diskette, etc. In yet other embodiments, the actual instructions are not present in the kit, but means for obtaining the instructions from a remote source, e.g. via the internet, are provided. An example of this embodiment is a kit that includes a web address where the instructions can be viewed and/or from which the instructions can be downloaded. As with the instructions, this means for obtaining the instructions is recorded on a suitable substrate.

In addition to the subject database, programming and instructions, the kits may also include one or more control samples and reagents, e.g., two or more control samples for use in testing the kit.

## EXAMPLES

*Methods*

Two identical samples of mouse genomic DNA were labeled with adapters. One sample used an adapter that has a redundantly synthesized region consisting of 7 bases (RYBDHVB), each of which could be one of two (for R and Y positions) or three (for B, D, H or V positions) bases (or a total of 972 different sequences) followed by the bases AC**A**; the second sample used an adapter that has a redundantly synthesized region consisting of 7 bases (RYBDHVB), each of which could be one of two (for R and Y positions) or three (for B, D, H or V positions) bases (or a total of 972 different sequences) followed by the bases AC**G**. Note the bolded underlined bases correspond to



IPR2025-01353; IPR2025-01355
CSHL EX2002

a synthetic polymorphic site. In these adapters, the sequence RYB serves as the DBR region and the DHVB serves as the MID. Thus, there were 12 possible DBR (2 x 2 x 3) codes and 81 different MIDs present (3 x 3 x 3 x 3).

The two samples were then mixed together in equal amounts to create, in effect, a perfect 50/50 heterozygote of **A** and **G** three bases downstream from the MID (i.e., the DHVB sequence). Different amounts of the mixture (100ng, 300ng, 600ng, 2500ng, 5000ng, and 10,000ng) were subjected to hybridization pull-down reactions followed by amplification through 10 cycles of PCR with TiA and TiB primers. The capture probes employed were 5'-biotinylated 60-mer reverse phase cartridge purified oligonucleotides (BioSearch). After amplification with TiA and TiB, a secondary PCR reaction with TiA and a sequence-specific primer tailed with TiB (5'-CCTATCCCCTGTGTGCCTTGGCAGTCTCAGGGACACCCAGCCAAGACAGC-3') was used to amplify a specific fragment. The PCR fragment generated from each sample in the hybridization pull-down/PCR was sent for 454 Ti Shot sequencing to determine the DBR, MID and A/G allele.

The amplicon sequence from the hybridization pull down/PCR is shown below. The DBR region is underlined, the MID is in bold, and the allele (R, which corresponds to A or G) is bold underline.

CCATCTCATCCCTGCGTGTCTCCGACTCAGRYB**DHVB**AC**R**TAGAATGTGCAT
GGATCGTATGAGCACCTGTGGGCAGGGCAAGTGGCAGATGCCTTAGTGGATC
TCACTGGAAGCCTGGCAGAAAGGTGGAGCTTGAAGGATGTAACGAAAGCCA
GCGGCCAGCAGGACAGACCCAGTGGTGGGGAGCACAGAACTTGTCGGCAGC
TACTCCACCTGAAGGACCGGTGTCTAATCAGCTGCTCTGTGCTTAGCCCCAGA
GCAGGTACAGCTATGGCTACAACTCCCTCCACCATTAGCTTGTTACAGAGAA
GGAAATCGGTCCTTGAGAGGCTGTCTTGGCTGGGTGTCCCTGAGACTGCCAA
GGCACACAGGGGATAGG

*Results*

Figure 1 shows the allele ratio for each MID in the sample. The numbers at the top of each of the 6 panels show the input mass of genomic DNA used (in nanograms).

The abscissa shows, for any particular MID, the fraction of A sequencing reads, or calls, (i.e., the number of reads of one of the polymorphic bases at the synthetic SNP position) to the total number of calls, e.g., the ratio of (A calls)/(A calls + G calls). This is called the allele ratio. The ordinate is the number of MIDs observed with a particular allele ratio (the total number of MIDs, as noted above, is 81). Because the input DNA was known to be at a 50/50 A/G ratio, the allele ratio for each sample should be 0.5.

At low input mass, the allele ratio is distorted – either too high or two low – because the molecules on the input to the first PCR step were limiting and so one of the two alleles was preferentially observed. As mass into the first step increases, both alleles are observed and at closer and closer to the expected 0.5 ratio. This analysis shows that there was considerable allele drop out at 100ngs, 300ngs and 600ngs, while little to no allele drop out occurred at the higher input levels.

Figure 2 shows the fraction of DBR sequences for each MID associated with each allele. Since the input material is nominally 50/50 A and G at the synthetic polymorphic position, we would expect each of the 81 effective MIDs to be associated with 50% A reads and 50% G reads (as noted above). Furthermore, since the 12 DBRs are random and are associated with each of the 81 different MIDs, one would expect, with sufficient input DNA copies, that all 12 DBRs would be observed for the A allele and all 12 for the G allele. Thus, the fraction of DBRs observed for a particular base in the ideal case would be 12/24 or 0.5 for each allele. To the extent that insufficient numbers of molecules go into the processing steps, there may be less than 12 DBRs for each MID allele, and so the ratio can stray away from the ideal.

In Figure 2, the abscissa shows, for any particular MID, the proportion of the DBRs actually observed for each allele divided by the total number of DBRs actually observed. The ordinate is the number of MIDs observed with a particular proportion or DBRs associated with an allele. There are, and should be, a total of 81 different MIDs seen. We observe that at low input mass, one frequently can see a distorted proportion – either too high or two low. This is likely due to limiting number of molecules input into the first PCR step, and so one of the two alleles was preferentially observed. As mass into the first step increases, both alleles are observed more frequently and associated with



IPR2025-01353; IPR2025-01355
CSHL EX2002

each allele is an observation of more and more of the DBRs, and thus a proportion closer to the 0.5.

A comparison of the date in Figures 1 and 2 shows an additional feature of employing DBRs as described herein. For a true heterozygote, one starts to see sooner (i.e., at lower mass) in the DBR analysis (Figure 2) that the distribution clusters around the expected 0.5 ratio. This makes sense because, for example, observation of 6 of the 12 DRBs for the A allele and 4 of the 12 DBRs of the G allele would result in a proportion of (6 / [6+4] ) = 0.60 which, in fact is fairly close to the expected 0.5. The net effect is that the use of DBRs gives much greater confidence in the presence of a true heterozygote or true homozygote.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

Accordingly, the preceding merely illustrates the principles of the invention. It will be appreciated that those skilled in the art will be able to devise various arrangements which, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope. Furthermore, all examples and conditional language recited herein are principally intended to aid the reader in understanding the principles of the invention and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents and equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure. The scope of the present invention, therefore, is not intended to be limited to the exemplary embodiments shown and



IPR2025-01353; IPR2025-01355
CSHL EX2002

described herein. Rather, the scope and spirit of present invention is embodied by the appended claims.





IPR2025-01353; IPR2025-01355
CSHL EX2002

## CLAIMS

That which is claimed is:

1. A method of determining the minimum number of polynucleotide molecules sequenced from a sample, the method comprising:

attaching an adapter to polynucleotide molecules in a sample, wherein the adapter comprises a degenerate base region (DBR) comprising at least one nucleotide base selected from: R, Y, S, W, K, M, B, D, H, V, and N;

sequencing a plurality of the adapter-attached polynucleotides, wherein the sequence of the DBR and at least a portion of the polynucleotide is obtained for each of the plurality of adapter-attached polynucleotides; and

determining the number of distinct DBR sequences present in the plurality of sequenced adapter-attached polynucleotides to determine the minimum number of polynucleotides from the sample that were sequenced in the sequencing step.

2. The method of claim 1, wherein the adapter further comprises a sequencing primer site.

3. The method of claim 1, wherein the adapter further comprises one or more of: an MID, a reflex sequence, and a promoter site.

4. The method of claim 1, wherein the attaching step further comprises:

attaching adapters to multiple different samples, wherein the adapter employed for each different sample comprises a unique MID; and

pooling the multiple different adapter attached samples prior to the sequencing step.

5. The method of claim 1, wherein the DBR comprises at least 2 nucleotide bases, wherein each of the at least 2 nucleotide bases are selected from: R, Y, S, W, K, M, B, D, H, V, and N.

6. The method of claim 5, wherein the DBR comprises from 3 to 20 nucleotide bases, wherein each of the 3 to 10 nucleotide bases is selected from: R, Y, S, W, K, M, B, D, H, V, and N.

7. The method of claim 1, wherein the determined minimum number of polynucleotides from the sample that were sequenced in the sequencing step is employed in a method for allele calling.

8. The method of claim 1, wherein the method further comprises amplifying the adapter-attached polynucleotides prior to the sequencing step.

9. The method of claim 1, wherein the sequencing step comprises performing a next generation sequencing process.

10. The method of claim 1, wherein the sample is genomic DNA sample.

11. The method of claim 4, wherein each of the multiple different samples is a genomic DNA sample derived from a different subject.

12. The method of claim 11, wherein the subject is a human.

13. The method of claim 1, further wherein the sample is enriched for a region of interest prior to the attaching step.

14. The method of claim 1, further comprising enriching the adapter-attached polynucleotides for a region of interest.

15. The method of claim 5, wherein the adapter is an asymmetric adapter, wherein a first adapter domain is present at a first end of the polynucleotide and a second adapter domain is present at a second end of the polynucleotide, wherein the DBR is a split DBR comprising one or more of the at least 2 nucleotide bases in the first adapter domain and one or more of the at least 2 nucleotide bases in the second adapter domain.

# INCREASING CONFIDENCE OF ALLELE CALLS WITH MOLECULAR COUNTING

## ABSTRACT OF THE DISCLOSURE

Aspects of the present invention include methods and compositions for establishing the minimum number of individual polynucleotide molecules originating from the same genomic region of the same original sample that have been sequenced in a particular sequence analysis configuration or process. In these aspects of the invention, a degenerate base region (DBR) is attached to the starting polynucleotide molecules that are subsequently sequenced (e.g., after certain process steps are performed, e.g., amplification and/or enrichment). The number of different DBR sequences present in a sequencing run establishes a minimum number of different starting polynucleotides that have been sequenced. DBRs can be used to enhance numerous different nucleic acid sequence analysis applications, including allowing higher confidence allele call determinations in genotyping applications.

F:\DOCUMENT\CGLC\025PRV\CGLC-025PRV Application as filed.doc



IPR2025-01353; IPR2025-01355
CSHL EX2002

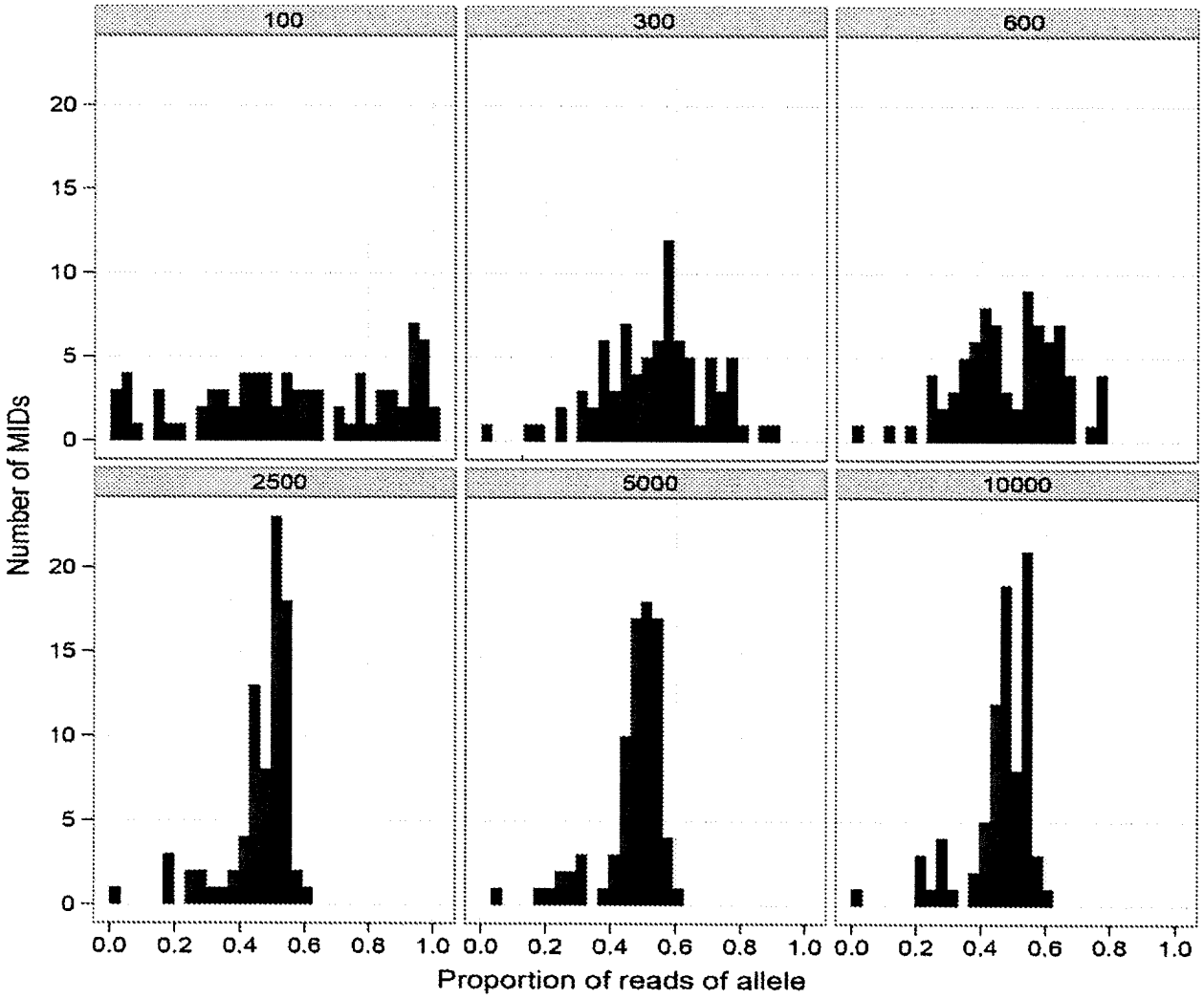

100
300
600
2500
5000
10000
Number of MIDs
0
5
10
15
20
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of reads of allele

Fig. 1

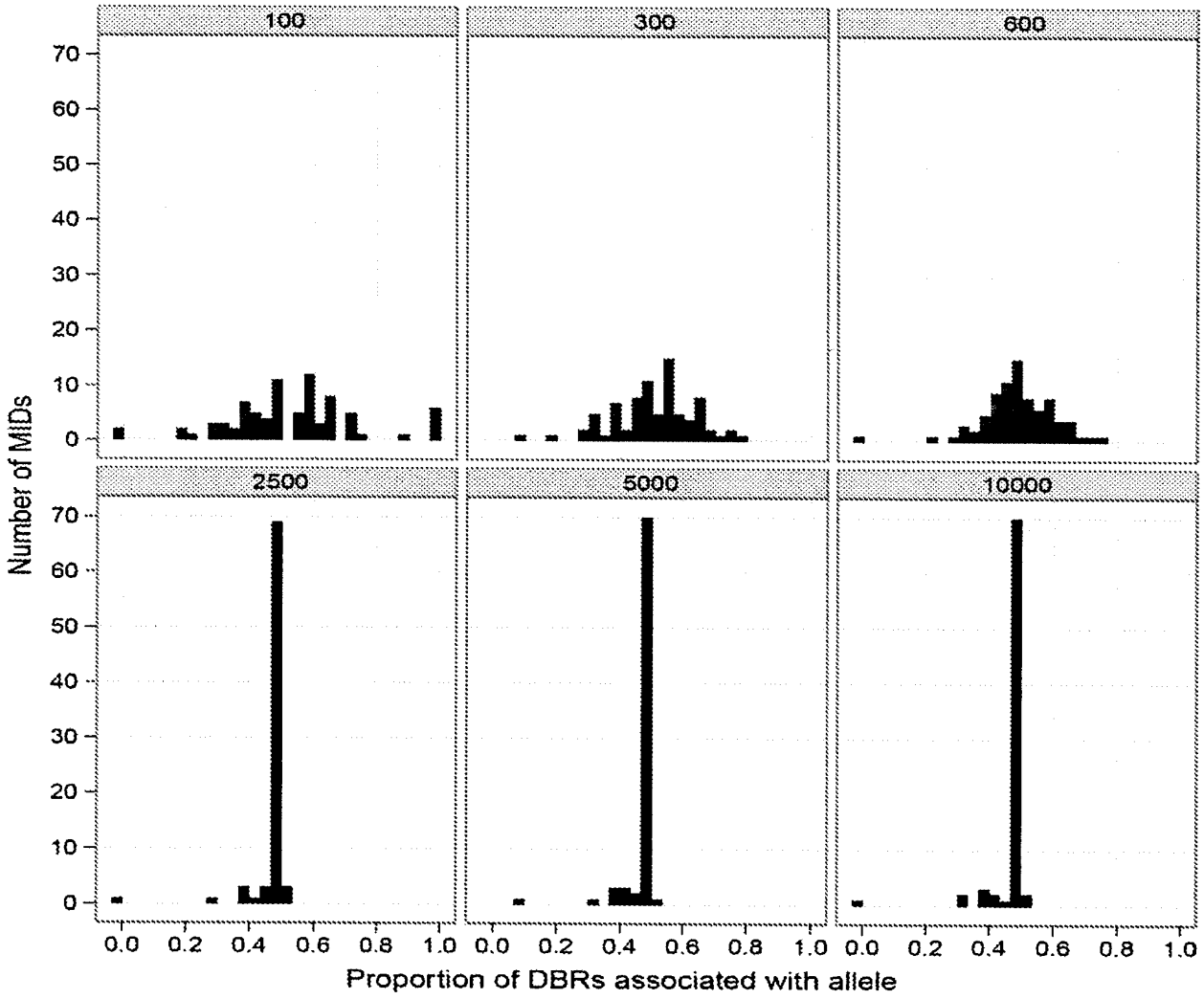



Fig. 2



IPR2025-01353; IPR2025-01355
CSHL EX2002

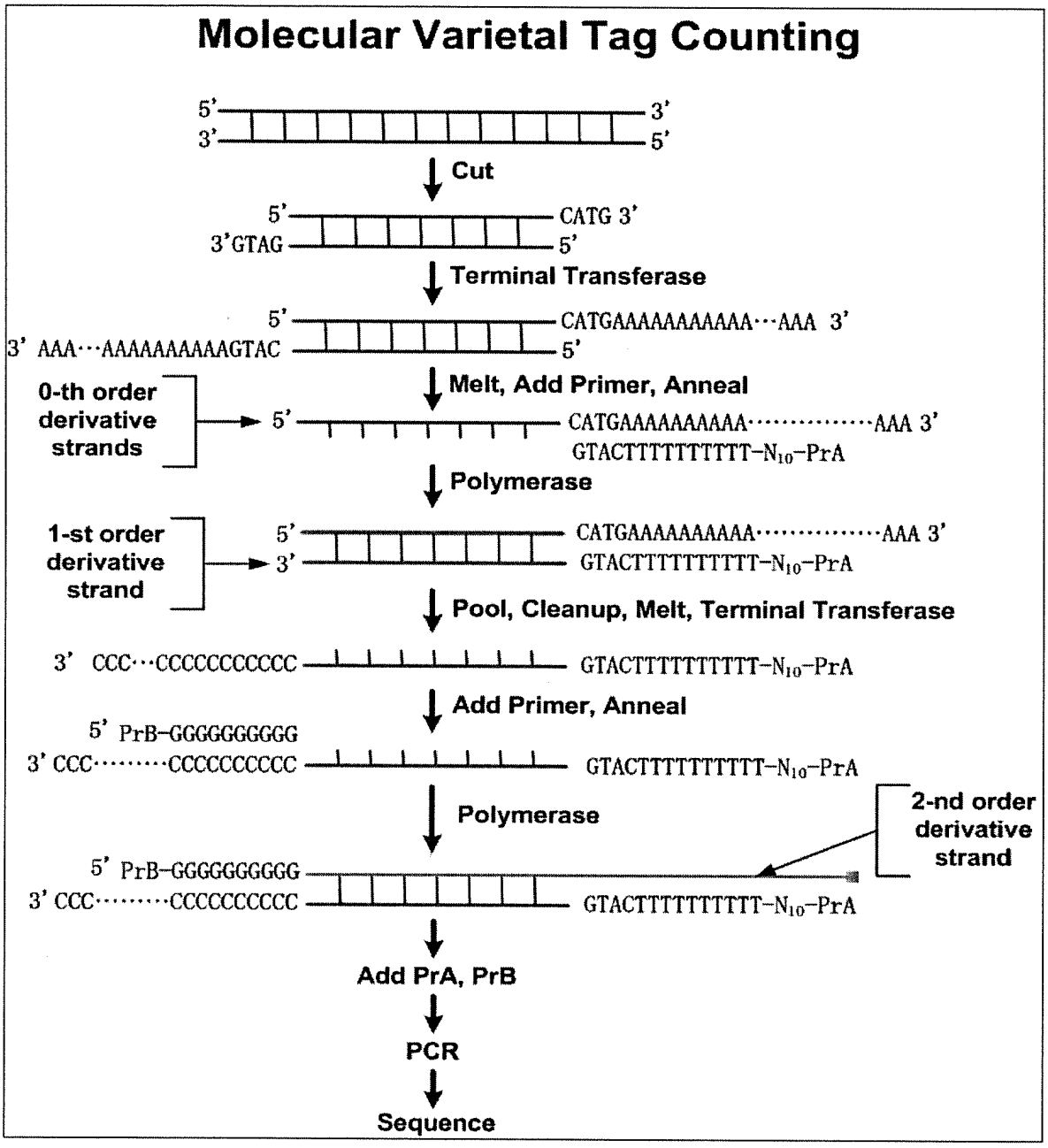

Molecular Varietal Tag Counting
5'
3'
3'
5'
Cut
5'
CATG 3'
3'GTAG
5'
Terminal Transferase
5'
CATGAAAAAAAAAAA···AAA 3'
3' AAA···AAAAAAAAAAGTAC
5'
Melt, Add Primer, Anneal
0-th order derivative strands
5'
CATGAAAAAAAAAA·····AAA 3'
GTACTTTTTTTTTT-N10-PrA
Polymerase
1-st order derivative strand
5'
3'
CATGAAAAAAAAAA·····AAA 3'
GTACTTTTTTTTTT-N10-PrA
Pool, Cleanup, Melt, Terminal Transferase
3' CCC···CCCCCCCCCC
GTACTTTTTTTTTT-N10-PrA
Add Primer, Anneal
5' PrB-GGGGGGGGG
3' CCC·····CCCCCCCCC
GTACTTTTTTTTTT-N10-PrA
Polymerase
2-nd order derivative strand
5' PrB-GGGGGGGGG
3' CCC·····CCCCCCCCC
GTACTTTTTTTTTT-N10-PrA
Add PrA, PrB
PCR
Sequence

Figure 11 illustrates one embodiment of this idea, where the $N_{10}$ designates an oligo-nucleotide tag. The first six nucleotides (4096 possible sequences), randomly generated, provide the varietal tag component, and the last four nucleotides (256 choices) are chosen to encode the micro-well, and, hence, a sample tag. The lengths of the varietal tag and the sample tag can be changed to suit needs, and are not limited to a total of ten nucleotides. The lengths are chosen for illustration. The sample tag allows pooling, with deconvolution occurring afterwards. The combination of the varietal tag with the sequence of the target molecule provides an essentially unique tag for the first derivative strand, that can only with very low probability be replicated by chance from another like molecule (a chance of roughly 1/N where N is the number of available tags). In the end, multiple copies of sequence-identical molecules in the input sample are each individually counted. In applications to mRNA counting, the length of the varietal component can be increased to accommodate possibly thousands of transcripts from the same gene.

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 | 5981/81503-A/GJG/MW |
|---|---|---|
| | Application Number | 13/278,333 |
| Title of Invention | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

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 |
|---|---|---|---|---|
| | James | | Hicks | |

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

| City | Locust Valley | State/Province | NY | Country of Residence | US |
|---|---|---|---|---|---|

**Mailing Address of Inventor:**

| Address 1 | 3 Great Meadow Road | | |
|---|---|---|---|
| Address 2 | | | |
| City | Locust Valley | State/Province | NY |
| Postal Code | 11560 | Country | US |

**Inventor 2** Remove

**Legal Name**

| Prefix | Given Name | Middle Name | Family Name | Suffix |
|---|---|---|---|---|
| | Nicholas | | Navin | |

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

| City | Houston | State/Province | TX | Country of Residence | US |
|---|---|---|---|---|---|

**Mailing Address of Inventor:**

| Address 1 | 1621 Wichita St. | | |
|---|---|---|---|
| Address 2 | | | |
| City | Houston | State/Province | TX |
| Postal Code | 77004 | Country | US |

**Inventor 3** Remove

**Legal Name**

| Prefix | Given Name | Middle Name | Family Name | Suffix |
|---|---|---|---|---|
| | Jennifer | | Troge | |

**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 | 5981/81503-A/GJG/MW |
|---|---|---|
| | Application Number | 13/278,333 |
| Title of Invention | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

| City | Oyster Bay | State/Province | NY | Country of Residence i | US |
|---|---|---|---|---|---|

**Mailing Address of Inventor:**

| Address 1 | 100 Simcoe St. | | |
|---|---|---|---|
| Address 2 | | | |
| City | Oyster Bay | State/Province | NY |
| Postal Code | 11771 | Country i | US |

**Inventor 4** Remove

**Legal Name**

| Prefix | Given Name | Middle Name | Family Name | Suffix |
|---|---|---|---|---|
| | Zihua | | Wang | |

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

| City | Huntington | State/Province | NY | Country of Residence i | US |
|---|---|---|---|---|---|

**Mailing Address of Inventor:**

| Address 1 | 13 Meadowlawn St. | | |
|---|---|---|---|
| Address 2 | | | |
| City | Huntington | State/Province | NY |
| Postal Code | 11743 | Country i | US |

**Inventor 5** Remove

**Legal Name**

| Prefix | Given Name | Middle Name | Family Name | Suffix |
|---|---|---|---|---|
| | Michael | | Wigler | |

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

| City | Cold Spring Harbor | State/Province | NY | Country of Residence i | US |
|---|---|---|---|---|---|

**Mailing Address of Inventor:**

| Address 1 | 52 Snake Hill Road | | |
|---|---|---|---|
| Address 2 | | | |
| City | Cold Spring Harbor | State/Province | NY |
| Postal Code | 11724 | Country i | US |

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

## Correspondence Information:

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

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

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 | 5981/81503-A/GJG/MW |
|---|---|---|
| | Application Number | 13/278,333 |
| Title of Invention | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

| | | | |
|---|---|---|---|
| **Customer Number** | 23432 | | |
| **Email Address** | | Add Email | Remove Email |

## Application Information:

| | | | |
|---|---|---|---|
| **Title of the Invention** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | | |
| **Attorney Docket Number** | 5981/81503-A/GJG/MW | **Small Entity Status Claimed** | ☒ |
| **Application Type** | Nonprovisional | | |
| **Subject Matter** | Utility | | |
| **Total Number of Drawing Sheets (if any)** | 11 | **Suggested Figure for Publication (if any)** | |

## Filing By Reference :

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 | 23432 | | |

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 | 5981/81503-A/GJG/MW |
|---|---|---|
| | Application Number | 13/278,333 |
| Title of Invention | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

## 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 | Expired | | Remove |
|---|---|---|---|
| Application Number | Continuity Type | Prior Application Number | Filing Date (YYYY-MM-DD) |
| | Claims benefit of provisional | 61510579 | 2011-07-22 |
| Prior Application Status | Expired | | Remove |
| Application Number | Continuity Type | Prior Application Number | Filing Date (YYYY-MM-DD) |
| | Claims benefit of provisional | 61406067 | 2010-10-22 |

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

## 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.

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 | 5981/81503-A/GJG/MW |
|---|---|---|
| | Application Number | 13/278,333 |
| Title of Invention | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

## 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

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 o f 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.

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 | 5981/81503-A/GJG/MW |
|---|---|---|
| | Application Number | 13/278,333 |
| Title of Invention | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

**Applicant 1**

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 | Cold Spring Harbor Laboratory | | |
| **Mailing Address Information For Applicant:** | | | |
| **Address 1** | One Bungtown Road | | |
| Address 2 | | | |
| **City** | Cold Spring Harbor | **State/Province** | NY |
| **Country** | US | Postal Code | 11724 |
| Phone Number | | Fax Number | |
| Email Address | | | |

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

## 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.

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

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 | 5981/81503-A/GJG/MW |
|---|---|---|
| | Application Number | 13/278,333 |
| Title of Invention | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

| 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.

## Signature:

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

| **Signature** | /Gary J. Gershik/ | | | Date (YYYY-MM-DD) | 2015-08-04 |
|---|---|---|---|---|---|
| First Name | Gary | Last Name | Gershik | Registration Number | 39992 |

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

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.**

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 13278333 |
| **Filing Date:** | 21-Oct-2011 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | John P. White/Jason Keith |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |

Filed as Small Entity

**Filing Fees for Utility under 35 USC 111(a)**

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| Notice of Appeal | 2401 | 1 | 400 | 400 |
| **Post-Allowance-and-Post-Issuance:** | | | | |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Extension-of-Time:** | | | | |
| Extension - 3 months with $100 paid | 2253 | 1 | 600 | 600 |
| **Miscellaneous:** | | | | |
| | | | **Total in USD ($)** | **1000** |


IPR2025-01353; IPR2025-01355
CSHL EX2002

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 23116131 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Jason Keith |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 04-AUG-2015 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 21:25:01 |
| **Application Type:** | Utility under 35 USC 111(a) |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | Deposit Account |
| Payment was successfully received in RAM | $1000 |
| RAM confirmation Number | 6097 |
| Deposit Account | 033125 |
| Authorized User | GERSHIK, GARY J |

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

Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)



IPR2025-01353; IPR2025-01355
CSHL EX2002

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Extension of Time | 150804_5981_81503_A_Petition_MW.pdf | 65770<br>b4e075b935b9e83d90b844bebe085d867c6e84e8 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Notice of Appeal Filed | 150804_5981_81503_A_Notice_Of_Appeal_MW.pdf | 56654<br>f8e4eebc78d0fa915b6b39808061fe56c59fbe19 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | | 150804_5981_81503_A_Amendment_MW.pdf | 2552466<br>490d0fb16181c48eacc94b09883fe5458e58eb25 | yes | 62 |

**Multipart Description/PDF files in .zip description**

| Document Description | Start | End |
|---|---|---|
| Response After Final Action | 1 | 1 |
| Claims | 2 | 9 |
| Applicant Arguments/Remarks Made in an Amendment | 10 | 10 |
| Applicant summary of interview with examiner | 11 | 14 |
| Applicant Arguments/Remarks Made in an Amendment | 15 | 62 |

| | | | | | |
|---|---|---|---|---|---|
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 4 | Affidavit-Rule 131-pre-AIA (FTI) ONLY | 150804_5981_81503_A_Declaration_MW.pdf | 2253768<br>a836058a9d6c00a8c8a65460d71fb8249b5aa817 | no | 51 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 5 | Application Data Sheet | 150804_5981_81503_A_ADS_MW.pdf | 390270<br>0aacff61f13e415ac458ece9b86d2eef1d8533dc | no | 7 |
| **Warnings:** | | | | | |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Information: | | | | | |
|---|---|---|---|---|---|
| This is not an USPTO supplied ADS fillable form | | | | | |
| 6 | Fee Worksheet (SB06) | fee-info.pdf | 32537<br>8d52bb2d964c119f077df655500dc268e0b75f04 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | | **Total Files Size (in bytes):** | 5351465 | |

**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.**

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| Applicant | : | Cold Spring Harbor Laboratory |
| Inventors | : | James Hicks, Nicholas Navin, Jennifer Troge, Zihua Wang, and Michael Wigler |
| U.S. Serial No. | : | 13/278,333 Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 Art Unit: 1634 |
| Confirmation No. | : | 4469 |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
August 4, 2015

**BY EFS**
Commissioner for Patents
Alexandria, VA 22313-1450

**PETITION FOR A THREE-MONTH EXTENSION OF TIME**

This Petition is submitted in connection with the Final Office Action issued February 4, 2015 by the U.S. Patent and Trademark Office in connection with the above-identified application.

A response to the February 4, 2015 Final Office Action was originally due May 4, 2015. A Petition for a One-Month Extension of Time with the one-month extension of time fee for a small entity of one hundred dollars ($100.00) was filed June 4, 2015.

Applicant hereby petitions for an extension of time of two additional months (i.e. a third month). The fee for a three-month extension of time for a small entity is seven hundred dollars ($700.00). As noted above, Applicant previously paid one hundred dollars ($100.00) for a one-month extension of time. Thus, the fee for an extension of time of a third month is the fee for a three-month extension ($700.00) less the previously paid, i.e. $100.00. The undersigned hereby authorizes the Patent Office to charge the


IPR2025-01353; IPR2025-01355
CSHL EX2002

amount of SIX HUNDRED DOLLARS ($600.00) to Deposit Account No. 03-3125. With a third month extension of time, a response to the February 4, 2015 Final Office Action is due August 4, 2015. Accordingly, the subject application is pending today for the purpose of filing a Notice of Appeal.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicants' undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee, other than the $600.00 fee for an extension of time of two additional months for which the authorization to charge Deposit Account No. 03-3125 has been given above, is deemed necessary in connection with the filing of this Petition. However, if any additional fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39.992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

Minerva Y. Wong 8/4/2015
Minerva Y. Wong Date

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| Applicant | : | Cold Spring Harbor Laboratory |
| Inventors | : | James Hicks, Nicholas Navin, Jennifer Troge, Zihua Wang, and Michael Wigler |
| U.S. Serial No. | : | 13/278,333 Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 Art Unit: 1634 |
| Confirmation No. | : | 4469 |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
August 4, 2015

**BY EFS**
Commissioner for Patents
Alexandria, VA 22313-1450

**NOTICE OF APPEAL FROM THE EXAMINER'S DECISION TO THE PATENT TRIAL AND APPEAL BOARD**

Applicants hereby appeal to the Patent Trial and Appeal Board from the Final Office Action mailed February 4, 2015 finally rejecting claims 1-7, 9, 11-13, 24, 26-28, and 68-71 in the above-identified application.

Applicants are concurrently filing a Petition For A Three-Month Extension of Time. With a three-month extension of time, the subject application is pending as of today, August 4, 2015. Accordingly, this Notice of Appeal is being timely filed.

The required fee for filing a Notice of Appeal under 37 C.F.R. § 41.20(b)(1) for a small entity is FOUR HUNDRED DOLLARS ($400.00) and authorization is hereby given to charge this fee to Deposit Account No. 03-3125.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicant's undersigned attorney invites the Examiner to telephone him at the number provided below.


IPR2025-01353; IPR2025-01355
CSHL EX2002

No fee, other than the $400.00 fee for filing a Notice of Appeal for which the authorization to charge Deposit Account No. 03-3125 has been given above, is deemed necessary in connection with the filing of this Notice of Appeal. However, if any additional fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39.992
Attorney for Applicant(s)
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

Minerva Y. Wong 8/14/2015
Minerva Y. Wong Date


IPR2025-01353; IPR2025-01355
CSHL EX2002

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<br>Substitute for Form PTO-875 | Application or Docket Number<br>13/278,333 | Filing Date<br>10/21/2011 | ☐ To be Mailed |
|---|---|---|---|

**ENTITY**: ☐ LARGE ☒ SMALL ☐ MICRO

## APPLICATION AS FILED – PART I

| | (Column 1) | (Column 2) | | |
|---|---|---|---|---|
| FOR | NUMBER FILED | 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

| | | (Column 1) | | (Column 2) | (Column 3) | | |
|---|---|---|---|---|---|---|---|
| | **08/04/2015** | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | RATE ($) | ADDITIONAL FEE ($) |
| AMENDMENT | Total (37 CFR 1.16(i)) | * 20 | Minus | ** 20 | = 0 | X $40 = | 0 |
| | Independent (37 CFR 1.16(h)) | * 2 | 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** |

| | | (Column 1) | | (Column 2) | (Column 3) | | |
|---|---|---|---|---|---|---|---|
| | | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | RATE ($) | ADDITIONAL FEE ($) |
| AMENDMENT | 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
/DORIS BURNS/

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.*



IPR2025-01353; IPR2025-01355
CSHL EX2002

Doc Code: A.NE.AFCP
Document Description: After Final Consideration Pilot Program Request

PTO/SB/434 (05-13)

## CERTIFICATION AND REQUEST FOR CONSIDERATION UNDER THE AFTER FINAL CONSIDERATION PILOT PROGRAM 2.0

| Practitioner Docket No.: | Application No.: | Filing Date: |
|---|---|---|
| 5981/81503-A/GJG/MW | 13/278,333 | October 21, 2011 |
| **First Named Inventor:** | **Title:** | |
| James Hicks | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

APPLICANT HEREBY CERTIFIES THE FOLLOWING AND REQUESTS CONSIDERATION UNDER THE AFTER FINAL CONSIDERATION PILOT PROGRAM 2.0 (AFCP 2.0) OF THE ACCOMPANYING RESPONSE UNDER 37 CFR 1.116.

1. The above-identified application is (i) an original utility, plant, or design nonprovisional application filed under 35 U.S.C. 111(a) [a continuing application (*e.g.*, a continuation or divisional application) is filed under 35 U.S.C. 111(a) and is eligible under (i)], or (ii) an international application that has entered the national stage in compliance with 35 U.S.C. 371(c).
2. The above-identified application contains an outstanding final rejection.
3. Submitted herewith is a response under 37 CFR 1.116 to the outstanding final rejection. The response includes an amendment to at least one independent claim, and the amendment does not broaden the scope of the independent claim in any aspect.
4. This certification and request for consideration under AFCP 2.0 is the only AFCP 2.0 certification and request filed in response to the outstanding final rejection.
5. Applicant is willing and available to participate in any interview requested by the examiner concerning the present response.
6. This certification and request is being filed electronically using the Office's electronic filing system (EFS-Web).
7. Any fees that would be necessary consistent with current practice concerning responses after final rejection under 37 CFR 1.116, *e.g.*, extension of time fees, are being concurrently filed herewith. [There is no additional fee required to request consideration under AFCP 2.0.]
8. By filing this certification and request, applicant acknowledges the following:
    - Reissue applications and reexamination proceedings are not eligible to participate in AFCP 2.0.
    - The examiner will verify that the AFCP 2.0 submission is compliant, *i.e.*, that the requirements of the program have been met (see items 1 to 7 above). For compliant submissions:
        - The examiner will review the response under 37 CFR 1.116 to determine if additional search and/or consideration (i) is necessitated by the amendment and (ii) could be completed within the time allotted under AFCP 2.0. If additional search and/or consideration is required but cannot be completed within the allotted time, the examiner will process the submission consistent with current practice concerning responses after final rejection under 37 CFR 1.116, *e.g.*, by mailing an advisory action.
        - If the examiner determines that the amendment does not necessitate additional search and/or consideration, or if the examiner determines that additional search and/or consideration is required and could be completed within the allotted time, then the examiner will consider whether the amendment places the application in condition for allowance (after completing the additional search and/or consideration, if required). If the examiner determines that the amendment does not place the application in condition for allowance, then the examiner will contact the applicant and request an interview.
            - The interview will be conducted by the examiner, and if the examiner does not have negotiation authority, a primary examiner and/or supervisory patent examiner will also participate.
            - If the applicant declines the interview, or if the interview cannot be scheduled within ten (10) calendar days from the date that the examiner first contacts the applicant, then the examiner will proceed consistent with current practice concerning responses after final rejection under 37 CFR 1.116.

| Signature | Date |
|---|---|
| Gary J. Gershik | August 4, 2015 |
| **Name (Print/Typed)** Gary J. Gershik | **Practitioner Registration No.** 39,992 |

***Note**: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4(d) for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.*

☐ * Total of _____ forms are submitted.



IPR2025-01353; IPR2025-01355
CSHL EX2002

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 23118044 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | John P. White/Jason Keith |
| **Filer Authorized By:** | John P. White |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 05-AUG-2015 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 10:33:31 |
| **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 | After Final Consideration Program Request | 150804_5981_81503_A_After_Final_MW.pdf | 67324<br>6d05608622d625e64600eef5b049ee0ae11807ef | no | 1 |

| | |
|---|---|
| **Warnings:** | |
| **Information:** | |



IPR2025-01353; IPR2025-01355
CSHL EX2002

| | |
|---|---|
| **Total Files Size (in bytes):** | 67324 |

**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.**

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

# NOTICE OF ALLOWANCE AND FEE(S) DUE

23432 7590 12/02/2015

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER |
|---|
| LU, FRANK WEI MIN |

| ART UNIT | PAPER NUMBER |
|---|---|
| 1634 | |

DATE MAILED: 12/02/2015

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

TITLE OF INVENTION: VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

| APPLN. TYPE | ENTITY STATUS | ISSUE FEE DUE | PUBLICATION FEE DUE | PREV. PAID ISSUE FEE | TOTAL FEE(S) DUE | DATE DUE |
|---|---|---|---|---|---|---|
| nonprovisional | SMALL | $480 | $0 | $0 | $480 | 03/02/2016 |

**THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.**

**THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.**

**HOW TO REPLY TO THIS NOTICE:**

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 1/2 the amount of undiscounted fees, and micro entity fees are 1/2 the amount of small entity fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

**IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.**

# PART B - FEE(S) TRANSMITTAL

**Complete and send this form, together with applicable fee(s), to: Mail** **Mail Stop ISSUE FEE**
**Commissioner for Patents**
**P.O. Box 1450**
**Alexandria, Virginia 22313-1450**
**or Fax (571)-273-2885**

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

23432 7590 12/02/2015
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

**Certificate of Mailing or Transmission**
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

| (Depositor's name) |
|---|
| (Signature) |
| (Date) |

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

TITLE OF INVENTION: VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

| APPLN. TYPE | ENTITY STATUS | ISSUE FEE DUE | PUBLICATION FEE DUE | PREV. PAID ISSUE FEE | TOTAL FEE(S) DUE | DATE DUE |
|---|---|---|---|---|---|---|
| nonprovisional | SMALL | $480 | $0 | $0 | $480 | 03/02/2016 |

| EXAMINER | ART UNIT | CLASS-SUBCLASS |
|---|---|---|
| LU, FRANK WEI MIN | 1634 | 435-006110 |

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 _____
2 _____
3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. The following fee(s) are submitted:
- ❑ Issue Fee
- ❑ Publication Fee (No small entity discount permitted)
- ❑ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ❑ A check is enclosed.
- ❑ Payment by credit card. Form PTO-2038 is attached.
- ❑ The director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ❑ Applicant certifying micro entity status. See 37 CFR 1.29 — NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
- ❑ Applicant asserting small entity status. See 37 CFR 1.27 — NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
- ❑ Applicant changing to regular undiscounted fee status. — NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____



PTOL-85 Part B (10-13) Approved for use through 10/31/2013. OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 12/02/2015

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER | |
|---|---|
| LU, FRANK WEI MIN | |
| **ART UNIT** | **PAPER NUMBER** |
| 1634 | |

DATE MAILED: 12/02/2015

**Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)**
(Applications filed on or after May 29, 2000)

The Office has discontinued providing a Patent Term Adjustment (PTA) calculation with the Notice of Allowance.

Section 1(h)(2) of the AIA Technical Corrections Act amended 35 U.S.C. 154(b)(3)(B)(i) to eliminate the requirement that the Office provide a patent term adjustment determination with the notice of allowance. See Revisions to Patent Term Adjustment, 78 Fed. Reg. 19416, 19417 (Apr. 1, 2013). Therefore, the Office is no longer providing an initial patent term adjustment determination with the notice of allowance. The Office will continue to provide a patent term adjustment determination with the Issue Notification Letter that is mailed to applicant approximately three weeks prior to the issue date of the patent, and will include the patent term adjustment on the patent. Any request for reconsideration of the patent term adjustment determination (or reinstatement of patent term adjustment) should follow the process outlined in 37 CFR 1.705.

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

## OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. 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, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. 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.

## 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 disclosure of these records is required by the Freedom of Information Act.
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 inspection 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.


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Notice of Allowability | Application No. 13/278,333 | Applicant(s) HICKS ET AL. | |
|---|---|---|---|
| | Examiner FRANK LU | Art Unit 1634 | AIA (First Inventor to File) Status No |

***-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--***

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *the amendment filed on 8/4/2015*.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on_____.
2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.
3. ☒ The allowed claim(s) is/are *1-7,9,11-13,24,26-28,31 and 68-71*. As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov .
4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   **Certified copies:**

   a) ☐ All b) ☐ Some *c) ☐ None of the:
   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
**THIS THREE-MONTH PERIOD IS NOT EXTENDABLE**.

5. ☐ CORRECTED DRAWINGS ( as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   **Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).**
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

**Attachment(s)**

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
4. ☒ Interview Summary (PTO-413), Paper No./Mail Date *10/7/2015* .
5. ☒ Examiner's Amendment/Comment
6. ☒ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

| /FRANK LU/<br>Primary Examiner, Art Unit 1634 | October 27, 2015 |
|---|---|

***Reasons for Allowance***

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with Mr. Gary GerShik (Reg. No. 39,992) on October 7, 2015.

2. The application has been amended as follows:

1. (Currently Amended) A method for obtaining [from genomic nucleic acids] genomic copy number information from genomic nucleic acids, comprising:

a) obtaining segments of the genomic nucleic acids;

b) tagging the segments with substantially unique tags [to generate], thereby generating tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic nucleic acids from step (a) and a substantially unique tag having a unique nucleotide sequence of the substantially unique tags, so as to permit one substantially unique tag having a unique nucleotide sequence of the substantially unique tags to differentiate [differentiation] from another tag having a unique nucleotide sequence of the substantially unique tags [and which tag is part of another tagged nucleic acid molecule];

c) subjecting the tagged nucleic acid molecules to [amplification by] a polymerase chain reaction (PCR), thereby generating amplified tagged nucleic acid molecules;

d) generating tag associated sequence reads by sequencing the [product] amplified tagged nucleic acid molecules of step (c);

e) assigning each of the amplified tagged nucleic acid molecules to a location on a genome associated with the genomic nucleic acids by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (d) corresponding to one segment of the segments of the genomic nucleic acids to a location on the genome; and

f) in the amplified tagged nucleic acid molecules, counting the number of the amplified tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome, thereby obtaining counts of the amplified tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location of the genome and the genomic copy number information from the genomic nucleic acids, wherein the different substantially unique tag is from the substantially unique tags, and wherein the genome has multiple locations.

2. (Currently Amended) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising multiple locations [more than one location on the genome, by counting the number of the tagged nucleic acid molecules having a different tag that have been assigned to the locations on the genome in step (f), and assigning as the copy number of the region the count obtained in step (f) for the locations within the region].

3. (Currently Amended) The method of claim 1, further comprising comparing a count from the counts obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Currently Amended) The method of claim 1, further comprising

[a] g) summing the counts from the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

[b] h) summing the counts from the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region [is comprised of] comprises a number of locations which is comparable to the number of locations of the first region;

[c] i) comparing the value obtained in step [(a)] g) to the value obtained in step [(b)] h), thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Currently Amended) The method of claim 4, wherein step [(b)] h) further comprises

[i] j) summing the counts from the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region [is comprised of] comprises a number of locations which is comparable to the number of locations of the first region; and

[ii] k) obtaining an average of the sum of the counts obtained in step h) and the counts obtained in step j) [(f) for locations which comprise the second region and the sum of the counts obtained in step (f) for locations which comprise the third region].

7. (Currently Amended) The method of claim 1, further comprising [i] l) summing the counts from the counts obtained in step (f) for locations which comprise a region of the genome [and ii] m) summing the counts obtained from a reference sample for the same region of the


IPR2025-01353; IPR2025-01355
CSHL EX2002

genome[[,]]; and iii) comparing the sums of the counts obtained in step l) and step m) [(i) and (ii)] , thereby estimating a relative genomic copy number of the region of the genome.

9. (Currently Amended) A method for obtaining [from mRNA transcripts] mRNA copy number information [unaffected by amplification distortion] from mRNA transcripts, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA transcripts under conditions that promote the formation of only one complement of the mRNA transcripts, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to an end of the first order derivative strands, thereby generating first order derivative strands comprising the polynucleotide tail; and

iii) subjecting the first order derivative strands comprising the polynucleotide tail to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement of the first order derivative strands comprising the polynucleotide tail, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique tags [primers, thereby generating] and the first order derivative strands and the second order derivative strands are the tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to a polymerase chain reaction (PCR) [amplification], thereby generating amplified tagged nucleic acid molecules;

c) generating tag associated sequence reads by sequencing the amplified tagged nucleic acid molecules[product] of step (b);

d) assigning each of the amplified tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (c) corresponding to a mRNA transcript of the mRNA transcripts to a location on the cDNA library; and

e) in the amplified tagged nucleic acid molecules, counting the number of the amplified tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library, thereby obtaining counts of the amplified tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library of the cDNA library and mRNA copy number information from the mRNA transcripts, wherein the different tag is from the substantially unique tags, and the cDNA library comprises multiple locations.

11. (Currently Amended) The method of claim 1, wherein said tagging the segments [to generate tagged nucleic acid molecules] comprises:

[a] n) adding a polynucleotide tail to the ends of the segments of the genomic nucleic acids [to generate], thereby generating zero-th order derivative strands comprising the polynucleotide tail;

[b] o) subjecting the zero-th order derivative strands of step [(a)] n) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement of the zero-th order derivative strands, thereby generating first order derivative strands;

[c] p) adding a polynucleotide tail to the first order derivative strands, thereby generating a first order derivative strands comprising the polynucleotide tail;

[d] q) subjecting the first order derivative strands comprising the polynucleotide tail to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement of the first order derivative strands comprising the polynucleotide tail, thereby generating second order derivative strands,

wherein the primers of at least one of steps [(b)] o) and [(d)] q) comprise the substantially unique tags [primers, thereby generating] and the first order derivative strands and the second order derivative strands are the tagged nucleic acid molecules.

12. (Currently Amended) The method of claim 1 [step (b)], wherein said tagging the segments [to generate the tagged nucleic acid molecules] comprises:

[a] r) adding a polynucleotide tail to the ends of the segments of genomic nucleic acids [to generate], thereby generating zero-th order derivative strands comprising the polynucleotide tail,

[b] s) subjecting the zero-th order derivative strands of step [(a)] r) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step [(a)] r) under conditions that promote the ligation of a primer to the 5' ends of the zero-th order derivative strands,

[c] t) subjecting the product of step [(b)] s) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step [(a)] r) under conditions which promote the formation of only one complement of the product of step s), wherein the primers of


IPR2025-01353; IPR2025-01355
CSHL EX2002

step s) and the primers of step t) have different nucleotide sequences [than the primers of step (b)], and wherein the polymerase has 3'-5' proofreading activity,

wherein the primers of at least one of steps [(b) and (c)] s) and t) comprise substantially unique [primers] tags, thereby generating tagged nucleic acid molecules.

13. (Currently Amended) The method of claim 12, wherein said adding the polynucleotide tail comprises the use of a terminal transferase.

24. (Currently Amended) The method of claim 1, wherein [i)] the tagged nucleic acid molecules are subject to hybrid capture prior to the PCR or [ii)] the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to step (d).

27. (Currently Amended) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to the PCR [amplification] or prior to said sequencing of the product of step (c).

28. (Currently Amended) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads [according to the sample tag].

68. (Currently Amended) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to the PCR in step(c).

70. (Currently Amended) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to said sequencing the amplified tagged nucleic acid molecules in step d).

3. The following is an examiner's statement of reasons for allowance:

Claims 1-7, 9, 11-13, 24, 26-28, 31, and 68-71 are allowable in light of applicant's amendments and the declaration under 37 C.F.R. 1.131 filed on August 4, 2015, and the examiner's amendment. The rejections under 35 U.S.C 112 (a) and 112 (b) have been withdrawn in view of applicant's amendments filed on August 4, 2015 and the examiner's amendment. No prior art in the record teaches a combination of steps a) to f) of claim 1 and a combination of steps a) to e) of claim 9. No prior art either alone or in combination with the other art in the record teaches or reasonably suggests a method for obtaining genomic copy number information and a method for obtaining mRNA copy number information which comprise all limitations recited in claims 1 and 9.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance".

4. Papers related to this application may be submitted to Group 1600 by facsimile transmission. Papers should be faxed to Group 1600 via the PTO Fax Center. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993)(See 37 CAR § 1.6(d)). The CM Fax Center number is (571)273-8300.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Frank Lu, Ph.D., whose telephone number is (571)272-0746. The examiner can normally be reached on Monday-Friday from 9 A.M. to 5 P.M.


IPR2025-01353; IPR2025-01355
CSHL EX2002

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dave Nguyen, can be reached on (571)272-0731.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Frank W Lu /
Primary Examiner, Art Unit 1634
October 27, 2015


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Examiner-Initiated Interview Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 13/278,333 | HICKS ET AL. | |
| | Examiner | Art Unit | |
| | FRANK LU | 1634 | |

All participants (applicant, applicant's representative, PTO personnel):

(1) *FRANK LU*. (3)_____.

(2) *Mr.Gary GerShik (Reg. No. 39,992)*. (4)_____.

Date of Interview: *07 October 2015*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☒112 ☐102 ☐103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *1-5, 7, 9, 11-13, 24, 27, 28, 68, and 70*.

Identification of prior art discussed: _____.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

*Mr. GerShik agreed with the examiner's amendment*.

**Applicant recordation instructions**: It is not necessary for applicant to provide a separate record of the substance of interview.

**Examiner recordation instructions**: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

| /FRANK LU/<br>Primary Examiner, Art Unit 1634 | October 7, 2015 |
|---|---|


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Search Notes | Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|---|
| | 13278333 | HICKS ET AL. |
| | Examiner | Art Unit |
| | FRANK LU | 1634 |

| CPC- SEARCHED | | |
|---|---|---|
| Symbol | Date | Examiner |
| Update | 10/27/2015 | /FL/ |

| CPC COMBINATION SETS - SEARCHED | | |
|---|---|---|
| Symbol | Date | Examiner |
| Update | 10/27/2015 | /FL/ |

| US CLASSIFICATION SEARCHED | | | |
|---|---|---|---|
| Class | Subclass | Date | Examiner |
| Update | | 10/27/2015 | /FL/ |

| SEARCH NOTES | | |
|---|---|---|
| Search Notes | Date | Examiner |
| Update | 10/27/2015 | /FL/ |

| INTERFERENCE SEARCH | | | |
|---|---|---|---|
| US Class/ CPC Symbol | US Subclass / CPC Group | Date | Examiner |
| 435 | 6.1, 6.11, 6.12, 91.1, 91.2, 91.51 | 10/27/2015 | /FL/ |
| 536 | 23.1, 24.3, 24.33, 25.3 | 10/27/2015 | /FL/ |

| | |
|---|---|
| | /FRANK LU/<br>Primary Examiner.Art Unit 1634 |


IPR2025-01353; IPR2025-01355
CSHL-EX2002

| Issue Classification | Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|---|
| | 13278333 | HICKS ET AL. |
| | **Examiner** | **Art Unit** |
| | FRANK LU | 1634 |

| CPC | | | | | | |
|---|---|---|---|---|---|---|
| **Symbol** | | | | | **Type** | **Version** |
| C12Q | | 1 | / | 6881 | F | 2013-01-01 |
| C12Q | | 1 | / | 6886 | I | 2013-01-01 |
| C12Q | | 1 | / | 683 | I | 2013-01-01 |
| C12Q | | 2600 | / | 156 | A | 2013-01-01 |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |
| | | | / | | | |

| CPC Combination Sets | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Symbol** | | | | | **Type** | **Set** | **Ranking** | **Version** |
| C12Q | | 1 | / | 683 | I | 1 | 1 | 2013-01-01 |
| C12Q | | 2525 | / | 155 | A | 1 | 2 | 2013-01-01 |
| C12Q | | 2525 | / | 191 | A | 1 | 3 | 2013-01-01 |
| C12Q | | 2535 | / | 122 | A | 1 | 4 | 2013-01-01 |
| C12Q | | 2537 | / | 159 | A | 1 | 5 | 2013-01-01 |
| C12Q | | 2537 | / | 16 | A | 1 | 6 | 2013-01-01 |

| NONE | | **Total Claims Allowed:** | |
|---|---|---|---|
| (Assistant Examiner) | (Date) | 20 | |
| /FRANK LU/<br>Primary Examiner.Art Unit 1634 | 10/27/2015 | O.G. Print Claim(s) | O.G. Print Figure |
| (Primary Examiner) | (Date) | 1 | None |



IPR2025-01353; IPR2025-01355
CSHL EX2002

| Issue Classification | Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|---|
| | 13278333 | HICKS ET AL. |
| | **Examiner** | **Art Unit** |
| | FRANK LU | 1634 |

| US ORIGINAL CLASSIFICATION | |
|---|---|
| **CLASS** | **SUBCLASS** |
| 435 | 91.2 |

**CROSS REFERENCE(S)**

| CLASS | SUBCLASS (ONE SUBCLASS PER BLOCK) | | | | |
|---|---|---|---|---|---|
| 435 | 6.1 | 6.11 | 6.12 | 91.1 | 91.51 |
| 536 | 23.1 | 24.3 | 24.33 | 25.3 | |

| INTERNATIONAL CLASSIFICATION | | | | | |
|---|---|---|---|---|---|
| **CLAIMED** | | | | | **NON-CLAIMED** |
| C | 1 | 2 | P | 19 / 34 (2006.01.01) | |
| C | 1 | 2 | Q | 1 / 68 (2006.01.01) | |
| C | 0 | 7 | H | 21 / 02 (2006.01.01) | |
| C | 0 | 7 | H | 21 / 04 (2006.01.01) | |
| C | 0 | 7 | H | 21 / 00 (2006.01.01) | |

| NONE | | **Total Claims Allowed:** | |
|---|---|---|---|
| (Assistant Examiner) | (Date) | 20 | |
| /FRANK LU/<br>Primary Examiner.Art Unit 1634 | 10/27/2015 | O.G. Print Claim(s) | O.G. Print Figure |
| (Primary Examiner) | (Date) | 1 | None |

| Issue Classification | Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|---|
| | 13278333 | HICKS ET AL. |
| | **Examiner** | **Art Unit** |
| | FRANK LU | 1634 |

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

| Final | Original | Final | Original | Final | Original | Final | Original | Final | Original | Final | Original | Final | Original | Final | Original |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 17 | 27 | | | | | | | | | | | | |
| 2 | 2 | 18 | 28 | | | | | | | | | | | | |
| 3 | 3 | 19 | 31 | | | | | | | | | | | | |
| 4 | 4 | 20 | 9 | | | | | | | | | | | | |
| 5 | 5 | | | | | | | | | | | | | | |
| 6 | 6 | | | | | | | | | | | | | | |
| 7 | 7 | | | | | | | | | | | | | | |
| 8 | 11 | | | | | | | | | | | | | | |
| 9 | 12 | | | | | | | | | | | | | | |
| 10 | 13 | | | | | | | | | | | | | | |
| 11 | 24 | | | | | | | | | | | | | | |
| 12 | 68 | | | | | | | | | | | | | | |
| 13 | 69 | | | | | | | | | | | | | | |
| 14 | 70 | | | | | | | | | | | | | | |
| 15 | 71 | | | | | | | | | | | | | | |
| 16 | 26 | | | | | | | | | | | | | | |

| | | | |
|---|---|---|---|
| NONE<br><br>(Assistant Examiner) | (Date) | **Total Claims Allowed:**<br>20 | |
| /FRANK LU/<br>Primary Examiner.Art Unit 1634<br><br>(Primary Examiner) | 10/27/2015<br><br>(Date) | O.G. Print Claim(s)<br>1 | O.G. Print Figure<br>None |



IPR2025-01353; IPR2025-01355
CSHL EX2002

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office<br>INFORMATION DISCLOSURE STATEMENT<br>(Use several sheets if necessary) | | |
|---|---|---|
| | Application Number | 13/278,333 |
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Ricks et al. |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/MW |

**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 |
|---|---|---|---|---|
| | 1 | 2004-110153 | 06-10-2004 | Dong et al. |
| | 2 | 2007-172873 | 07-26-2007 | Brenner et al. |
| | 3 | 2010-178655 | 07-15-2010 | Hamilton et al. |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | 4 | WO 2007/062445 | 06-07-2007 | Alexander Seitz | |

**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] |
|---|---|---|---|
| | 5 | February 13, 2014 Extended European Search Report, issued in connection with European Patent Application No. 11835243.4 | |
| | 6 | Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 | |

**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 |
|---|---|---|---|---|
| | 7 | 5,639,603 | 06-17-1997 | Dower et al. |
| | 8 | 7,537,897 | 05-26-2009 | Brenner et al. |
| | 9 | 7,622,281 | 11-24-2009 | Ronaghi et al. |
| | 10 | 2006/0073506 | 04-06-2006 | Christians et al. |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Exhibit D

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office<br><br>INFORMATION DISCLOSURE STATEMENT<br>(Use several sheets if necessary) | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks et al. |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/MW |

**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] |
|---|---|---|---|
| | 11 | Casbon J.A., Osborne R.J., Brenner S., Lichtenstein C.P. A method for counting PCT template molecules with application to next-generation sequencing. Nucleic Acids Research Advance Access, published April 13, 2011 | |
| | 12 | Cheung V.G., Nelson S.F. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. PNAS, vol. 93:14676-14679, 1996 | |
| | 13 | Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," Science, vol. 323 pp. 133-138 (2009) | |
| | 14 | Supporting Online Material for Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," accessible at www.sciencemag.org/cgi/content/full/1162986/DC1, published November 20, 2008 | |
| | 15 | Fu G.K., Hu J., Wang P., Fodor S.P.A. Counting individual DNA molecules by the stochastic attachment of diverse labels. PNAS, vol. 108:9026-9031, 2011 | |
| | 16 | Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S.. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135 | |
| | 17 | McCloskey M.L., Stoger, R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767 | |
| | 18 | Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130 | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office<br><br>INFORMATION DISCLOSURE STATEMENT<br>(Use several sheets if necessary) | | |
|---|---|---|
| | Application Number | 13/278,333 |
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks et al. |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/MW |

**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] |
|---|---|---|---|
| | 19 | Paunio T., Reima I., Syvanen A. Preimplanation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. Clinical Chemistry 42:9:1382-1390 (1996) | |
| | 20 | Technology Backgrounder: Single Molecule Real Time (SMRT™) DNA Sequencing. Pacific Biosciences 2009 | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 June 2007 (07.06.2007)

PCT

(10) International Publication Number
WO 2007/062445 A1

(51) International Patent Classification:
*C12Q 1/68* (2006.01)

(21) International Application Number:
PCT/AT2006/000494

(22) International Filing Date:
29 November 2006 (29.11.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
A 1923/2005 29 November 2005 (29.11.2005) AT

(71) Applicant and
(72) Inventor: **SEITZ, Alexander** [AT/AT]; St.-Veit-Gasse 18/2, A-1130 Vienna (AT).

(74) Agent: **SONN & PARTNER PATENTANWÄLTE**; Riemergasse 14, A-1010 Vienna (AT).

(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, 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).

**Declaration under Rule 4.17:**
— *of inventorship (Rule 4.17(iv))*

**Published:**
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: POLYNUCLEOTIDE AMPLIFICATION

(57) Abstract: The present invention provides a method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of a) obtaining a sample or RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining a sample of full length cDNA, b) tailing the 3' end of the transcribed cDNA with a polynucleotide tail after the 3' end, c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (51) primer is specific for the a upstream portion of the polynucleotide tail and the next 1 to 10 nucleotides upstream of the 3' polynucleotide tail of the cDNA.

WO 2007/062445 A1



Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,353
Filed: October 21, 2011
SIDS Exhibit 1
IPR2025-01353, IPR2025-01355
CSHL EX2002

# Polynucleotide amplification

The present invention relates to the field of specific polynucleotide amplification.

To understand the causalities of the different states life presents itself, one needs to find the factors and processes that lead to it. One central element in this quest is the genome where information is stored in such a fashion that it can be propagated through the generations in a highly conserved manner. In executing the information contained in the genome, DNA is transcribed into RNA and further translated into protein. Studying the genome (DNA), transcriptome (RNA) and proteome (protein) has in recent years greatly contributed to our understanding of the molecular basis of life.

An array of possibilities exists to compare RNA samples in vitro. An early approach is the method of differential hybridization or differential screening (Rebagliati et al., 1985). It involves picking at random clones (or sequences) from cDNA libraries and checking for their presence in different concentration in samples of RNA. An enormous effort has to be put into identifying a single differentially expressed transcript as their population is thought to be below 1%. Subtractive hybridization has been proposed to overcome this problem by enriching for differentially expressed transcripts prior to screening. Another way of addressing the problem of abundance was made by introducing high-density micro arrays. In using micro array technology 100s to 1000s of cDNAs, corresponding to RNAs of genes of interest, are spotted and bound onto a surface. They are normally derived from sequenced EST libraries and - if non redundant - can cover as many transcripts as spots fit onto the membrane. In short, RNA extracted from cells or tissues of interest are labeled, e.g. during reverse transcription, and hybridized to these microarrays. The signal intensity of corresponding spots between two or more RNA preparations can then be compared. Even though this technology has produced massive amounts of data, there are two fundamental drawbacks of all these hybridization techniques. One is cross hybridization. This means that two DNA molecules do not need to be complementary over their entire sequence to hybridize. Therefore this introduces ambiguity towards the identity of the molecule that hybridizes (binds) to



IPR2025-01353; IPR2025-01355
CSHL EX2002

the spot. Secondly one can only compare molecules that have been spotted, limiting the search to sequences that are present in the library used to generate the array.

An alternative approach was introduced in 1992 (Liang et al., 1992; Welsh et al., 1992) that today is generally known as "differential display" (DD). In DD the total mRNAs population of a sample is first divided into defined pools of cDNAs during reverse transcription (RT). This is accomplished by priming the RT with one, two or x base anchored oligo dT primers, subdividing the samples into 3, 12 or $3x4^{x-1}$ fractions. These anchor bases hybridize to the bases immediately upstream of the poly A tail. Therefore selecting for mRNA molecules to be reverse transcribed whose bases immediately upstream of the poly A tail are complementary to the anchor bases of the primer. In a subsequent Polymerase chain reaction (PCR) arbitrary primers and anchored primers are used to amplify defined pools of cDNA. Corresponding pools are run side by side on a gel that separates them by size. Differences of expression levels are reflected in differences in band intensity. Bands of interest are cut out of the gel, reamplified and sequenced. One drawback of the original DD is that it uses arbitrary primers for amplification. Such a shotgun approach can reach full transcriptome coverage only statistically, by covering the transcriptome several times. Therefore some RNA molecules will necessarily be presented several times in the display and others might not be present at all.

Recently methods have been introduced that amplify pools of RNA in a more systematic fashion (Prashar et al., 1996; Prashar et al., 1999). In principle, three additional steps are involved. After first strand cDNA synthesis second strand synthesis is carried out to get double stranded DNA that can be cut with a restriction enzyme. In a third step an adaptor is ligated to the 5' ends of the fragments. These fragments are then PCR amplified with one primer being complementary to the 5' adaptor sequence and an anchor primer that was used for RT. It was estimated by the authors that each 6-base cutting restriction enzyme cuts 8% of the cDNAs at a position between 50 and 400 bases from the 3' poly A tail of the mRNA. Therefore more than 12 restriction enzymes will be necessary to approach a complete coverage of the transcriptome. A somewhat more elegant variety of this method was described by Matz et al. (1997)with the intro-

duction of "ordered differential display" that takes advantage of the PCR suppression effect.

However, besides covering the transcriptome only statistically at one level or another, all afore mentioned DD methods suffer from another drawback. They will give the investigator only part of the mRNA sequence that is differentially expressed. Downstream of these DD technologies one will still have to either screen cDNA libraries or where available online sequence libraries for full length clones or employ methodologies that can discover 5' and 3' sequences from a partially known mRNA sequence such as RACE (Frohman et al., 1988). Even identifying such full length transcripts will still not yield certainty at least in eukaryotes that the mRNA found is the mRNA that generated the signal as in eukaryotes a major contributor to genome complexity is splicing. Therefore the sequence found on the display can be common to several splice variants. Furthermore different splice variants of a gene may have different functions, making it highly desirable to know the correct splice variant that generated the signal.

A goal of the present invention is to provide a method that can amplify (enrich for) defined subpopulations of transcripts in their entire length with full coverage of all mRNA molecules present in the sample.

The present invention now provides a method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of:

a) obtaining (or providing) a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining (or providing) a sample of full length cDNA,

b) tailing the 3' end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 10 nucleotides (i.e. the next 1 or 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides), preferably 1 to 5 nucleotides, even more preferred 3 to 5 nucleotides, upstream of the 3' polynucleotide tail of the cDNA. Of course the second primer

can also comprise more specific nucleotides further upstream of the 1 to 10 nucleotide region upstream of the tail specific region (complement).

A full length cDNA can be created by a reverse transcription process described below, or by methods known in the state of the art. Methods to generate full length cDNA are abundantly known to the skilled man in the art, e.g. disclosed in patents US 5,962,271 and US 5,962,272. Such generated cDNA can also be used as starting material to pool full-length cDNA into different fractions. These methods use the template switching ability of reverse transcriptase to add a defined sequence to the 5' end of the cDNA in a cap dependent manner. The primer that is needed in the subsequent amplification will therefore have a portion that will be complementary to the so generated 3' end of the cDNA and will define what fraction of cDNA will be amplified through several, e.g. 1 to 10, additional bases defined by the 5' end of the mRNA molecules amplified.

Therefore it is preferred that a tail is added to the cDNA by the template switching ability of the reverse transcriptase, preferably in a cap dependent manner and the second amplification primer (5' primer) comprises complementary nucleotides to the so added tail and oligo-dCs and 1 to 10 nucleotides upstream of the oligo-dC sequence.

As used herein, the term "reverse transcriptase" relates to any polymerase that has reverse transcriptase activity and can be used to synthesize cDNA from RNA.

As used herein, the term "full length cDNA" is defined as DNA that includes a sequence complementary to the RNA sequence from the first base to the last base of the RNA. In case of RNA molecules that have a cap and/or a polyA tail as is the case for e.g. most eukaryotic mRNA, "full length cDNA" is defined as DNA that includes a sequence complementary to the RNA sequence from the first base after the cap, e.g. the RNA 7-methylguanosine cap or RNA m7G cap, to the last base before the poly A tail of the RNA used as a template.

As used herein, the term "full length amplification product" is defined as DNA that includes a sequence complementary to the RNA sequence from the first base to the last base of the RNA. In case of RNA molecules that have a cap and/or a polyA tail as is the case for e.g. most eukaryotic mRNA, "full length amplifica-

tion product" is defined as amplification product that includes a sequence complementary to the RNA sequence from the first base after the cap to the last base before the poly (A) tail of the RNA used as a template.

A central element of this invention is the fact that full length RNA molecules can be represented in subpopulations that are defined on their 5' and/or 3' end. In principle that means that after the generation of full length cDNA by any means known in the art, the 5' and/or 3' sequence can be used to formulate a set of primer combinations (primer matrix) that has the following quality.

1) Each cDNA will be amplified by only one primer pair in the absence of miss priming. Therefore full transcriptome coverage can be reached without redundancy.
2) Each cDNA sequence within such a defined pool will encode the full amino acid sequence of the protein translated from the corresponding RNA and also
3) Define its transcription start site on its gene (there could be different RNA molecules transcribed from the same gene that can have different transcription start sites).

In comparing quality (1) to original DD one can state that therein the arbitrary primers have the property of either not being complementary to a specific cDNA, being complementary once or being complementary more than once. Therefore each cDNA can be represented either not at all, or once or more than once. When using alternative methodologies that include a restriction enzyme, each restriction enzyme will cut a specific cDNA either not at all or once or more than once. Therefore truly complete transcriptome coverage can never be reached. And increasing the probability for mRNAs to be represented at least one time in the display will mean to increase redundancy.

Qualities (2) and (3) can also not be reached by differential display techniques used to date without employing downstream technologies such as RACE.

The present invention provides methods for amplifying and analyzing the transcriptome through qualitative and quantitative detection of differentially expressed RNA molecules. The technique according to the invention is an organized full length expression display.

In a preferred embodiment the method according to the inven-

tion comprises a primer, which is specific for the (preferably 3' poly A) tail of the RNA or mRNA and the next 1 to 10 nucleotides, preferably 1 to 5 nucleotides, even more preferred 3 to 5 nucleotides, upstream of the (poly A) tail, that is used for the reverse transcription and/or one amplification primer is specific for a corresponding (5' poly-T) stretch of the cDNA, which is complementary to the (3' poly A) tail of the RNA or mRNA, and the next 1 to 10 nucleotides, preferably 1 to 5 nucleotides, even more preferred 3 to 5 nucleotides, downstream of the corresponding (5' poly-T) stretch of the cDNA. Accordingly in addition to the selective 3' primer described above a selective 5' primer can be used to increase the selective power of the primer pair and thus increase the number of cDNA pools defined by the specific nucleotides next to the end sequence anchors.

Preferably the primer for reverse transcription consists of deoxynucleotides.

Generally, the primers throughout this specification can be of any nucleotides, such as deoxyribonucleotides or ribonucleotides. Although deoxynucleotides are preferred in most cases, ribonucleotide primers are also feasible. Therefore a deoxyribonucleotide in a given primer, such as given in examples below, may be substituted by one of its ribonucleotide analogue (e.g. dG by G, dA by A, dC by C, dT by T or U) or other nucleotide analogues.

Preferably the tailing of the 3' end is performed using terminal transferase. - Although other tailing methods are also disclosed, like ligation of a tail sequence, which can be e.g. a defined arbitrary sequence. The terminal transferase can add a certain number of nucleotides preferably uniformly selected from one nucleotide type. Any other means for tailing, adding a tail sequence can also be used, e.g. employing the template switching ability of the reverse transcriptase or by ligating the tail sequence which can be uniformly of one type of nucleotides or of varying nucleotides. Such a tail is preferably a sequence in the range between 5 and 500 nucleotides, more preferred less than 400, less than 300, less than 200, less than 100, less than 50 or less than 30 nucleotides.

In a preferred embodiment the amplification of the cDNA sequences using a selective primer pair is performed by PCR. PCR or polymerase chain reaction is a well established technique

known in the state of the art to amplify polynucleotides using a polymerase, nucleotidetriphosphates, primers, buffer substances and cofactors such as Mg ions and a temperature protocol for hybridization, polymerisation and melting or strand separation.

Preferably the first primer and/or second primer for the amplification consists of deoxynucleotides.

In a preferred embodiment the method according to the invention comprises an additional step of separation of the amplification products according to their length. The separation of the created polynucleotides renders a characterization, isolation and sequence determination possible for each pool after the selective polymerisation which allows a significant advantage for this preferred embodiment in practice.

Preferably the separation is performed by gel electrophoresis, preferably agarose gel electrophoresis, polyacrylamide gel electrophoresis, or capillary electrophoresis.

Even more preferred the method according to the invention comprises an additional step of determining the identity or sequence of the amplification products. Sequence determination can be done for example by gel sequencing, e.g. PCR with dideoxynucleotides, or in an automated sequencer.

Preferably the step of determining the identity or sequence of the amplification products is performed by an automated process on a chip.

The sample used for reverse transcription preferably contains total RNA or mRNA, preferably purified RNA or mRNA, from a specimen. Total RNA includes, but is not limited to, protein coding RNA also called coding RNA such as messenger RNA (mRNA) and non protein coding RNA (non coding RNA or ncRNA), such as ribosomal RNA (rRNA), transfer RNA (tRNA), micro RNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA). Every one of these classes, alone or together with one or more classes, can be analysed according to methods that are within the scope of the invention.

These RNA molecules are preferably purified by affinity chromatography using e.g. an oligo-dT-cellulose column. Thus mRNA with a poly A tail can be preselected.

It is preferred in the inventive method that reverse transcription is carried out at high stringency conditions to in-

crease selectivity of the anchored primers, which can be used as a pool of anchored primers in one polymerisation reaction, as well as separate primers for different polymerisation reactions. In the present invention the phrases "stringent hybridization conditions" or "stringent conditions" refer to conditions under which a primer of the invention will hybridize to its target sequence, but to a minimal number of other sequences. High stringency can be achieved for example by using a high annealing temperature alone or in combination with a chemical additive such as dimethylsulfoxide (DMSO), formamide, tetramethylammonium chloride (TMAC) and Tetramethylammonium oxalate (TMAO) (Kovarova et al. 2000). Examples for high stringency hybridization conditions are given in Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, N.Y. 1989) and Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley and Sons, N.Y. 1994. Under standard PCR conditions often an annealing temperature is chosen that is about 5 degree Celsius below the melting temperature of the primer. Keeping reaction conditions constant the melting temperature in turn is mainly dependent on the length and the GC-content of the oligonucleotides, which can be calculated by common means known in the art. Under more stringent conditions the annealing temperature is raised above the temperature used in standard conditions to decrease the likelihood of mispriming. The annealing temperature can be raised to a temperature that will still allow enough priming events for the amplification to occur. For instance, if a primer has a melting temperature of 50°C, and an annealing temperature of 45°C would be chosen under standard conditions, an annealing temperature of 50°C would thus be more stringent then 45°C. For example primers of the type $dT_{16}VN$ can lead to strong mispriming under normal RT conditions and therefore decrease the specificity of the cDNA pools synthesized (Mizuno et al., 1999). Analogous results have been reported for AMV and HIV-1 reverse transcriptase. Primer promiscuity can be decreased by increasing temperature during annealing and reverse transcription. For this purpose thermostable enzymes can be used that exhibit reverse transcriptase activity or reverse transcriptase enzymes can be thermo protected by adding thermo protectants such as trehalose to the reaction (Carninci et al., 1998). An added benefit of using trehalose is that the

amount of full length cDNA increases (Carninci et al., 1998). This property was also described by Spiess et al. (2002). In addition Spiess showed that betaine can also increase the amount of full length cDNA and that a combination of betaine and trehalose gave the best result (Spiess et al., 2002). Such RT can be derived from moloney murine leukemia virus (M-MLV) (US 4,943,531 A) or in a more preferred embodiment M-MLV lacking RNAse H activity (US 5,405,776 A). Yet another possibility to raise primer specificity is to include competitive oligo nucleotide blockers as described by Mizuno et al. (1999). Therefore the present invention also comprises a method wherein high stringency conditions are used for the processes of reverse transcription and/or the amplification.

Preferably, in the method according to the invention the selectivity of the reverse transcription and/or amplification reaction, in particular the polymerase chain reaction, is increased by the utilization of trehalose, betaine, tetramethylammonium chloride, tetramethylammonium oxalate, formamide and oligo-blockers, or by the utilization of dimethylsulfoxide during the polymerase chain reaction, to reduce the occurrence of mispriming and increase the amount of full length cDNA.

As all reverse transcriptases have a high error rate it is preferred that a proofreading activity is included in the reaction. Enzymes that have such an activity through their 3'-->5' exonuclease function are e.g. polymerases such as PFU, Ultma, Vent, Deep Vent, PWO and Tli, or E.coli. exonuclease III. In addition to reducing errors introduced into a cDNA during the RT the inclusion of a proofreading activity can also enhance full length cDNA synthesis (Hawkins, 2003).

It is further preferred that such a proofreading activity is also included in the PCR step of the method. On the one hand this will reduce errors introduced into the sequence during PCR as well as extend the size of cDNA molecules that can be amplified (Barnes, 1994). As such a 3'-->5' exonuclease activity can degrade primers used in the reaction it is preferred to protect primers from degradation, e.g. by introducing a phosphorothioate bond at the 3' terminal of the primer (Skerra, 1992; Di Giusto, 2004)or by using LNAs (locked nucleic acids).

It is preferred that during reverse transcription and/or PCR, enzymes or enzyme combinations are used that have a high

processivity and fidelity.

The present invention provides a method for full transcriptome coverage and no redundancy at least in principle. However, methodical properties such as mispriming and polymerase stop can influence the outcome nevertheless. Coverage of long mRNAs therefore depends on the used enzymes. Proof reading enzymes, used for long PCR methods, enlarge the scope of transcriptome coverage to at least 99% under ideal conditions for the given enzymes and method parameters (e.g. long extension times). Misannealing, which can be minimized by e.g. stringency and proof reading activities, may lead to a limited redundancy of the covered transcriptome.

The maximal length of an RNA molecule that can be displayed depends in particular on the quality of the reverse transcription and the amplification reaction, e.g. PCR. In principle long PCR can amplify DNA up to at least 35kb (Barnes., 1994) and RT-PCR up to at least 20kb (Thiel et al., 1997), so at least 99.9% of all transcripts can be displayed. However there exist exceptional long RNA molecules such as titin (NCBI accession number X90568) that has about 80kb length and even though in theory displayable, for practical reasons one will probably not tune the display to such transcripts.

Preferably, the reverse transcription is performed by a reverse transcriptase (RT) derived from the moloney murine leukemia virus, AMV, HIV-1, HIV-2, wherein the RNase H activity of these RTs is or is not present or is reduced. Reverse transcription can be performed by any enzyme that has a reverse transcriptase activity, e.g. Tth DNA polymerase that exhibits reverse transcriptase activity in the presence of $Mn^{2+}$. In addition the RNA can be degraded after the reverse transcription step, for example by an RNA degrading enzyme such as RNase to facilitate long PCR.

Preferably for reverse transcription as well as the amplification designed enzymes can be used, which are independent of a thermal cycling procedure, i.e. are capable of isothermal polymerisation.

It is further preferred that primers used for reverse transcription have a wobble in one or more bases, preferably 1, 2, 3, 4 or 5 bases, on their 3' end preferable between position 2 to 4 upstream of the last dT of the terminal dT stretch to fur-

ther increase specificity.

As used herein, the term "primer" refers to a mixture of oligonucleotides that have the same sequence.

As used herein, the term "wobble primer" refers to a mixture of oligonucleotides that have the same sequence, except for the base or bases that are part of the "wobble".

As used herein, the term "wobble" refers to a base within a primer pool that is present as a mixture of two or three or four different nucleotides.

As an example a wobble primer can consist of oligonucleotides of the sequence ACACACN, where N is referred to as the wobble and N is A or C or T or G. This wobble primer would then consist of a mixture of nucleotides of the sequence ACACACA, ACACACC, ACACACT and ACACACG.

When using N nucleotides as a mixture (N is a mixture of dA, dC, dG and dT) in wobble primers, nucleotides with universal bases can be used instead, such as deoxyinosine, 3-nitropyrrole 2'-deoxynucloside and 5-nitroindole 2'-deoxynucleoside. Universal bases will basepair with any nucleotide (dA, dC, dG, dT).

Therefore, according to the invention the primers used for reverse transcription or amplification preferably have a wobble substitution, e.g. dT may be substituted by dG, dA, dC and vice versa, in one or two bases on their 3' end, preferably between position 2 and 4 upstream of the last nucleotide, which is preferably dT for a 5' RT primer (for 5'-->3' strand synthesis) and a 3' amplification primer and dA for a 5' amplification primer. In the PCR step of the method such "wobble" primers can be used to increase specificity for both the 5' and 3' end of the cDNA molecules to be amplified.

Wobbles can also be used to anchor the primer that is complementary to the (e.g. poly A) tail of the original RNAs to the bases immediately upstream of the tail. In this case the selection into different pools can be achieved by primers selective to bases corresponding to bases after the cap of the original RNA. The same rational can be applied to the 3' end of the cDNA where wobbles can anchor a primer to the bases corresponding to bases downstream of the cap of the original RNA. Therefore the selection into different pools would be done through primers that select for bases that correspond to bases immediately upstream of the (e.g. poly A) tail of the original RNAs.

Even further preferred is a method according to the invention, wherein the wobble base represents a substitution by universal bases, preferably deoxyinosine, 3-nitropyrrole 2'-deoxynucloside or 5-nitroindole 2'-deoxynucleoside, which can basepair to any regular nucleobase such as dA, dT, dG or dC.

In a preferred embodiment specificity of the cDNAs amplified can be enhanced by using a combination of a primer that is exonuclease resistant on its 3' end (e.g. by including a phosphorothioate modification or a locked nucleic acid) with a 3'-->5' exonuclease function, such an activity being a property of e.g. Pfu DNA Polymerase. This will provide an "off switch" for amplification if a single mismatch, or multiple mismatches occur in the first 8 bases from the 3' end of the primer. (Yang et al., 2005). As the bases at the 3' prime end of the primers are the ones that discriminate the cDNA pools in methodologies presented in this invention, it is preferred that such an "off switch" is included to stop the extension of primers that mismatched to a cDNA, thus further enhancing specificity. Preferably in the inventive method a polymerase sensitive for a mismatch between nucleotides of the primer and the RNA or cDNA is used (either during reverse transcription or amplification), and the polymerase does not extend a mismatched exonuclease resistant primer, wherein it is preferred that the primers are resistant to exonucleases, especially preferred by a phosphorothioate modification or a locked nucleic acid.

In a preferred embodiment the primer used for reverse transcription consists of an oligo dT stretch or tail complementary stretch without an anchor. Since the selection of specific RNAs occurs on the cDNA level via the amplification step described above, specificity (of the anchor) at the RT level is not a necessity. Of course mispriming should be avoided, especially where the (oligo T) primer aligns with the sequence upstream of the (poly A) tail stretch of the RNA. Mispriming can be avoided using stringent conditions and/or using an enzyme with proof reading properties. Preferably the annealing temperature during the RT process lies in the range of 30 to 50 °C, depending on the melting temperature of the primer used.

Even though utilization of stringent conditions, such as the use of higher annealing temperatures, wobbled primers and oligoblockers, can decrease mispriming events during reverse tran-

scription (RT), a preferred solution is to not separate into different pools during RT. As the reverse transcriptase reaction allows for more misprimed sequences to be extended than the amplification reaction (e.g. PCR reaction), it is most preferred that the separation into pools is done during the amplification reaction. This also decreases the RT reactions necessary to one per sample to be displayed.

In a preferred embodiment the generation of full-length cDNA can be further advanced. Under non optimal conditions reverse transcriptase might stop polymerisation and/or fall off the RNA strand before reaching the 5' end of the RNA molecule. Generally this is attributed to secondary structure that has not melted and that the polymerase cannot traverse. Therefore generating also partial cDNA's. To lower the probability that such transcripts will interfere with subsequent (PCR) amplification one can exploit the intrinsic ability of reverse transcriptase to add a few ($Mn^{2+}$ and $Mg^{2+}$ dependent 1-6 nucleotides, Schmidt et al., 1999) dC once it reaches the 5' cap end of the mRNA molecule (Clark, 1988). The cap (Furuichi et al., 1975) structure greatly increases the nontemplated nucleotide adding capability of RT compared to mRNAs that do not have the cap structure. This can be used to further enrich for full-length cDNAs. In a subsequent PCR reaction cDNAs can be selected that have these Cs present by including dG nucleotides in the 5' primer at the complementary position. In such a case the tailing can be done with any nucleotide except dC, preferably with dT or dA. The general structure of a 5' PCR primer that would select for cDNAs that have three Cs added to their 5' end, then will be $dT_xdGdGdGHN_y$ or $dA_xdGdGdGHN_y$ or $dC_xdGdGdGHN_y$; H can either be dA or dC or dT.

In a preferred embodiment the method according to the invention, includes the generation of cDNA with a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, during reverse transcription, dependent on a 5' cap of the template RNA or mRNA.

Preferably the ability of a reverse transcriptase to generate the 3' sequence of oligo-dC is increased by the addition of $Mn^{2+}$ ions during the process of reverse transcription. The number of nucleotides added can be increased by including Mn ions in the RT reaction as described in Schmidt et al., 1999. Therefore in a preferred embodiment Mn ions are included in the RT reac-

tion to increase quantitatively the nontemplated addition of dC on the 3' end of cDNA in a cap dependent manner. Subsequent tailing is carried out with any nucleotide but dC in order to allow the selection of the cap dependent dC. The selection of added nucleotides during tailing can be done by providing just the wanted nucleotide during this process. One primer that is used for (PCR) amplification can therefore have the general formula $dP_xdG_y HN_z$; $P_x$ representing nucleotides complementary to the tail; dG being deoxyguanylate and y an integer between 0 and 10, preferably between 1 and 5 more preferable between 3 and 5; H can either be dA or dC or dT; and N is a sequence of nucleotides, its members independently selected from either dA, dC, dG, dT; z is an integer between 0 and 10 preferably between 0 and 5.

In a preferred embodiment the cap dependent 3' sequence of oligo-dC is used to isolate fully transcribed cDNA by using a second primer during the amplification which is specific for the oligo-dC sequence. This can be done by using a primer as described above.

Preferably the tailing of the 3' end of the cDNA is performed with a polynucleotide sequence, which is characterized by the absence of dC.

Even more preferred the DNA polymerase used for the amplification reaction is Taq DNA polymerase, Tfl DNA polymerase, aTaq DNA polymerase, Sequenase or Klentaq.

In special embodiments an enzyme with proof reading activity is used, preferably selected from PFU, Ultma, Vent, Deep Vent, PWO and Tli polymerases and E.coli. exonuclease III, is used during the amplification reaction. Such an enzyme can be used in addition to a polymerase, preferably in lower concentrations than the polymerase, or, if it is a polymerase itself, instead. Polymerases with proof reading activity prevent stops during DNA polymerisation and ensure long transcripts with significantly reduced errors. Preferably a combination of such an enzyme with proof reading capability is used with a polymerase above.

In a more preferred embodiment a primer used for the amplification reaction, which is preferably a polymerase chain reaction, is characterized by the general formula $dP_xdG_yHN_z$, wherein $P_x$ represents nucleotides complementary to the tail added to the 3' end of the cDNA, e.g. by the terminal transferase or by the

template switching property of the reverse transcriptase or by any other means, dG being deoxyguanylate and y an integer, including 0, preferably between 2 and 5, H being either dT or dA or dC, N being a sequence of nucleotides with a length z wherein its member nucleotides are independently selected from dA, dC, dG and dT, and z being an integer between 0 and 10, preferably between 0 and 5.

The 5′ end of primers used for RT and PCR can also be anchored to any sequence that will allow RT and PCR to be carried out. 5' and 3' PCR primers may have different melting temperatures and PCR conditions may be optimized by equalizing melting temperatures of primer pairs through the addition of an anchor to one or both of the primers. For certain downstream application it might also be of benefit to add a specific sequence such as a promoter that can start RNA Polymerase transcription (e.g. T7 or T3), or the inclusion of a restriction site.

In a most preferred embodiment the amplification is performed by PCR and the PCR primers contain anchoring sequences, which allow proper placement at the end of the 3' or 5' oligonucleotide end stretch of the cDNA.

Preferably an amplification primer or a reverse transcription primer contains a promoter for RNA polymerases. This optional feature allows the transcription of the cDNA and its PCR products.

Preferably the promoter enhances transcription by the T7, T3 or SP6 RNA polymerase.

Primers as well as PCR products can be labeled by any reporter group known in the art that enables their detection. Examples of reporter groups include fluorescent, chemiluminescent or radioisotopes and others known in the art. Any reporter group can be used that can be detected by such techniques as fluorescent measurement, light emission measurement, scintillation counting and other means known in the art.

Accordingly, the amplification primers and/or the amplification products are labeled by a reporter group, preferably fluorescence markers, chemiluminescence markers or radioisotopes in a preferred embodiment of the present invention. Preferably, these reporter groups are used to detect the DNA.

Recently, it has been shown that during transcription at least some eukaryotes such as Arabidopsis (Jin et al., 2004) can

add arbitrary nucleotides to RNA molecules on their 3' end immediately before the poly A tail, in most cases one or two. And even within a certain gene such nucleotide additions can be heterogenic. Therefore using a set of primers that discriminate for the first and/or second base upstream of the poly A during reverse transcription the intended pooling of a mRNA species into certain pools of cDNA will fail for certain transcripts. However a higher degree of pooling can be achieved using nucleotides upstream of these arbitrarily added nucleotides for the primers that define the fractions. A preferred primer for reverse transcription and PCR amplification that will be complementary to the poly A tail and select for bases above these additions be of following general structure $dT_xVN_yM_z$, where dT depicts deoxythymidylate with x repetitions; V and N are wobble bases where V is either dA, dC or dG ; N is a sequence of dA, dC, dG, dT, independently selected for each member; y is an interger equal to or greater than 0, preferable between 1 and 5, and M is either dA, dC, dG, dT; z is an integer above 1, preferably between 1 and 10. In such a primer $M_z$ will specify the cDNA pool to be amplified. As the number of nucleotides added to RNA molecules before polyadenylation might differ even for a specific transcript, a primer can be a mixture of oligonucleotides where the amount of y in N is different. An additional benefit is that in order for the Polymerase to be able to extend the chain the pressure for correct hybridization of bases after the wobble will strongly increase, therefore decreasing mispriming of the bases that define the fractions that will be amplified, hence specificity will be increased.

In a preferred method according to the invention arbitrary nucleotides, preferably one to three nucleotides, are added to the 3' end of the RNA or mRNA posttranscriptionally but before the addition of a (e.g. poly A) tail, and the primers for reverse transcription and/or PCR contain a sequence with a wobble corresponding to the arbitrary sequence.

A simplified version of the invention that will not specifically select for full length cDNA can be formulated by not using the 3' (dC) nontemplated adding function of the RT enzyme. There will still be a portion of full length cDNAs that have no dC added under normal RT condition. Tailing and PCR can be used as described to selectively amplify pools of cDNA. Each cDNA

will still be present in exactly one pool. However the extent of PCR products present that will correspond to full length RNA molecules will be determined by the quality of the RT reaction alone. As one does not select in the PCR reaction that follows for full length cDNA copies, more than one cDNA molecule might give rise to PCR products that represent a single RNA molecule, therefore introducing redundancy. Still each RNA molecule will be represented at least once. Redundancy can be minimized by employing methods that favour full length reverse transcription, such as addition of trehalose and/or betaine and/or increasing temperature as disclosed above. Even though redundancy might not be abolished in this embodiment, still each RNA molecule will be present at least once. This is an improvement compared to other DD techniques that can reach a representation for a certain mRNA molecule only statistically.

Also there exist species of RNA molecules that do not have a poly A tail and a cap structure. However such RNA molecules will also have 5' and 3' prime end that can be used during subsequent amplification to enrich for defined fractions as is intended by this invention.

Examples of RNA that may or may not have a cap and may have or may not have a poly A tail are non protein coding RNAs (ncRNA) such as ribosomal RNA (rRNA), transfer RNA (tRNA), micro RNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA) and small nuclear RNA (snRNA).

In RNA molecules that do not have a poly A tail, a tail such as a poly A tail, is added synthetically (e.g. enzymatically) prior to reverse transcription. Preferably a polynucleotide tail is added where the nucleotides are selected independently from A or C or G or U.

There exist species of RNA that have a certain size range, such as miRNA that are about 19 to 23 bases long. Therefore size separation of pools of amplification products derived from miRNAs is limited. When displaying miRNAs with a size range of 19 to 23 bases, a maximum of 5 molecules per pool can be resolved according to size. Therefore a quite large primer set (primer matrix) has to be used to resolve all miRNA molecules.

In an alternative approach the nucleotides of one end of the miRNA can be used to artificially enlarge the amplification products for a specific number of nucleotides according to the

terminal sequence of the miRNA. A general formula of such a primer is $P_xN_y$; with P representing nucleotides complementary to a tail; N represents nucleotides independently selected from either A, C, G, T; y is an integer between 0 and 10; x defines the number of nucleotides and is different for all $N_y$. If the 5′ end of the miRNA is used for such a size enlargement, a primer mix containing certain or all possible 5′ primers is used in combination with a 3′ primer that is selective for nucleotides at the end of the miRNA molecule, to define a certain pool. One possible primer matrix according to the principles of this invention is shown in figure 6.

Therefore in a preferred embodiment of the invention, amplification products - in particular full length amplification products - are generated that are differentially enlarged by certain numbers of nucleotides depending on the 5′ or 3′ sequences of the RNA or the full length cDNA.

Preferably, amplification products are generated that are differentially enlarged by defined numbers of nucleotides at their 5' and/or 3' end as compared to the RNA, by using different primers which have different lengths and recognize different 5' or 3' endsequences of the RNA or the full length cDNA. Also, pools of amplification products can be generated that are enlarged by defined numbers of nucleotides at their 5' and or 3' end as compared to the RNA, preferably by using as a 5' and/or as a 3' primer a mixture of primers, where each primer recognizes a different endsequence and has a different length. Preferably the primers have a defined sequence that is not complementary to the cDNA (during the amplification step) or RNA (during the reverse transcription step).

In prokaryotes mRNA has no poly A tail. Methods have been devised (Amara et al., 1997) that can make full-length cDNA from such molecules, such as adding a synthetic poly A tail. It is also within the scope of this invention to apply said methods to such molecules.

Preferably in a method according to the invention RNA or mRNA samples are used, which do not have a poly A tail or a tail sequence and, preferably a polynucleotide tail, especially preferred a poly A tail, is added synthetically (enzymatically) prior to reverse transcription.

To decrease the necessary amount of RNA to perform the dis-

play a preamplification step can be introduced before the amplification reaction, e.g. the PCR. Whether or not separation into pools was done in the reverse transcription (RT) step, the preamplification step has to be designed in a way that will amplify all the cDNA that is intended to be separated into defined pools in the amplification (PCR) step. The preamplification step can be introduced before the tailing or after the tailing, before the main amplification with the specific primers.

For example a promoter such as T7 RNA Polymerase can be added to the 5'- end of the primer that is used for priming the RT reaction. After the RT and before the amplification (PCR) RNA can be synthesized by an RNA polymerase specific to the promoter, for example T7 RNA polymerase. Such obtained amplified RNA has to be reverse transcribed again and depending whether or not a sequence is present downstream of the cDNA base that is complementary to the first base of the RNA and that can be used to prime the amplification reaction (PCR) on this end, such a sequence has to be added, for example by the template switching ability of the reverse transcriptase or by tailing with terminal deoxynucleotidyltransferase.

Another example for a preamplification step is the use of a preamplification PCR. In case of a terminal deoxynucleotidyltransferase tailed cDNA one primer has to be specific to the so added tail. For example a poly T tailed cDNA will need a primer such as $dA_x$ and a primer complementary to the poly A tail of the original RNA for example $dT_x$. Preferably one can preamplify only full length cDNA by including variable amounts of G, preferably between 1 and 5, to enable the primer to hybridize only to cDNA that was fully reverse transcribed and thus where the reverse transcriptase added Cs through its non templated nucleotide adding capacity in a cap dependent manner. Additionally the primer that can hybridize to the sequence complementary to the original RNA poly A tail can be anchored to the 5'-end of the poly A tail by including one or more wobbles. This will help to increase the length of the amplification (PCR) products obtained, by not needlessly amplifying long stretches of poly A tail. Such a primer has the general formula of $dT_xVN_y$ (x is the amount of dT, V either G, C or A, and N any nucleotide).

The discrimination of the cDNA into different pools can also be accomplished by using more than one amplification (PCR) step.

Each amplification (PCR) step would be more specific to the pool to be amplified. For instance in a first amplification (PCR) a primer such as $dT_x$ could be used for hybridizing to cDNA on the side that is complementary to the poly A tail side of the RNA. This primer would be not discriminating. In a next step, 3 reactions could be used with one being primed with $dT_xG$, one with $dT_xA$ and one with $dT_xC$. Thus creating 3 pools. In another PCR step the pools can be further divided. For example the $dT_xG$ pool can be divided in to 4 pools by priming with $dT_xGA$, $dT_xGC$, $dT_xGG$ and $dT_xGT$. Thus creating a total of 12 pools. The same rational can be applied to the other end of the cDNA and the pooling of both sides can be combined if desired. Such a sequential pooling can on the one hand further decrease the amount of starting material (RNA) needed or help to achieve a pooling tuned to the complexity of the sample to be investigated or the depth the investigator wants to cover.

It is within the scope of the invention that two or more samples of RNA that are combined by methods that yield normalization or subtraction between such RNA or their derived cDNA or PCR products can also be displayed according to disclosed methods. For example RNA can be (pre)treated by reducing RNA species in high concentration.

Therefore the present invention also relates to a method wherein the RNA in the sample or the cDNA is preamplified, preferably by PCR, or wherein RNA samples are normalized or subtracted. These (pre)treatment steps are preferably introduced before the main (discriminating) amplification c).

In a further embodiment of the method according to the invention the step of fractionation of the amplification product is performed by gel electrophoresis, pulsed-field electrophoresis, agarose gel electrophoresis, PAGE, capillary electrophoresis or pulsed-field capillary electrophoresis. Fractionation and comparison of samples generated during PCR can be done by any methods that allow distinguishing DNA molecules according to specific properties, such as sequence or size. Such methodologies include e.g. gel electrophoresis, such as polyacrylamide gel electrophoresis (PAGE). Especially suited is capillary gel electrophoresis because of its high resolving power and because many samples can easily be analyzed in parallel (Behr et al., 1999). Standard capillary electrophoresis machines today can

process 384 samples in one run. And through the use of appropriate protocols one can achieve the necessary resolving power over a wide range of sizes. Especially long fragments can be resolved sufficiently with pulsed-field capillary electrophoresis (Heller et al., 2001; Morris et al., 2001). Also methods have been devised to collect fractions of the molecules separated in a capillary (Irie et al., 2000; Magnusdottir et al., 2001). Such fractions that represent transcripts of interest (e.g. differentially expressed genes) can then be subjected to further analysis, such as sequencing, to determine the identity of the molecule.

In a preferred embodiment the detection of differences between two cDNA pools can be included into the PCR step of the method. Such methodologies, that can measure DNA levels and/or the change of such levels during PCR, are generally referred to as "real time PCR"(e.g. Wittwer et al. 1997a, Wittwer et al. 1997b). One procedure applicable measures fluorescence from SYBR Green I dye, that has been added to the PCR reaction. This dye will show a strong increase in fluorescence when intercalating with double stranded DNA. The more DNA present in the reaction the stronger the fluorescence. Therefore providing the means to calculate the amount of DNA present in the reaction. The measurement is usually done at the end of the polymerization step as the products will be double stranded. As the amount of DNA increases during subsequent cycles, so will the fluorescence. This data can be compared e.g by blotting on a curve and curves can be compared between different samples. If there is only one cDNA amplified the curve will have an exponential phase. By comparing the start of the exponential phase between samples one can calculate relative amounts of DNA. If two or more products will be present in a given sample, curves will add onto each other and will result in a change of steepness. Still such data can be compared between different samples, if the cDNA pool amplified is not too complex. It is preferred that the amount of cDNAs amplified in one PCR reaction is reduced to such a point that differences in amplification of each product can still be detected. Another way of analysing a sample, when using real time PCR, is the possibility of calculating a melting curve of the DNA present at the end of each cycle. Each PCR product present in the reaction will melt at a specific temperature,

that is defined by the sequence. As a product is melting into its single strands the dye is not intercalating with DNA anymore, whereby a drop of fluorescence will result. Products that will melt at different temperatures will be represented as an increase and then a decrease of the steepness at different areas of the melting curve. Comparing curves between cycles and different samples of the same cDNA pool to be amplified, one can identify samples with a different profile. A melting curve analysis can also be done only at the end of the PCR, however differences in products that have plateaued might not be detected.

The use of such analysis methods preferably in an automated procedure can rapidly screen for samples that show differences during or at the end of PCR and eliminate the need to analyse all of them through a fractionation procedures such as gel electrophoresis.

Therefore in a preferred embodiment the detection of differences between two cDNA pools, preferably corresponding pools between two or more samples, is done during or at the end of the amplification step, preferably being a PCR step, of the method, preferably by real time PCR.

Clearly the methods described in this invention can be integrated into a partially or fully automated process. For example the size fractionation, comparison of different samples, collection of fractions, and sequencing can be all done on a chip. The technologies have already been demonstrated in principle (Xie et al., 2001).

Such an integration of the method onto a chip can also be done for the PCR step of the method. It is preferred that the primers, or primer pairs are spotted and/or bound onto a solid support to enable solid phase DNA amplification (e.g. by PCR). Through microarraying or nanoarraying or single molecule arraying the whole set of primers onto a solid support, a single reaction vessel can be created that contains the whole primer matrix or parts thereof. Therefore instead of using an own reaction vessel for each 5′-3′ primer pair, a single reaction vessel contains all 5′-3′ primer pairs. It is preferred that each primer pair is located in a defined area.

A primer matrix can also be created where each 5′-3′ primer pair is bound onto its own solid support (e.g. a bead) and the solid support itself is then arranged in such a manner that each

primer combination or primer pair can be addressed independently. For instance a matrix of beads can be arranged randomly in microwells and the location of each bead decoded by methods known to the art (e.g. Gunderson et al., 2004). The method as described by Gunderson can thus be adapted and incorporated by the present invention. This method requires only a few labels and several sequential hybridizations to identify thousands of different DNA sequences with great accuracy. Regarding the primer matrix, the primers on the areas may comprise the information for decoding themselves (e.g. as sequence which can be decoded by other labeled oligo nucleotide by several steps of hybridisation, detection and washing) or together with the primers of the invention decoding oligonucleotides can be bound onto the carrier.

Therefore in a preferred embodiment a 5'primer or a 3'primer or a 5'-3' primer pair is spotted and/or bound onto a solid support, preferably the spot or the solid support is located in a defined 2-dimensional or 3-dimensional area and it is preferred that such an area contains information to decode the identity of the primer or primer pair located in that area.

In a preferred embodiment the primer matrix is generated by direct on chip synthesis of the 5' primers or 3' primers or 5'-3'primer pairs.

In a preferred method according to the invention the steps of reverse transcription and amplification are performed in one step.

It is obvious to the one skilled in the art that other possibilities exist to enrich for full-length cDNA. Such methods can easily be adapted to amplify subsequently pools of full length cDNA that are defined through nucleotides that are complementary to the 5' end and/or nucleotides on the 3' end upstream of the poly A tail.

For example the CapSelect method (Schmidt et al., 1999) can be used to substitute the steps of reverse transcription and tailing the 3' end of the cDNA. This method is also drawn on the nontemplated addition of C nucleotides of reverse transcriptase, and addition of a poly N tail, preferably a poly A tail. However the poly A tail consists of ribonucleotides instead of deoxyribonucleotids as it is claimed that terminal transferase will preferably add only 3 to 4 ribonucleotides, compared to an

unspecified amount of deoxyribonucleotids. This is important for the method as in a next step an adaptor will be specifically linked to the 3' end of the cDNA selecting for 3'-AAACCC-5'. Using a primer that is complementary to this adaptor and a primer complementary to the poly A tail it is intended to amplify full length cDNA. Instead of using universal primers to amplify all cDNAs one can use sets of primers that select for the 5' and/or 3' nucleotides and will amplify pools of cDNAs. The primer structure for hybridizing to the poly A tail and selecting for x amount of bases before the poly A tail will be the same as described. For amplifying the 5' end of mRNAs one can use the adaptor sequence followed by 3 to 4 Ts and 3 to 4 Gs and an x amount of bases that correspond to the 5' sequence of the RNA transcripts. Increasing x will decrease the number of transcripts amplified through such a primer.

Preferably the polynucleotide tail added by the terminal transferase consists of ribonucleotides, preferably 2 to 6 ribonucleotides, preferably adenosine or thymidine, after a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, which is generated during reverse transcription, dependent on a 5' cap of a template RNA or mRNA, and the second amplification primer (5' primer) comprises complementary nucleotides to the ribonucleotides and oligo-dCs and 1 to 10 nucleotides upstream of the olio-dC sequence.

The method for amplification of certain cDNAs, derived from mRNA, can also be used to determine the expression level of individual mRNAs. Preferably for measuring the expression level of messenger RNA a method according to the invention is employed, including the separation and/or identification of amplification results. Like in common DD results of certain cell lines can be compared for better visualisation of different expressions.

Furthermore the present invention provides a kit for a method according to the invention comprising a DNA polymerase, preferably Taq polymerase, a reverse transcriptase or mixtures of different polymerases, cofactors or salts of metal ions, preferably $Mg^{2+}$ and $Mn^{2+}$, requested by a polymerase, primers and optionally a terminal transferase and additional buffer substances. The primers are preferably selective for the 5' and 3' end of a RNA or cDNA as described above. Even more preferred the primers of the primerset are selective for the first, second,

third, fourth and/or fifth terminal nucleotide of the RNA or cDNA (of either end, independently from each other). Preferably the kit comprises 1, 2, 3, 4, 5, 10, more than 10, more than 20 or more than 30 primer pairs. With this kit a display method according to the invention can easily and quickly be performed. Preferably instructions for the execution of the method are also included. Preferably the full primerset (primer matrix) or parts thereof - preferably one of each primer pair - are bound to a solid support such as a chip, preferably a glass chip.

In another aspect of the present invention there is provided a solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a pair of primers, wherein different primers are selective for at least two, preferably at least three, even more preferred at least four, most preferred at least 10, different 3' or 5' end sequences of target polynucleotides, preferably selected from RNA or cDNA, and each area comprises primers for at least two different end sequences (e.g. a pair of a 5' and a 3' primer), preferably one primer is selective for a (e.g. poly A) tail or a complement thereof and even more preferred selective for the next 1 to 10 nucleotides, preferably the next 1, 2, 3 or 4 nucleotides (of either the 5' and/or the 3' primer. The primers can be specific as described above. These primers may also, e.g. in these combinations, be used in the kit or the method. A primer pair in one area selective for mRNA is specific for the poly A tail and the 5' (cap) end, as well as the next 1 nucleotide (different in each area). Preferably the solid surface comprises at least two, at least 4, at least 10, at least 20 or at least 40 different primer pairs.

Provided is further a solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a 5' primer and a 3' primer or a pair of primers, wherein the primers are as defined above being specific for both ends of a cDNA or a RNA, especially comprising formula $dP_xdG_y\,HN_z$; in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or primer pairs.

Also provided is a solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas,

preferably in spots, a 5' primer or a 3' primer comprising the general formula as defined above ($dP_xdG_yHn_z$) or a complement thereof, preferably selected from a primer comprising one of the following sequences: AGA GAT TTT TTT TTT TTT TT GA, AGA GAT TTT TTT TTT TTT T GG, AGA GAT TTT TTT TTT TTT T GC, AGA GAT TTT TTT TTT TTT TT GT, AGA GAT TTT TTT TTT TTT TT CA, AGA GAT TTT TTT TTT TTT TT AG, AGA GAT TTT TTT TTT TTT TTT AA, AAA AAA AAA AAA AAA GGG ATA, AAA AAA AAA AAA AAA GGG ACA, AAA AAA AAA AAA AAA GGG AGA, AAA AAA AAA AAA AAA GGG AAA, AAA AAA AAA AAA AAA GGG GTA, AAA AAA AAA AAA AAA GGG CTA, AAA AAA AAA AAA AAA GGG TTA, in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or primer pairs.

In a further aspect a method for amplifying a pool of polynucleotide molecules in a sample is provided, comprising the steps of:

a) obtaining (or providing) a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining (or providing) a sample of full length cDNA,

b) tailing the 3' end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides (i.e. nucleotide 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10), upstream of the 3' polynucleotide tail of the cDNA, preferably also comprising unspecific wobble nucleotides, especially preferred in the upstream portion within the 1 to 100 nucleotide range - wherein at least 1, 2 ,3, 4, 5, 6, 7, 8, 9 or 10 nucleotide (s) is/are specific for a nucleotide of the cDNA; the method further preferably defined as given above.

The method preferably includes the use of a primer, which is specific for a 3' tail of the RNA (optionally an added 3' tail) or mRNA (preferably the endogenous 3' poly A tail) and the next

1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides (i.e. nucleotide 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10), upstream of the 3' tail is used for the reverse transcription and/or a first amplification primer is specific for a 5' tail of the cDNA, which is complementary to the 3' tail of the RNA or mRNA, and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides (i.e. nucleotide 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10), downstream of the 5' tail. Of course, also both or independently from one another, in the 5' and the 3' primers, the nucleotide stretch (1 to 100 nucleotides) next to the tail (or tail complement) may comprise unspecific nucleotides (such as wobble nucleotides) in addition to at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more specific nucleotides selecting for different pools (with or without pool separation) of the amplification product. Such primers can also of course be used for the reverse transcription as well. Products of such wobble primers (where any nucleotide may be present at a corresponding position to the wobble position) still fulfil the requirement of full-length cDNA or full-length amplification product although the sequence may not be preserved at positions that correspond to the unspecific nucleotide positions. Using longer unspecific nucleotide primer stretches primer-loops may occur thus resulting in a product with an altered length which is however still considered to be equivalents to full-length cDNA or amplification products for the present invention.

The present invention is further illustrated by the following example and figure without being limited thereto.

Fig. 1 shows a graphical representation of a procedure using the principles of the present invention to amplify full length cDNA into defined subpopulations. The oligo-T and oligo-G stretches indicated in the primers are terminal anchors of the primers and the N, NN and $NN_y$ indicate nucleotides selected from A, T, G, C for defining the subpopulations.

Fig. 2: Gel elecrophoresis product of a display of a pool of artificial mRNA molecules according to example 2: In lane 1-7 primer nr. 8 was used together with primers nr. 1-7. In lanes 8-13 primer nr. 1 was used together with primers 9-14.

Fig. 3: Gel elecrophoresis product of an organised full length expression display of mouse liver RNA according to example 3.

Fig. 4: Gel elecrophoresis product of an organised full length expression display of mouse liver RNA according to example 4 with PCR preamplification.

Fig. 5: NCBI nucleotide database screening for non redundant full length mouse cDNA clones. 99% of the 16854 clones present in the database are between 200 and 6000 bases in length and 99.9% below 7500 bases and no clone was above 9500 bases.

Fig. 6 shows how RNA without a cap and a tail, such as miRNA, is divided into subpopulations. a) Steps of the procedure. Primer nomenclature as in fig.1 b) One possible Primer matrix that increases the resolving power by increasing the size range of the amplification products. The B in the 5' primers indicate a mixed base (wobble) that is selected from T, G, C.

E x a m p l e s

Expample 1: Principle of procedure

An overview of an example of a polynucleotide amplification procedure according to the present invention is given in fig. 1. Anchored primers for reverse transcription can follow the formula $dT_xV$ or $dT_xVN_y$ (with x defining a repetition of deoxythymidylate ("dT") that has in combination with the V or $VN_y$ part a high enough specificity and melting temperature to successfully prime the RT reaction; V is either deoxyguanylate ("dG"), deoxyadenylate ("dA") or deoxycytidylate ("dC"); N is a sequence of y nucleotides selected independently from dA, dC, dG or dT; if more than one base is used for anchored primers y defines their number; increasing y will increase the subpopulations of reverse transcripts. The tailing of the 3' end of such synthesized cDNA with a poly N tail can be done using terminal transferase - N stands for a sequence of uniformly nucleotides selected from dT, dA, dC or dG, independently form the "N" indicated in the reverse transcription primers. The amplification of the cDNA pools, optionally preselected by the $VN_y$ stretches defined by polymerase chain reaction (PCR) is performed using primers that hybridize to the 5' and 3' end of the cDNA by selecting to terminal sequences like the nucleotides appended by the terminal

transferase (3' of the cDNA; corresponding to the 5' end of the original mRNA) and the poly T stretch (5' of the cDNA; complementary to the 3' poly A tail of the original mRNA) and selecting for an oligo $dN_y$ sequence, wherein N is a stretch of several nucleotides, again independently selected from dT, dA, dC or dG and define different mRNA populations through y bases immediately adjacent to the tails.

Example 2: Amplification of all the RNA molecules present in a sample.

As all the mRNA molecules present in a sample are not known, even in well characterized tissues such as liver from human or mouse, and their number maybe in the thousands, an approach was chosen to use a pool of 14 artificially synthesized mRNA molecules to show that all RNA molecules present in that sample are detected. For this purpose stretches with different lengths of genomic DNA from the mouse GAPDH gene (GenBank Accession: NC_000072) were amplified by standard PCR. 5' primers were used that had an anchored T7 polymerase promoter. Following the promoter a combination of 4 bases were included that would define each mRNA on its 5' end. Primers chosen for the 3' end included, after the GAPDH specific sequence, a combination of 4 bases, that would define these molecules upstream of the poly A tail. Following these 4 bases 21 dT or in one of the RNAs 22 dT were included for the poly A tail. The PCR products obtained were gel purified and subjected to standard T7 RNA transcription, that included 7-methyl-guanosine (the cap). Such T7 reaction products were again gel purified. A mixture of such obtained synthetic mRNA molecules was used as starting material for the reverse transcription (RT). To show that the methods disclosed are specific for the 5' as well as 3' end of the mRNA molecules, 7 of the 14 molecules had a fixed sequence on the 5' end and varied in the 3' end and 7 molecules had a fixed sequence on the 3' end and varied at the 5' end. For sequence of the resulting artificial mRNA molecules see table 1.

Reverse transcription was carried out in a 25µl reaction consisting of 50mM Tris-HCl (ph 8.3), 75 mM KCl, 3mM $MgCl_2$, 10mM DTT, 0.6M trehalose, 2M betaine, 0.5mM dATP, 0.5mM dTTP, 0.5mM dCTP, 0.5mM dGTP, 40nM primer (oligo $dT_{(21)}$). The artificial mRNA mix was included at a level of about 4 ng/µl reaction volume.



IPR2025-01353; IPR2025-01355
CSHL EX2002

The reaction (without the RT) was heated to 70 °C for 20 sec, 32°C for 2 min. Then the Molony Murine Leukemia Virus Reverse Trancriptase lacking RNase H activity (M-MLV RT (-H)) was added a concentration of 4-8 units/µl reaction volume and the reaction continued for 2 min at 32°C. At this point 0.5 U of Pfu polymerase was added. Four cycles at 50°C for 10min and 60°C for 10 sec were used to fully reverse transcribe the RNA, after that the reaction was cooled to 4°C.

Table 1: Sequences of artificial RNAs to be displayed

| RNA Nr. | 5' sequence ..... 3 sequence | Length of RNA | Predicted Length of PCR product |
|---|---|---|---|
| 1 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CTG TGG GCA AGG TC **TC** $A_{(21)}$ | 1515 | 1531 |
| 2 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....TGT CCC TTT TGG TC **CC** $A_{(21)}$ | 1333 | 1348 |
| 3 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....TTG TGG AAG GGC TC **GC** $A_{(21)}$ | 1243 | 1258 |
| 4 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CGA TGC CCC CAT TC **AC** $A_{(21)}$ | 1112 | 1128 |
| 5 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CAC CAC CAT GGA TC **TG** $A_{(21)}$ | 942 | 958 |
| 6 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....CAA CGG GAA GCC TC **TA** $A_{(21)}$ | 843 | 859 |
| 7 | $^{m7}$G **AT** AT CCG CCC TGT GCT.....ACG ACC CCT TCA TC **TT** $A_{(21)}$ | 736 | 753 |
| 8 | $^{m7}$G **AT** AT GAT GCG CAA AGG.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 687 | 703 |
| 9 | $^{m7}$G **AC** AT GGC CTA AGC AAG.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 786 | 802 |
| 10 | $^{m7}$G **AG** AT GAG CAC GGA AGC.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 880 | 896 |
| 11 | $^{m7}$G **AA** AT CCG CTT TGA TTT.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 994 | 1010 |
| 12 | $^{m7}$G **GT** AT CAG CTC CCC ATC.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 1070 | 1086 |
| 13 | $^{m7}$G **CT** AT GGG CCT GGG TCA.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 1178 | 1194 |
| 14 | $^{m7}$G **TT** AT GGG CGG AGT GGA.....TCC TGC GGC CCA TC **TC** $A_{(21)}$ | 1287 | 1303 |

The cDNA was phenol/chloroform extracted, ethanol precipitated and dissolved in 10 µl of $H_2O$. Poly T tailing was carried in a 20 µl reaction containing 0.2mM dTTP, 100mM cacodylate buffer (pH 6.8), 1mM $CoCl_2$, 0.1mM DTT and 15 units of terminal deoxynucleotidyl transferase for 15 min at 37 °C. The reaction was Phenol/Chloroform extracted, ethanol precipitated and dissolved in 100 µl $H_2O$.

PCR was carried in a 10 µl reaction containing 2.5µl tailed cDNA, 50mM Tris (pH 8.8), 14mM $(NH_4)_2SO_4$, 4mM $MgCl_2$, 0.2mM of each dNTP, 0.2 µM of each primer (table 2, fig. 2), 0.5 units Taq

polymerase and 0.03 units Pfu polymerase. Samples were denatured at 93°C for 30 sec, and cycled 21 times at 93°C for 10 sec, 55°C for 30 sec, 72°C for 10 min. The initial annealing temperature of 55°C was decreased each cycle by 0.5°C min. Thereafter followed 19 cycles at 93°C for 10 sec, 49°C for 30 sec, 72°C for 10 min with a final extension step at 72°C for 2 min.

Table 2: PCR primers

| Primer nr. | Sequence |
|---|---|
| 1 | AGA GAT TTT TTT TTT TTT TT **GA** |
| 2 | AGA GAT TTT TTT TTT TTT T **GG** |
| 3 | AGA GAT TTT TTT TTT TTT T **GC** |
| 4 | AGA GAT TTT TTT TTT TTT TT **GT** |
| 5 | AGA GAT TTT TTT TTT TTT TT **CA** |
| 6 | AGA GAT TTT TTT TTT TTT TT **AG** |
| 7 | AGA GAT TTT TTT TTT TTT TTT **AA** |
| | |
| 8 | AAA AAA AAA AAA AAA GGG **ATA** |
| 9 | AAA AAA AAA AAA AAA GGG **ACA** |
| 10 | AAA AAA AAA AAA AAA GGG **AGA** |
| 11 | AAA AAA AAA AAA AAA GGG **AAA** |
| 12 | AAA AAA AAA AAA AAA GGG **GTA** |
| 13 | AAA AAA AAA AAA AAA GGG **CTA** |
| 14 | AAA AAA AAA AAA AAA GGG **TTA** |

2.5 µl of the reaction product was added to 2.5 µl 100% formamide loading buffer and denatured at 96 °C for 2 min and cooled on ice, loaded on a 3.5 % acrylamide, 7M urea gel and run at 180V for 2h and subsequently silver stained.

The results can be seen in figure 2. In lane 1-7 primer nr. 8 was used together with primers nr. 1-7. In lanes 8-13 primer nr. 1 was used together with primers 9-14. In each lane the PCR amplified the specific product, as the predicted length of the PCR products shown in table 1 perfectly matches the lengths of the products on the gel.

The primer combination in lane 1 amplifies the 2 specific products corresponding to the RNA to be displayed. The reaction carried out in lane 1 is the most demanding reaction of the set

as two products have to be amplified, the ones representing RNA 1 and 8. In addition as can be seen in table 1 the RNA molecules were designed in a way to enhance miss priming of the PCR reaction represented in lane 1. All RNA molecules (except RNA Nr.6)have on position 3 and 4 before the T tail the same bases as on position 1 and 2 of RNA number 1 and 8. As primer nr. 1 uses its last two bases to select the cDNA to be amplified, it potentially can miss prime with any of the other cDNA present in the sample. And as cDNA moleclules 1-8 have the same bases on its other end it is only the two last bases of primer 1 that selects between these 8 molecules in reaction 1. However only the two RNAs to be amplified are displayed, showing the great specificity that can be reached in such a display.

Lane 2 shows the limits of this reaction. In addition to the band to be amplified a faint band that corresponds to RNA nr. 4 was amplified. However the band to be amplified is several times more intense than the band that resulted through the miss priming event. Additionally even though a touch down protocol was used for the thermal cycling, the same profile was used for all the different primer combinations. To show that in principle the reaction can be optimized for a specific primer set another reaction was carried that included 1.302M betaine and 1.3% DMSO. The product is shown in lane 14 and only the correct band is displayed.

Example 3: Amplification of total RNA from mouse liver.

Total RNA (13µg) from mouse liver was reverse transcribed, tailed and PCR amplified under the same conditions as for the whole artificial RNA set in example 2.

The same conditions for RT, tailing and PCR were used to show that it is possible to display RNA molecules from a complex mixture under the same stringent conditions as it is possible with artificial RNA. The PCR products were run on a 0.7% agarose gel to resolve longer molecules. As can be seen in figure 3 each primer combination amplifies a unique profile.

The length distribution of the RNA molecules displayed in figure 3, matches very well the length distribution of cDNA in the NCBI nucleotide database screening for non redundant full length mouse cDNA clones (fig. 5).

As on averaged about 10-15 bands can be seen on each lane

the full set of 768 primer combinations ($3x4^4$) displays about 7680-11520 PCR products representing full length RNA molecules expressed in liver. Being able to resolve such profiles on a simple 0.7% agarose gel shows the power of the method and the ease, with what it can be performed in the laboratory.

Example 4: Utilization of PCR Preamplification step.

Preamplification was carried out in a 10µl reaction containing 2.5µl of a purified tailed cDNA generated as in example 3, 50mM Tris (pH 8.8), 14mM $(NH_4)_2$ $SO_4$, 4mM $MgCl_2$, 0.2mM of each dNTP, 0.2µM of Primer $dT_{(21)}$, 0.2µM of Primer $dA_{(21)}$, 0.5 units Taq polymerase and 0.03 units Pfu polymerase. Samples were denatured at 93°C 30 sec, and cycled 11 times at 93°C 10 sec, 42°C 30 sec, 72°C 11 min with a final extension step at 72°C for 2 min. This reaction was diluted to 1ml by adding 990µl $H_2O$. 2.5µl of this dilute was used in the following PCR using the same PCR conditions as in example 3 with the difference that the final 2 cycles where omitted. As can be seen in figure 4 a distinctive banding can be seen for each primer combination. The preamplification step as used in the example shows that 500ng of total RNA are enough to provide material for even a large display with 768 reactions per RNA sample, as might be needed to resolve highly complex RNA mixtures from tissues like brain. The amount of RNA used can easily be reduced by increasing the cycles during preamplification. Therefore the sensitivity of the display should meet even the most demanding applications, where only limited amounts of RNA are available. Furthermore the amount of bands per primer pair has increased compared to example 3. As the high annealing temperatures used in the discriminating PCR step, in both example 3 and 4, exert an extremely high stringency, some low level transcripts might be just below the detection limit when no preamplification step is used.

Expample 5: miRNA Display

An overview of an example of a polynucleotide amplification procedure according to the present invention, where the RNA has no cap and no poly A tail, in this case microRNA (miRNA), is given in fig. 6.

To be able to reverse transcribe full length cDNA from miRNA that has no Poly A tail, a synthetic tail is added by using poly

A polymerase. In a second step the Poly A tailed RNA is reverse transcribed into cDNA using an oligo dT primer. After polynucleotide tailing of the cDNA by terminal transferase, the cDNA samples have sufficient long nucleotide terminals that allow for primers to anneal and amplify the cDNA sequences into different pools according to the 5′ and 3′ sequences of the original miRNA molecules (fig. 6a).

As microRNA (miRNA) molecules have a size of about 19-23 bases, a sufficiently large primer matrix (all possible combinations of 5′ and 3′ primers) has to be used to separate all miRNAs into different pools where differences in expression between samples can be measured.

Alternatively one end of the miRNA sequence can be used to artificially enlarge the amplification products for a specific number of nucleotides according to the terminal sequence of the miRNA. Figure 6b shows an example of such a matrix. In this case the 5′ primer is used to split miRNA amplification products into different lengths according to the second and third nucleotide of the miRNA. In cases where the comparison of different pools is done by an additional size fractionation (such as capillary gel electrophoresis) all 16 5′ primers of fig. 6b can be used as a primer mix. This 5′ primer mix together with one 3′ primer defines a pool. Therefore 192 pools have a similar resolution when compared to 192 x 16 = 3072 pools, where each 5′ primer would be used together with each 3′ primer.

R e f e r e n c e s

Amara et al., 1997, Nucleic Acids Res. 25, 3465-3470.
Barnes, 1994, PNAS 91, 2216-2220.
Behr et al., 1999, Electrophoresis 20, 1492-1507.
Carninci et al., 1998, Proc. Natl. Acad. Sci. U. S. A 95, 520-524.
Clark, 1988, Nucleic Acids Res. 16, 9677-9686.
Di Giusto et al., 2004 Nucleic Acids Res. 32, e32 1-8.
Frohman et al., 1988, Proc. Natl. Acad. Sci U. S. A 85, 8998-9002.
Furuichi et al., 1975, Nature 253, 374-375.
Gunderson et al., 2004 Genome Res. 14, 870-877.
Hawkins et al., 2003, Biotechniques 34, 768-773.

Heller et al., 2001, Methods Mol. Biol. 162, 293-305.
Irie et al., 2000, Electrophoresis 21, 367-374.
Jin et al., 2004, RNA. 10, 1695-1697.
Kovarova et al., 2000, Nucleic Acids Res. 28, e70i-e70iii.
Liang et al., 1992, Science 257, 967-971.
Magnusdottir et al., 2001, Methods Mol. Biol. 162, 323-331.
Matz et al., 1997, Nucleic Acids Res. 25, 2541-2542.
Mizuno et al., 1999, Nucleic Acids Res. 27, 1345-1349.
Morris et al., 2001, Methods Mol. Biol. 162, 307-321.
Prashar et al., 1996, Proc. Natl. Acad. Sci. U. S. A 93, 659-663.
Prashar et al., 1999, Methods Enzymol. 303, 258-272.
Rebagliati et al., 1985, Cell 42, 769-777.
Schmidt et al., 1999, Nucleic Acids Res. 27, e31.
Spiess et al., 2002, Anal.Biochem. 301, 168-174.
Skerra, 1992, Nucleic Acids Res. 20, 3551-54.
Thiel et al., 1997, Anal.Biochem. 252, 62-70.
Welsh et al., 1992, Nucleic Acids Res. 20, 4965-4970.
Wittwer et al., 1997a, Biotechniques 22, 130-138.
Wittwer et al., 1997b, Biotechniques 22, 176-181.
Xie et al., 2001, Methods Mol. Biol. 162, 67-83.
Yang et al., 2005, Biochem. Biophys. Res. Commun. 328, 265-272.


IPR2025-01353; IPR2025-01355
CSHL EX2002

C l a i m s :

1. Method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of:

a) obtaining a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining a sample of full length cDNA,

b) tailing the 3′ end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 10 nucleotides upstream of the 3' polynucleotide tail of the cDNA.

2. Method according to claim 1 characterized in that a primer, which is specific for a 3' poly A tail of the RNA or mRNA and the next 1 to 10 nucleotides upstream of the poly A tail is used for the reverse transcription and/or a first amplification primer is specific for a 5' poly-T stretch of the cDNA, which is complementary to the 3' poly A tail of the RNA or mRNA, and the next 1 to 10 nucleotides downstream of the 5' poly-T stretch.

3. Method according to any one of claims 1 or 2 characterized in that the tailing of the 3' end is performed using terminal transferase.

4. Method according to any one of claims 1 to 3 characterized in that the amplification is performed by PCR.

5. Method according to any one of claims 1 to 4 comprising an additional step of separation of the amplification products according to their length, preferably by gel electrophoresis, preferably polyacrylamide gel electrophoresis, capillary electrophoresis, pulsed-field electrophoresis, agarose gel electrophoresis, PAGE or pulsed-field capillary electrophoresis.

6. Method according to any one of claims 1 to 5 comprising an additional step of determining the identity or sequence of the

amplification products, preferably by an automated process on a chip.

7. Method according to any one of claims 1 to 6, wherein the sample contains total RNA or mRNA, preferably purified RNA or mRNA, from a specimen.

8. Method according to any one of claims 1 to 7, wherein high stringency conditions are used for the processes of reverse transcription and/or the amplification.

9. Method according to any one of claims 1 to 8, wherein the selectivity of the reverse transcription and/or amplification, preferably a polymerase chain reaction, is increased by the utilization of trehalose, betaine, tetramethylammonium chloride, tetramethylammonium oxalate, formamide and oligo-blockers, or by the utilization of dimethylsulfoxide during the polymerase chain reaction, to reduce the occurrence of mispriming.

10. Method according to any one of claims 1 to 9, wherein the reverse transcription is performed by a reverse transcriptase (RT) derived from the moloney murine leukemia virus, AMV, HIV-1, HIV-2, wherein the RNase H activity of these RTs is present, reduced or not present.

11. Method according to any one of claims 1 to 10, wherein primers used for reverse transcription or amplification have a wobble substitution, preferably dT is substituted by dG, dA or dC or vice versa, in one or two bases on its 3' end, preferably between position 2 and 4 upstream of the last nucleotide, wherein the wobble base preferably represents a substitution by universal bases, preferably deoxyinosine, 3-nitropyrrole 2'-deoxynucloside or 5-nitroindole 2'-deoxynucleoside, which can basepair to any regular nucleobase such as dA, dT, dG or dC.

12. Method according to any on of the claims 1 to 11, using a polymerase sensitive for a mismatch between nucleotides of the primer and the RNA or cDNA, and the polymerase does not extend a mismatched exonuclease resistant primer, wherein it is preferred that the primers are resistant to exonucleases, especially pre-

ferred by a phosphorothioate modification or a locked nucleic acid.

13. Method according to any one of claims 1 to 12, whereby cDNA with a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, is generated during reverse transcription, dependent on a 5' cap of the template RNA or mRNA, and preferably wherein the cap dependent 3' sequence of oligo-dC is used to isolate fully transcribed cDNA by using a second primer during the amplification which is specific for the oligo-dC sequence.

14. Method according to any one of claims 1 to 13, wherein the tailing of the 3' end of the cDNA is performed with a polynucleotide sequence, which is characterized by the absence of dC.

15. Method according to any one of claims 1 to 14, wherein the DNA polymerase used for the amplification reaction is Taq DNA polymerase, Tfl DNA polymerase, aTaq DNA polymerase, Sequenase or Klentaq and/or an enzyme with proof reading activity, preferably selected from the PFU, Ultma, Vent, Deep Vent, PWO and Tli polymerases and E.coli. exonuclease III.

16. Method according to any one of claims 1 to 15, wherein the RNA in the sample or the cDNA is preamplified, preferably by PCR, or wherein RNA samples are normalized or subtracted.

17. Method according to any one of claims 1 to 16, wherein a primer used for the amplification reaction is characterized by the general formula $dP_xdG_yHN_z$, wherein $P_x$ represents nucleotides complementary to the sequence added to the 3' end of the cDNA, dG being deoxyguanylate and y an integer, preferably between 2 and 5, H being either dT or dA or dC, N being a sequence of nucleotides with a length z wherein its member nucleotides are independently selected from dA, dC, dG and dT, and z being an integer between 0 and 10, preferably between 0 and 5.

18. Method according to any one of claims 1 to 17, wherein the amplification is performed by PCR and the PCR primers contain anchoring sequences, which allow proper placement at the end of the 3' or 5' oligonucleotide end stretch of the cDNA.

19. Method according to any one of claims 1 to 18, wherein an amplification primer or a reverse transcription primer contains a promoter for RNA polymerases, preferably wherein the promoter enhances transcription by the T7 or T3 or SP6 RNA polymerases.

20. Method according to any one of claims 1 to 19, wherein the amplification primers and/or the amplification products are labeled by a reporter group, preferably fluorescence markers, chemiluminescence markers or radioisotopes, preferably wherein these reporter groups are used to detect the DNA.

21. Method according to any one of claims 1 to 20, wherein arbitrary nucleotides, preferably one to three nucleotides, are added to the 3' end of the RNA or mRNA posttranscriptionally but before the addition of a (e.g. poly A) tail, and the primers for reverse transcription and/or PCR contain a sequence with a wobble corresponding to the arbitrary sequence.

22. Method according to any one of the claims 1 to 21, wherein amplification products are generated that are differentially enlarged by defined numbers of nucleotides at their 5' and/or 3' end as compared to the RNA, by using different primers which have different lengths and recognize different 5' or 3' endsequences of the RNA or the full length cDNA.

23. Method according to any one of claims 1 to 22, wherein RNA or mRNA samples are used, which do not have a poly A tail or a tail sequence, and preferably a polynucleotide tail, especially preferred a poly A tail, is added synthetically prior to reverse transcription.

24. Method according to any one of claims 1 to 23, wherein the detection of differences between two cDNA pools, preferably corresponding pools between two samples, is done during or at the end of the amplification step, preferably by real time PCR.

25. Method according to any one of claims 1 to 24, wherein a 5' primer or a 3' primer or a 5'-3' primer pair, is spotted and/or bound onto a solid support, preferably the spot or the solid


IPR2025-01353; IPR2025-01355
CSHL EX2002

support is located in a defined 2-dimensional or 3-dimensional area and preferably such an area contains information to decode the identity of the primer or primer pair located in that area.

26. Method according to any one of claims 1 to 25, wherein the steps of reverse transcription and amplification are performed in one step.

27. Method according to any one of claims 1 to 26, wherein the polynucleotide tail added by the terminal transferase consists of ribonucleotides, preferably 2 to 6 ribonucleotides, preferably adenosine or thymidine, after a 3' sequence of oligo-dC, preferably with a length of 1 to 6 dC nucleotides, which is generated during reverse transcription, dependent on a 5' cap of a template RNA or mRNA, and the second amplification primer (5' primer) comprises complementary nucleotides to the ribonucleotides and oligo-dCs and 1 to 10 nucleotides upstream of the olio-dC sequence.

28. Method according to any one of claims 1 to 27, wherein a tail is added to the cDNA by the template switching ability of the reverse transcriptase, preferably in a cap dependent manner, and the second amplification primer (5' primer) comprises complementary nucleotides to the so added tail and oligo-dCs and 1 to 10 nucleotides upstream of the oligo-dC sequence.

29. Method for measuring the expression level of messenger RNA characterized by employing a method according to any one of claims 1 to 28, including the separation and/or identification of amplification results.

30. Kit for a method according to any one of claims 1 to 29 comprising a DNA polymerase, preferably Taq polymerase, a reverse transcriptase or mixtures of different polymerases, cofactors or salts of metal ions, preferably $Mg^{2+}$ and $Mn^{2+}$, requested by a polymerase, primers, preferably being selective for the 5' and/or 3' end of a RNA or cDNA, even more preferred being selective for the first, second, third, fourth and/or fifth terminal nucleotide of the RNA or cDNA, and optionally a terminal transferase and additional buffer substances.

31. A solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a 5' primer and a 3' primer or a pair of primers, wherein the primers are defined as in claim 1 point c), claim 2, claim 11, claim 12 or claim 17, in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or different primer pairs.

32. A solid surface, preferably a chip, in particular preferred a glass chip, comprising in defined areas, preferably in spots, a 5' primer or a 3' primer comprising the general formula as defined in claim 17 or a complement thereof, preferably selected from a primer comprising one of the following sequences: AGA GAT TTT TTT TTT TTT TT GA, AGA GAT TTT TTT TTT TTT T GG, AGA GAT TTT TTT TTT TTT T GC, AGA GAT TTT TTT TTT TTT TT GT, AGA GAT TTT TTT TTT TTT TT CA, AGA GAT TTT TTT TTT TTT TT AG, AGA GAT TTT TTT TTT TTT TTT AA, AAA AAA AAA AAA AAA GGG ATA, AAA AAA AAA AAA AAA GGG ACA, AAA AAA AAA AAA AAA GGG AGA, AAA AAA AAA AAA AAA GGG AAA, AAA AAA AAA AAA AAA GGG GTA, AAA AAA AAA AAA AAA GGG CTA, AAA AAA AAA AAA AAA GGG TTA, in particular preferred the solid surface comprises at least two, more preferred at least 9, especially preferred at least 20, most preferred at least 60, different areas with different primers or different primer pairs.

33. Method for amplifying a pool of polynucleotide molecules in a sample, characterized by the steps of:

a) obtaining a sample of RNA and reverse transcription of entire RNA molecules thus creating full length cDNA or obtaining a sample of full length cDNA,

b) tailing the 3' end of the transcribed cDNA with a polynucleotide tail after the 3' end,

c) amplification of the cDNA using a pair of primers, wherein a first (3') primer is specific for the 5' end of the cDNA and a second (5') primer is specific for the upstream portion of the polynucleotide tail and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides, upstream of

the 3' polynucleotide tail of the cDNA, preferably also comprising unspecific wobble nucleotides, especially preferred in the upstream portion within the 1 to 100 nucleotide range - wherein at least 1 nucleotide is specific for a nucleotide of the cDNA; the method further preferably defined as in any one of claims 1 to 29.

34. Method according to claim 1 or 33, characterized in that a primer, which is specific for a 3' tail of the RNA (optionally an added 3' tail) or mRNA (preferably the endogenous 3' poly A tail) and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides, upstream of the 3' tail is used for the reverse transcription and/or a first amplification primer is specific for a 5' tail of the cDNA, which is complementary to the 3' tail of the RNA or mRNA, and the next 1 to 100 nucleotides, preferably the next 1 to 50 nucleotides, especially preferred the next 1 to 20 nucleotides, in particular preferred the next 1 to 10 nucleotides, downstream of the 5' tail.



IPR2025-01353; IPR2025-01355
CSHL EX2002

SEQUENCE LISTING

<110> Seitz, Alexander

<120> Polynucleotide Amplifikation

<130> r48881

<150> AT A 1923/2005
<151> 2005-11-29

<160> 34

<170> PatentIn version 3.3

<210> 1
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<220>
<221> misc_feature
<222> (7)..(7)
<223> n is a, c, g, or t

<400> 1
acacacn 7

<210> 2
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 2
acacaca 7

<210> 3
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 3
acacacc 7

<210> 4
<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 4
acacact 7

<210> 5



IPR2025-01353; IPR2025-01355
CSHL EX2002

<211> 7
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 5
acacacg 7

<210> 6
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 6
gatatccgcc ctgtgct 17

<210> 7
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 7
ctgtgggcaa ggtctca 17

<210> 8
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 8
tgtccctttt ggtccca 17

<210> 9
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 9
ttgtggaagg gctcgca 17

<210> 10
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 10
cgatgccccc attcaca 17

<210> 11
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 11
caccaccatg gatctga 17

<210> 12
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 12
caacgggaag cctctaa 17

<210> 13
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 13
acgacccctt catctta 17

<210> 14
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 14
gatatgatgc gcaaagg 17

<210> 15
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 15
gacatggcct aagcaag 17

<210> 16
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 16


IPR2025-01353; IPR2025-01355
CSHL EX2002

gagatgagca cggaagc 17

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

<220>
<223> RNA-terminus

<400> 17
gaaatccgct ttgattt 17

<210> 18
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 18
ggtatcagct ccccatc 17

<210> 19
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 19
gctatgggcc tgggtca 17

<210> 20
<211> 17
<212> DNA
<213> Artificial

<220>
<223> RNA-terminus

<400> 20
gttatgggcg gagtgga 17

<210> 21
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 21
agagattttt ttttttttt ga 22

<210> 22
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer



IPR2025-01353; IPR2025-01355
CSHL EX2002

<400> 22
agagattttt ttttttttg g 21

<210> 23
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 23
agagattttt ttttttttg c 21

<210> 24
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 24
agagattttt tttttttttt gt 22

<210> 25
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 25
agagattttt tttttttttt ca 22

<210> 26
<211> 22
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 26
agagattttt tttttttttt ag 22

<210> 27
<211> 23
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 27
agagattttt tttttttttt taa 23

<210> 28
<211> 21
<212> DNA
<213> Artificial



IPR2025-01353; IPR2025-01355
CSHL EX2002

<220>
<223> primer

<400> 28
aaaaaaaaaa aaaaagggat a 21

<210> 29
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 29
aaaaaaaaaa aaaaagggac a 21

<210> 30
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 30
aaaaaaaaaa aaaaagggag a 21

<210> 31
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 31
aaaaaaaaaa aaaaagggaa a 21

<210> 32
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 32
aaaaaaaaaa aaaaaggggt a 21

<210> 33
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 33
aaaaaaaaaa aaaaagggct a 21

<210> 34
<211> 21
<212> DNA



IPR2025-01353; IPR2025-01355
CSHL EX2002

<213> Artificial

<220>
<223> primer

<400> 34
aaaaaaaaaa aaaaagggtt a 21



IPR2025-01353; IPR2025-01355
CSHL EX2002

# INTERNATIONAL SEARCH REPORT

International application No
PCT/AT2006/000494

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

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, BIOSIS, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 03/035841 A (SENTION INC [US]; SLEPNEV VLADMIR I [US]) 1 May 2003 (2003-05-01) Composition of claims 60-73 | 30 |
| A | US 5 962 272 A1 (CHENCHIK ALEX [US] ET AL) 5 October 1999 (1999-10-05) cited in the application the whole document | |
| A | DE 199 20 611 A1 (ROCHE DIAGNOSTICS GMBH [DE]) 9 November 2000 (2000-11-09) | |
| A | WO 02/065093 A (BAYLOR COLLEGE MEDICINE [US]; RES FOUND MENTAL HYGIENE [US]; GINSBERG) 22 August 2002 (2002-08-22) | |

-/--

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document 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 |
|---|---|
| 19 March 2007 | 03/04/2007 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL – 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | Cornelis, Karen |

Form PCT/ISA/210 (second sheet) (April 2005)



IPR2025-01353, IPR2025-01355
752 of 819
CSHL EX2002

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | EP 1 369 480 A (BIONEER CORP [KR]) 10 December 2003 (2003-12-10) figure 5 | |
| X | EP 1 186 673 A (AGILENT TECHNOLOGIES INC [US]) 13 March 2002 (2002-03-13) table 8 | 30-32 |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/AT2006/000494

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 03035841 A | 01-05-2003 | CA 2464691 A1 | 01-05-2003 |
| | | EP 1490511 A2 | 29-12-2004 |
| | | JP 2005507661 T | 24-03-2005 |
| | | US 2003077611 A1 | 24-04-2003 |
| | | US 2006105380 A1 | 18-05-2006 |
| US 5962272 A1 | | NONE | |
| DE 19920611 A1 | 09-11-2000 | EP 1072686 A2 | 31-01-2001 |
| | | JP 2000325079 A | 28-11-2000 |
| | | US 2003203389 A1 | 30-10-2003 |
| | | US 6558927 B1 | 06-05-2003 |
| WO 02065093 A | 22-08-2002 | CA 2437737 A1 | 22-08-2002 |
| | | EP 1366196 A2 | 03-12-2003 |
| | | JP 2004527236 T | 09-09-2004 |
| EP 1369480 A | 10-12-2003 | NONE | |
| EP 1186673 A | 13-03-2002 | US 2003065449 A1 | 03-04-2003 |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 24292881 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Joong Youn Cho |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 08-DEC-2015 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 13:06:52 |
| **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 | | 151208_5981_81503_A_SIDS_MW.PDF | 557823<br>e3b3dd9fd37320001dde859bec03c2b17375e389 | yes | 15 |

| | Multipart Description/PDF files in .zip description | | |
|---|---|---|---|
| | **Document Description** | **Start** | **End** |
| | Transmittal Letter | 1 | 12 |
| | Information Disclosure Statement (IDS) Form (SB08) | 13 | 15 |

| | | | | | |
|---|---|---|---|---|---|
| **Warnings:** | | | | | |
| The page size in the PDF is too large. The pages should be 8.5 x 11 or A4. If this PDF is submitted, the pages will be resized upon entry into the Image File Wrapper and may affect subsequent processing | | | | | |
| **Information:** | | | | | |
| 2 | Foreign Reference | 151208_5981_81503_A_SIDS_Exhibit_1_MW.pdf | 4734624<br>526d750c1c417080e6084a1993b3dec2604935db | no | 53 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_2_MW.pdf | 1041456<br>accc49476473f1c32d250ebcf5586ec28298bc80 | no | 12 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 4 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_3_MW.pdf | 2288558<br>9595596e3dfa1c738e9b75893e5f90b4314e269c | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 5 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_4_MW.pdf | 976119<br>78ebda1fdf464ab0b384258f4d5bfa7c23eff7cc | no | 8 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 6 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_5_MW.pdf | 741592<br>805357b57102447a7802f5bdfb6425ea42a3a9de | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 7 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_6_MW.pdf | 1039513<br>e6f7d4a893fe40adae27584dfa0b7acb91ac7751 | no | 6 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 8 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_7_MW.pdf | 1208253<br>a05964dfdaa54de0a26db4342f6ad9706e07e5a6 | no | 21 |
| **Warnings:** | | | | | |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| | | | | | |
|---|---|---|---|---|---|
| **Information:** | | | | | |
| 9 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_8_MW.pdf | 1010382<br>5f38cbdb82dc8537d0369ba86a8c283859e1d812 | no | 6 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 10 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_9_MW.pdf | 494165<br>401b9b41e04eacb6f1e120e7b1c9c0b5da73a799 | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 11 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_10_MW.pdf | 593509<br>807c74150a78c8b9b3cb9ed08b6062798ad9b1d4 | no | 7 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 12 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_11_MW.pdf | 6329200<br>1c996edd49d03d2946f4d6427cf798e58ccefcae | no | 9 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 13 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_12_MW.pdf | 1332636<br>dad1b02d7353d53aaf8a3253442ff1dda0a47e70 | no | 9 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 14 | Non Patent Literature | 151208_5981_81503_A_SIDS_Exhibit_13_MW.pdf | 1301635<br>b88a0b01a05bfd7eb2935f10fa6d2c4e255f1343 | no | 12 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 23649465 | | |

**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.**

Docket No. 5981/81503-A/GJG/MW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
December 8, 2015

**BY EFS**
Commissioner for Patents
Alexandria, VA

**COMMUNICATION REQUESTING INITIALED COPIES OF FORMS PTO-1449**

Certain Forms PTO-1449 filed with Information Disclosure Statements in connection with the subject application have not been returned with the Examiner's initials to applicants.

Specifically, applicants submitted an April 24, 2014 Supplemental Information Disclosure Statement including Form PTO-1449, and an October 21, 2011 Information Disclosure Statement including Form PTO-1449 to the United States Patent and Trademark Office in connection with the subject application. Neither has been returned to applicants to date initialed by the Examiner.

Applicants attach hereto as **Exhibit A** a copy of the Form PTO-1449 which was originally filed by applicants with the April 24, 2014 Supplemental Information Disclosure Statement.

Applicants also attach as **Exhibit B** a copy of the Form PTO-1449 which was originally filed by applicants with the October 21, 2011 Information Disclosure Statement. Additionally, a copy of the Electronic Acknowledgment Receipt confirming filing of the October 21, 2011 IDS Form PTO-1449 Substitute is attached hereto as **Exhibit C**.

Accordingly, applicants direct the Examiner's attention to the following previously submitted reference items **1-20** listed below which are also listed on the Form PTO-1449 attached hereto as **Exhibit D**.

Reference items **1-6** are a re-submission of previously submitted reference items **1-6** in the Supplemental Information Disclosure Statement originally filed on April 24, 2014.

Reference items **7-20** are a re-submission of previously submitted reference items **1-14** in the Information Disclosure Statement originally filed on October 21, 2011.

Reference items **1-3** and **7-10** are U.S. Patent Application Publications. Pursuant to 37 C.F.R. § 1.98(a)(2), copies of these reference items are not being submitted.

Copies of reference items **4-6** and **11-20** are attached hereto as **Exhibits 1-13**, respectfully.

1. U.S. Patent Application Publication No. 2004-110153, published June 10, 2004, Dong et al.;
2. U.S. Patent Application Publication No. 2007-172873, published July 26, 2007, Brenner et al.;
3. U.S. Patent Application Publication No. 2010-178655, published July 15, 2010, Hamilton et al.;
4. PCT International Patent Application Publication No. WO 2007/062445, published June 7, 2007, Alexander Seitz (**Exhibit 1**);
5. February 13, 2014 Extended European Search Report, issued in connection with European Patent Application No. 11835243.4 (**Exhibit 2**);
6. Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 (**Exhibit 3**);
7. U.S. Patent No. 5,639,603, issued June 17, 1997 (Dower et al.);

8. U.S. Patent No. 7,537,897, issued May 26, 2009 (Brenner et al.).
9. U.S. Patent No. 7,622,281, issued November 24, 2009 (Ronaghi et al.);
10. U.S. Patent Application Publication No. 2006/0073506, published April 6, 2006 (Christians et al.);
11. Casbon J.A., Osborne R.J., Brenner S., Lichtenstein C.P. A method for counting PCT template molecules with application to next-generation sequencing. Nucleic Acids Research Advance Access, published April 13, 2011 **(Exhibit 4)**;
12. Cheung V.G., Nelson S.F. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. PNAS, vol. 93:14676-14679, 1996 **(Exhibit 5)**;
13. Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," Science, vol. 323 pp. 133-138 (2009) **(Exhibit 6)**;
14. Supporting Online Material for Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," accessible at www.sciencemag.org/cgi/content/full/1162986/DC1, published November 20, 2008 **(Exhibit 7)**;
15. Fu G.K., Hu J., Wang P., Fodor S.P.A. Counting individual DNA molecules by the stochastic attachment of diverse labels. PNAS, vol. 108:9026-9031, 2011 **(Exhibit 8)**;
16. Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S.. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135 **(Exhibit 9)**;
17. McCloskey M.L., Stoger, R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767 **(Exhibit 10)**;
18. Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130 **(Exhibit 11)**;
19. Paunio T., Reima I., Syvanen A. Preimplanation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. Clinical Chemistry 42:9:1382-1390 (1996) **(Exhibit 12)**; and

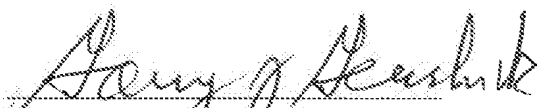
20. Technology Backgrounder: Single Molecule Real Time (SMRT™) DNA Sequencing. Pacific Biosciences 2009 **(Exhibit 13)**.

The Examiner is respectfully requested to make of record each item listed on the Form PTO-1449 Substitute by initialing and dating the attached Form PTO-1449 Substitute, and returning a copy of the initialed and dated form to applicants' undersigned attorney.

If a telephone interview would be of assistance in advancing prosecution of the subject application, applicants' undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this Information Disclosure Statement. However, if any additional fee is required authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

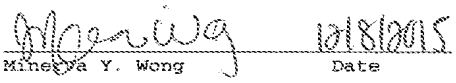
Respectfully submitted,



Gary J. Gershik
Registration No. 39,992
Attorney for Applicants
Cooper & Dunham, LLP
30 Rockefeller Plaza
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.



Minerva Y. Wong 10/18/2015
Date

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/REB |

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 |
|---|---|---|---|---|
| | 1 | 2004-110153 | 06-10-2004 | Dong et al. |
| | 2 | 2007-172873 | 07-26-2007 | Brenner et al. |
| | 3 | 2010-178655 | 07-15-2010 | Hamilton et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

FOREIGN PATENT DOCUMENTS

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | 4 | WO 2007/062445 | 06-07-2007 | Alexander Seitz | |
| | | | | | |
| | | | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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 of 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

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Exhibit A

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/REB |

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] |
|---|---|---|---|
| | 5 | February 13, 2014 Extended European Search Report, issued in connection with European Patent Application No. 11835243.4 | |
| | 6 | Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | Not Yet Known |
|---|---|---|
| | Filing Date | Herewith |
| | First Named Inventor | |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | Art Unit | |
| | Examiner Name | |
| | Attorney Docket No. | 5981/81503-A/GJG/ACK |

**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 |
|---|---|---|---|---|
| | 1 | 5,639,603 | 06-17-1997 | Dower et al. |
| | 2 | 7,537,897 | 05-26-2009 | Brenner et al. |
| | 3 | 7,622,281 | 11-24-2009 | Ronaghi et al. |
| | 4 | 2006/0073506 | 04-06-2006 | Christians et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |

EXAMINER SIGNATURE | DATE CONSIDERED

*EXAMINER: Initial if citation 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 of 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.

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Exhibit B

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | Not Yet Known |
|---|---|---|
| | Filing Date | Herewith |
| | First Named Inventor | |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | Art Unit | |
| | Examiner Name | |
| | Attorney Docket No. | 5981/81503-A/GJG/ACK |

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] |
|---|---|---|---|
| | 5 | Casbon J.A., Osborne R.J., Brenner S., Lichtenstein C.P. A method for counting PCR template molecules with application to next-generation sequencing. Nucleic Acids Research Advance Access, published April 13, 2011 | |
| | 6 | Cheung V.G., Nelson S.F. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. PNAS, vol. 93:14676-14679, 1996 | |
| | 7 | Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," Science, vol. 323 pp. 133-138 (2009) | |
| | 8 | Supporting Online Material for Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," accessible at www.sciencemag.org/cgi/content/full/1162986/DC1, published November 20, 2008 | |
| | 9 | Fu G.K., Hu J., Wang P., Fodor S.P.A. Counting individual DNA molecules by the stochastic attachment of diverse labels. PNAS, vol. 108:9026-9031, 2011 | |
| | 10 | Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135 | |
| | 11 | McCloskey M.L., Stoger, R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767 | |
| | 12 | Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130 | |
| | 13 | Paulo T., Reima I., Syvanen A. Preimplanation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. Clinical Chemistry 42:9:1382-1390 (1996) | |
| | 14 | Technology Backgrounder: Single Molecule Real Time (SMRT™) DNA Sequencing. Pacific Biosciences 2009 | |

| EXAMINER SIGNATURE | DATE CONSIDERED: |
|---|---|

*EXAMINER: Initial if citation 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.

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| Applicants | : | James Hicks et al. | |
| U.S. Serial No. | : | 13/278,333 | Examiner: Frank W. M. Lu |
| Filing Date | : | October 21, 2011 | Art Unit: 1634 |
| Confirmation No. | : | 4469 | |
| For | : | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION | |

30 Rockefeller Plaza
20th Floor
New York, New York 10112
December 10, 2015

BY EFS
Commissioner for Patents
Alexandria, VA

**AMENDMENT UNDER 37 C.F.R. § 1.312 IN RESPONSE TO DECEMBER 2, 2015 NOTICE OF ALLOWANCE**

This Amendment is submitted pursuant to 37 C.F.R. § 1.312, after the mailing of a Notice of Allowance but before payment of the issue fee in connection with the above-identified application. The issue fee is due March 2, 2016 and has not yet been paid. Accordingly, this Amendment is being timely filed.

This Amendment corrects a typographical error in claim 7, and therefore should be entered under 37 C.F.R. § 1.312.

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

**Remarks** begin on page **10** of this paper.



IPR2025-01353; IPR2025-01355
CSHL EX2002

**Amendments to the Claims:**

Pursuant to 37 C.F.R. § 1.121(c), this listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims**

1. (Previously Presented) A method for obtaining genomic copy number information from genomic nucleic acids, comprising:

   a) obtaining segments of the genomic nucleic acids;

   b) tagging the segments with substantially unique tags, thereby generating tagged nucleic acid molecules, such that each of the tagged nucleic acid molecules comprises one segment of the segments of the genomic nucleic acids from step (a) and a substantially unique tag having a unique nucleotide sequence of the substantially unique tags, so as to permit one substantially unique tag having a unique nucleotide sequence of the substantially unique tags to differentiate from another tag having a unique nucleotide sequence of the substantially unique tags;

   c) subjecting the tagged nucleic acid molecules to a polymerase chain reaction (PCR), thereby generating amplified tagged nucleic acid molecules;

   d) generating tag associated sequence reads by sequencing the amplified tagged nucleic acid molecules of step (c);

   e) assigning each of the amplified tagged nucleic acid molecules to a location on a genome associated with the genomic nucleic acids by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (d) corresponding to one segment of the segments of the genomic nucleic acids to a location on the genome; and


IPR2025-01353; IPR2025-01355
CSHL EX2002

f) in the amplified tagged nucleic acid molecules, counting the number of the amplified tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location on the genome,

thereby obtaining counts of the amplified tagged nucleic acid molecules having a different substantially unique tag that have been assigned to the same location of the genome and the genomic copy number information from the genomic nucleic acids, wherein the different substantially unique tag is from the substantially unique tags, and wherein the genome has multiple locations.

2. (Previously Presented) The method of claim 1, further comprising estimating a genomic copy number of a region of the genome comprising multiple locations.

3. (Previously Presented) The method of claim 1, further comprising comparing a count from the counts obtained in step (f) for a location on the genome to a count for the same location obtained from a reference sample, thereby estimating a relative genomic copy number of the location.

4. (Previously Presented) The method of claim 1, further comprising

g) summing the counts from the counts obtained in step (f) for locations on the genome which comprise a first region of the genome, wherein the first region comprises more than one location;

h) summing the counts from the counts obtained in step (f) for locations on the genome which comprise a second region of the genome, wherein the second region comprises a number of locations which is comparable to the number of locations of the first region;

i) comparing the value obtained in step (g) to the value


IPR2025-01353; IPR2025-01355
CSHL EX2002

obtained in step (h),

thereby estimating the relative genomic copy number of the first region of the genome to the genomic copy number of the second region of the genome.

5. (Previously Presented) The method of claim 4, wherein step (h) further comprises

j) summing the counts from the counts obtained in step (f) for locations on the genome which comprise a third region of the genome, wherein the third region comprises a number of locations which is comparable to the number of locations of the first region; and

k) obtaining an average of the sum of the counts obtained in step (h) and the counts obtained in step (j).

6. (Previously Presented) The method of claim 4, wherein the second region of the genome comprises a centromere.

7. (Currently Amended) The method of claim 1, further comprising l) summing the counts from the counts obtained in step (f) for locations which comprise a region of the genome; m) summing the counts obtained from a reference sample for the same region of the genome; and [[iii) ]]comparing the sums of the counts obtained in step (l) and step (m), thereby estimating a relative genomic copy number of the region of the genome.

8. (Cancelled)

9. (Previously Presented) A method for obtaining mRNA copy number information from mRNA transcripts, comprising:

a) generating tagged nucleic acid molecules, comprising:

i) subjecting the mRNA transcripts to a polymerase reaction in the presence of primers capable of hybridizing to the polyA tail of the mRNA

transcripts under conditions that promote the formation of only one complement of the mRNA transcripts, thereby generating first order derivative strands;

ii) adding a polynucleotide tail to an end of the first order derivative strands, thereby generating first order derivative strands comprising the polynucleotide tail; and

iii) subjecting the first order derivative strands comprising the polynucleotide tail to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step (ii) under conditions that promote the formation of only one complement of the first order derivative strands comprising the polynucleotide tail, thereby generating second order derivative strands,

wherein the primers of at least one of steps (i) and (iii) comprise substantially unique tags, and the first order derivative strands and the second order derivative strands are the tagged nucleic acid molecules;

b) subjecting each of the tagged nucleic acid molecules to a polymerase chain reaction (PCR), thereby generating amplified tagged nucleic acid molecules;

c) generating tag associated sequence reads by sequencing the amplified tagged nucleic acid molecules of step (b);

d) assigning each of the amplified tagged nucleic acid molecules to a location on a cDNA library associated with the mRNA transcripts by mapping the subsequence of each tag associated sequence read of said tag associated sequence reads of step (c) corresponding to a mRNA

transcript of the mRNA transcripts to a location on the cDNA library; and

e) in the amplified tagged nucleic acid molecules, counting the number of the amplified tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library,

thereby obtaining counts of the amplified tagged nucleic acid molecules having a different tag that have been assigned to the same location on the cDNA library and mRNA copy number information from the mRNA transcripts, wherein the different tag is from the substantially unique tags, wherein the cDNA library comprises multiple locations.

10. (Cancelled)

11. (Previously Presented) The method of claim 1, wherein said tagging the segments comprises:

n) adding a polynucleotide tail to the ends of the segments of the genomic nucleic acids, thereby generating zero-th order derivative strands comprising the polynucleotide tail;

o) subjecting the zero-th order derivative strands of step (n) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands under conditions that promote the formation of only one complement of the zero-th order derivative strands, thereby generating first order derivative strands;

p) adding a polynucleotide tail to the first order derivative strands, thereby generating a first order derivative strands comprising the polynucleotide tail;

q) subjecting the first order derivative strands comprising the polynucleotide tail to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the first order derivative strands under conditions that promote the formation of only one complement of the first order derivative strands comprising the polynucleotide tail, thereby generating second order derivative strands,

wherein the primers of at least one of steps o) and q) comprise the substantially unique tags, and the first order derivative strands and the second order derivative strands are the tagged nucleic acid molecules.

12. (Previously Presented) The method of claim 1, wherein said tagging the segments comprises:

r) adding a polynucleotide tail to the ends of the segments of genomic nucleic acids, thereby generating zero-th order derivative strands comprising the polynucleotide tail,

s) subjecting the zero-th order derivative strands of step (r) to a ligation reaction in the presence of primers capable of hybridizing to the polynucleotide tail of the zero-th order strands added in step (r) under conditions that promote the ligation of a primer to the 5′ ends of the zero-th order derivative strands,

t) subjecting the product of step (s) to a polymerase reaction in the presence of primers capable of hybridizing to the polynucleotide tail added in step r) under conditions which promote the formation of only one complement of the product of step (s), wherein the primers of step (s) and the primers of step (t) have different nucleotide sequences, and wherein the polymerase has 3′-5′ proofreading activity,

wherein the primers of at least one of steps (s) and (t) comprise substantially unique tags, thereby generating tagged nucleic acid molecules.

13. (Previously Presented) The method of claim 12, wherein said adding the polynucleotide tail comprises the use of a terminal transferase.

14-23. (Cancelled)

24. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to the PCR or the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to step (d).

25. (Cancelled)

26. (Previously Presented) The method of claim 1, wherein the substantially unique tags further comprise a sample tag.

27. (Previously Presented) The method of claim 26, wherein the tagged nucleic acid molecules are pooled with a plurality of tagged nucleic acid molecules having a different sample tag prior to the PCR or prior to said sequencing of the product of step (c).

28. (Previously Presented) The method of claim 27, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads.

29-30. (Cancelled)

31. (Previously Presented) The method of claim 1, wherein the tagged nucleic acid molecules are generated from a single cell.

32-67. (Cancelled)


IPR2025-01353; IPR2025-01355
CSHL EX2002

68. (Previously Presented) The method of claim 24, wherein the tagged nucleic acid molecules are subject to hybrid capture prior to the PCR in step (c).

69. (Previously Presented) The method of claim 68, wherein each of the tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on an array.

70. (Previously Presented) The method of claim 24, wherein the amplified tagged nucleic acid molecules produced in step (c) are subject to hybrid capture prior to said sequencing the amplified tagged nucleic acid molecules in step (d).

71. (Previously Presented) The method of claim 70, wherein each of the amplified tagged nucleic acid molecules subject to hybrid capture has a nucleic acid sequence which is complementary to a capture probe on a microarray.

**Remarks**

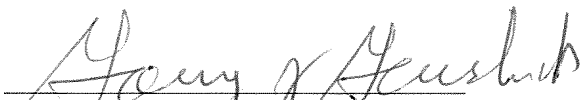
Claims 1-7, 9, 11-13, 24, 26-28, 31 and 68-71 are indicated as allowed in the December 2, 2015 Notice of Allowance. In this Amendment under 37 C.F.R. § 1.312, submitted prior to paying the issue fee, applicants amended claim 7 to correct a minor typographical error.

Applicants maintain that this Amendment under 37 C.F.R. § 1.312 raises no issue of new matter and respectfully request the entry of this Amendment under 37 C.F.R. § 1.312.

If a telephone interview would be of assistance in advancing consideration of the Amendment of the subject application under 37 C.F.R. § 1.312, applicants' undersigned attorney invites the Examiner to telephone him at the number provided below.

No fee is deemed necessary in connection with the filing of this Amendment. However, if any additional fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

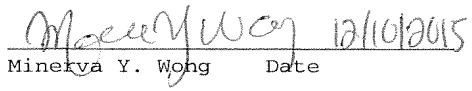
Respectfully submitted,



Gary J. Gershik
Registration No. 39.992
Attorney for Applicants
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.



Minerva Y. Wong Date



IPR2025-01353; IPR2025-01355
CSHL EX2002

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 24322473 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Joong Youn Cho |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 10-DEC-2015 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 14:54:04 |
| **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 | | 151210_5981_81503_A_Amendment_MW.pdf | 291998<br>096302c577acd589ebef36d21b1c6243c3613881 | yes | 10 |

| | Multipart Description/PDF files in .zip description | | |
|---|---|---|---|
| | **Document Description** | **Start** | **End** |
| | Amendment after Notice of Allowance (Rule 312) | 1 | 1 |
| | Claims | 2 | 9 |
| | Applicant Arguments/Remarks Made in an Amendment | 10 | 10 |
| **Warnings:** | | | |
| **Information:** | | | |
| | **Total Files Size (in bytes):** | 291998 | |

**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.**

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office<br><br>INFORMATION DISCLOSURE STATEMENT<br>(Use several sheets if necessary) | | |
|---|---|---|
| | Application Number | 13/278,333 |
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks et al. |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/MW |

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 |
|---|---|---|---|---|
| | 1 | 2004-110153 | 06-10-2004 | Dong et al. |
| | 2 | 2007-172873 | 07-26-2007 | Brenner et al. |
| | 3 | 2010-178655 | 07-15-2010 | Hamilton et al. |

FOREIGN PATENT DOCUMENTS

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | 4 | WO 2007/062445 | 06-07-2007 | Alexander Seitz | |

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] |
|---|---|---|---|
| | 5 | February 13, 2014 Extended European Search Report, issued in connection with European Patent Application No. 11835243.4 | |
| | 6 | Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 | |

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 |
|---|---|---|---|---|
| | 7 | 5,639,603 | 06-17-1997 | Dower et al. |
| | 8 | 7,537,897 | 05-26-2009 | Brenner et al. |
| | 9 | 7,622,281 | 11-24-2009 | Ronaghi et al. |
| | 10 | 2006/0073506 | 04-06-2006 | Christians et al. |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 12/11/2015 |
|---|---|---|---|

*EXAMINER: Initial if citation 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.

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Exhibit D

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | First Named Inventor | James Hicks et al. |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/MW |

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] |
|---|---|---|---|
| | 11 | Casbon J.A., Osborne R.J., Brenner S., Lichtenstein C.P. A method for counting PCT template molecules with application to next-generation sequencing. Nucleic Acids Research Advance Access, published April 13, 2011 | |
| | 12 | Cheung V.G., Nelson S.F. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. PNAS, vol. 93:14676-14679, 1996 | |
| | 13 | Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," Science, vol. 323 pp. 133-138 (2009) | |
| | 14 | Supporting Online Material for Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," accessible at www.sciencemag.org/cgi/content/full/1162986/DC1, published November 20, 2008 | |
| | 15 | Fu G.K., Hu J., Wang P., Fodor S.P.A. Counting individual DNA molecules by the stochastic attachment of diverse labels. PNAS, vol. 108:9026-9031, 2011 | |
| | 16 | Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S.. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135 | |
| | 17 | McCloskey M.L., Stoger, R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767 | |
| | 18 | Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130 | |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 12/11/2015 |
|---|---|---|---|

*EXAMINER: Initial if citation 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office<br><br>INFORMATION DISCLOSURE STATEMENT<br>(Use several sheets if necessary) | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks et al. |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/MW |

**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] |
|---|---|---|---|
| | 19 | Paunio T., Reima I., Syvanen A. Preimplanation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. Clinical Chemistry 42:9:1382-1390 (1996) | |
| | 20 | Technology Backgrounder: Single Molecule Real Time (SMRT™) DNA Sequencing. Pacific Biosciences 2009 | |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 12/11/2015 |
|---|---|---|---|

*EXAMINER: Initial if citation 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.

Duplicated documents



**Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office**

**INFORMATION DISCLOSURE STATEMENT**
**(Use several sheets if necessary)**

| | |
|---|---|
| Application Number | 13/278,333 |
| Filing Date | October 21, 201 |
| First Named Inventor | James Hicks |
| Art Unit | 1634 |
| Examiner Name | Frank W. M. Lu |
| Attorney Docket No. | 5981/81503 A/GJG/REB |

**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 |
|---|---|---|---|---|
| | 1 | 2004-110153 | 06-10-2004 | Dong et al. |
| | 2 | 2007-172873 | 07-26-2007 | Brenner et al. |
| | 3 | 2010-178655 | 07-15-2010 | Hamilton et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | 4 | WO 2007/062445 | 06-07-2007 | Alexander Seitz | |
| | | | | | |
| | | | | | |

EXAMINER SIGNATURE

DATE CONSIDERED

*EXAMINER: Initial if citation 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 of 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | Application Number | 13/278,333 |
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks |
| | Art Unit | 1634 |
| | Examiner Name | Frank W. Mc Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/RWB |

**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] |
|---|---|---|---|
| | 5 | February 13, 2014 Extended European Search Report issued in connection with European Patent Application No. 11835243.4 | |
| | 6 | Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

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 NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

23432 7590 12/15/2015

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

| EXAMINER |
|---|
| LU, FRANK WEI MIN |

| ART UNIT | PAPER NUMBER |
|---|---|
| 1634 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 12/15/2015 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Applicant-Initiated Interview Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 13/278,333 | HICKS ET AL. | |
| | Examiner | Art Unit | |
| | FRANK LU | 1634 | |

All participants (applicant, applicant's representative, PTO personnel):

(1) *FRANK LU*. (3)_____.

(2) *Mr. Gary GerShik (Reg. No. 39,992)*. (4)_____.

Date of Interview: *28 July 2015*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☐112 ☐102 ☒103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *1-7, 9, 11-13, 24, 26-28, 31, and 68-71* .

Identification of prior art discussed: *US Patent No. 8,741,606 B2*.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

*After reviwing the office action mailed on 6/11/2015, Mr. Gershik told the examiner that he wanted to remove the phrase "amplification distortion" from the claims. The examiner indicated that US Patent No. 8,741,606 B2 could be used as a potential reference against the claims pending in the application if the phrase "amplification distortion" is removed from the claims. Mr. Gershik will file a declaration under 37 C.F.R. 1.131*.

**Applicant recordation instructions:** The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

**Examiner recordation instructions**: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

| /FRANK LU/<br>Primary Examiner, Art Unit 1634 | July 28, 2015 |
|---|---|


IPR2025-01353, IPR2025-01355
CSHL EX2002

# Summary of Record of Interview Requirements

**Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record**
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

**Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews**
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

---

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

**Examiner to Check for Accuracy**

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

| Response to Rule 312 Communication | Application No. | Applicant(s) |
|---|---|---|
| | 13/278,333 | HICKS ET AL. |
| | Examiner | Art Unit |
| | FRANK LU | 1634 |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address –*

1. ☒ The amendment filed on *10 December 2015* under 37 CFR 1.312 has been considered, and has been:

a) ☒ entered.

b) ☐ entered as directed to matters of form not affecting the scope of the invention.

c) ☐ disapproved because the amendment was filed after the payment of the issue fee.

Any amendment filed after the date the issue fee is paid must be accompanied by a petition under 37 CFR 1.313(c)(1) and the required fee to withdraw the application from issue.

d) ☐ disapproved. See explanation below.

e) ☐ entered in part. See explanation below.

Note that IDS fiiled on 10/21/2011 suggested by applicant cannot be considered as IDS since applicant has not put these references on the form-1449 and no reference listed in the IDS has been submitted in that time.

| | |
|---|---|
| Please attach signed form-1449 and an interview summary in this office action. | /FRANK LU/<br>Primary Examiner, Art Unit 1634<br>December 11, 2015 |

| Applicant-Initiated Interview Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 13/278,333 | HICKS ET AL. | |
| | Examiner | Art Unit | |
| | FRANK LU | 1634 | |

All participants (applicant, applicant's representative, PTO personnel):

(1) *FRANK LU*. (3)_____.

(2) *Mr. Gary GerShik (Reg. No. 39,992)*. (4)_____.

Date of Interview: *28 July 2015*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☐112 ☐102 ☒103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *1-7, 9, 11-13, 24, 26-28, 31, and 68-71* .

Identification of prior art discussed: *US Patent No. 8,741,606 B2*.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

*After reviwing the office action mailed on 6/11/2015, Mr. Gershik told the examiner that he wanted to remove the phrase "amplification distortion" from the claims. The examiner indicated that US Patent No. 8,741,606 B2 could be used as a potential reference against the claims pending in the application if the phrase "amplification distortion" is removed from the claims. Mr. Gershik will file a declaration under 37 C.F.R. 1.131*.

**Applicant recordation instructions:** The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

**Examiner recordation instructions**: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

| /FRANK LU/<br>Primary Examiner, Art Unit 1634 | July 28, 2015 |
|---|---|


IPR2025-01353, IPR2025-01355
CSHL EX2002

# Summary of Record of Interview Requirements

**Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record**
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

**Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews**
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

---

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:

1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

**Examiner to Check for Accuracy**

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

duplicated documents



| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503.2 A/GJG/REB |

**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 |
|---|---|---|---|---|
| | 1 | 2004-110153 | 06-10-2004 | Dong et al. |
| | 2 | 2007-172873 | 07-26-2007 | Brenner et al. |
| | 3 | 2010-178655 | 07-15-2010 | Hamilton et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**FOREIGN PATENT DOCUMENTS**

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | 4 | WO 2007/062445 | 06-07-2007 | Alexander Seitz | |
| | | | | | |
| | | | | | |

EXAMINER SIGNATURE

DATE CONSIDERED

*EXAMINER: Initial if citation 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 of 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.

Applicant(s): James Hicks, et al.
U.S. Serial No.: 13/278,333
Filed: October 21, 2011
Exhibit A

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | Art Unit | 1634 |
| | Examiner Name | Frank W. M Lu |
| | Attorney Docket No. | 9961/81503-A/GJG/REB |

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] |
|---|---|---|---|
| | 5 | February 13, 2014 Extended European Search Report, issued in connection with European Patent Application No. 11835243.4 | |
| | 6 | Miner et al. (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, Vol. 32, No. 17 e135, pages 1-4 | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

Form PTO-1449 (Substitute) U.S. Department of Commerce
Patent and Trademark Office

| | |
|---|---|
| Application Number | Not Yet Known |
| Filing Date | Herewith |
| First Named Inventor | |
| Art Unit | |
| Examiner Name | |
| Attorney Docket No. | 5981/01503-A/GJG/ACK |

# INFORMATION DISCLOSURE STATEMENT
(Use several sheets if necessary)

## 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 |
|---|---|---|---|---|
| | 1 | 5,639,603 | 06-17-1997 | Dower et al. |
| | 2 | 7,537,897 | 05-26-2009 | Brenner et al. |
| | 3 | 7,622,281 | 11-24-2009 | Ronaghi et al. |
| | 4 | 2006/0073506 | 04-06-2006 | Christians et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

## FOREIGN PATENT DOCUMENTS

| Examiner Initials* | Cite No.[1] | Foreign Patent Document Country Code[3] Number[4] Kind Code[5] (if known) | Publication Date MM-DD-YYYY | Name of Patentee or Applicant of Cited Document | T[6] |
|---|---|---|---|---|---|
| | | | | | |

EXAMINER SIGNATURE

DATE CONSIDERED

*EXAMINER: Initial if citation 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 of 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.

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | Not Yet Known |
|---|---|---|
| | Filing Date | Herewith |
| | First Named Inventor | |
| INFORMATION DISCLOSURE STATEMENT (Use several sheets if necessary) | Art Unit | |
| | Examiner Name | |
| | Attorney Docket No. | 5981/01503-A/GJG/ACK |

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] |
|---|---|---|---|
| | 5 | Casbon J.A., Osborne R.J., Brenner S., Lichtenstein C.P. A method for counting PCR template molecules with application to next-generation sequencing. Nucleic Acids Research Advance Access, published April 13, 2011 | |
| | 6 | Cheung V.G., Nelson S.F. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. PNAS, vol. 93:14676-14679, 1996 | |
| | 7 | Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," Science, vol. 323 pp. 133-138 (2009) | |
| | 8 | Supporting Online Material for Eid et al., "Real-Time DNA Sequencing from Single Polymerase Molecules," accessible at www.sciencemag.org/cgi/content/full/1162986/DC1, published November 20, 2008 | |
| | 9 | Fu G.K., Hu J., Wang P., Fodor S.P.A. Counting individual DNA molecules by the stochastic attachment of diverse labels. PNAS, vol. 108:9026-9031, 2011 | |
| | 10 | Miner B.E., Stoger, R.J., Burden, A.F., Laird, C.D., Hansen R.S. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Research, 2004:32(17):e135 | |
| | 11 | McCloskey M.L., Stoger R., Hansen, R.S., Laird, C.D. Encoding PCR Products with Batch-stamps and Barcodes. Biochem. Genet. 2007:45:761-767 | |
| | 12 | Parameswaran et al. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Research. 2007; 35(19): e130 | |
| | 13 | Paunio T., Reima I., Syvanen A. Preimplanation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. Clinical Chemistry 42:9:1382-1390 (1996) | |
| | 14 | Technology Backgrounder: Single Molecule Real Time (SMRT) DNA Sequencing. Pacific Biosciences 2009 | |

| EXAMINER SIGNATURE | DATE CONSIDERED |
|---|---|

*EXAMINER: Initial if citation 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.

# PART B - FEE(S) TRANSMITTAL

**Complete and send this form, together with applicable fee(s), to: Mail** Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
**or Fax** (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

23392 3590 3/02/2016
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

OPAP AKIMA-18 MAR 0 2 2016 PATENT & TRADEMARK OFFICE

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

**Certificate of Mailing or Transmission**
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

| |
|---|
| (Depositor's name) |
| (Signature) |
| (Date) |

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 3981/81503-A/GJG/GC | 4469 |

TITLE OF INVENTION: VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

| APPN. TYPE | ENTITY STATUS | ISSUE FEE DUE | PUBLICATION FEE DUE | PREV. PAID ISSUE FEE | TOTAL FEES DUE | DATE DUE |
|---|---|---|---|---|---|---|
| nonprovisional | SMALL | $480 | $0 | $0 | $480 | 03/02/2016 |

| EXAMINER | ART UNIT | CLASS-SUBCLASS |
|---|---|---|
| LU, FRANK WEI MIN | 1634 | 435-006110 |

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. Use of a Customer Number is required.

2. For printing on the patent front page, list:
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 Gary J. Gershik
2 Cooper & Dunham LLP
3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)
PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: Cold Spring Harbor Laboratory
(B) RESIDENCE: (CITY and STATE OR COUNTRY): Cold Spring Harbor, New York

Please check the appropriate assignee category or categories (will not be printed on the patent): ☐ Individual ☑ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
- ☑ Issue Fee
- ☐ Publication Fee (No small entity discount permitted)
- ☐ Advance Order - # of Copies .....

4b. Payment of Fee(s): (Please first reapply any previously paid issue fee shown above)
- ☐ A check is enclosed.
- ☐ Payment by credit card. Form PTO-2038 is attached.
- ☐ The director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number 03-3125 (enclose an extra copy of this form).

5. Change in Entity Status (from status indicated above)
- ☐ Applicant certifying micro entity status. See 37 CFR 1.29
- ☐ Applicant asserting small entity status. See 37 CFR 1.27
- ☐ Applicant changing to regular undiscounted fee status.

NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature: Gary J. Gershik Date: March 2, 2016
Typed or printed name: Gary J. Gershik Registration No.: 39,992

03/03/2016 INTEFSW 00002404 13278333
01 FC:2501 480.00 DA



PTOL-85 Part B (10-13) Approved for use through 10/31/2013 OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

# PART B - FEE(S) TRANSMITTAL

**Complete and send this form, together with applicable fee(s), to: Mail** Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
**or Fax** (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

23432 7590 12/02/2015

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

**Certificate of Mailing or Transmission**
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

TITLE OF INVENTION: VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

| APPLN. TYPE | ENTITY STATUS | ISSUE FEE DUE | PUBLICATION FEE DUE | PREV. PAID ISSUE FEE | TOTAL FEES DUE | DATE DUE |
|---|---|---|---|---|---|---|
| nonprovisional | SMALL | $480 | $0 | $0 | $480 | 03/02/2016 |

| EXAMINER | ART UNIT | CLASS-SUBCLASS |
|---|---|---|
| LU, FRANK WEI MIN | 1634 | 435-006110 |

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 Gary J. Gershik
2 Cooper & Dunham LLP
3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: Cold Spring Harbor Laboratory

(B) RESIDENCE: (CITY and STATE OR COUNTRY): Cold Spring Harbor, New York

Please check the appropriate assignee category or categories (will not be printed on the patent): ☐ Individual ☑ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
- ☑ Issue Fee
- ☐ Publication Fee (No small entity discount permitted)
- ☐ Advance Order - # of Copies

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ☐ A check is enclosed.
- ☐ Payment by credit card. Form PTO-2038 is attached.
- ☑ The director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number 03-3125 (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ☐ Applicant certifying micro entity status. See 37 CFR 1.29
  NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
- ☐ Applicant asserting small entity status. See 37 CFR 1.27
  NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
- ☐ Applicant changing to regular undiscounted fee status.
  NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature: Gary J. Gershik
Typed or printed name: Gary J. Gershik
Date: March 2, 2016
Registration No.: 39,992



PTOL-85 Part B (10-13) Approved for use through 10/31/2013. OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 13278333 |
| **Filing Date:** | 21-Oct-2011 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | Gary J. Gershik/Susan Chiarello |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |

Filed as Small Entity

**Filing Fees for Utility under 35 USC 111(a)**

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| Utility Appl Issue Fee | 2501 | 1 | 480 | 480 |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Extension-of-Time:** | | | | |
| **Miscellaneous:** | | | | |
| | | | **Total in USD ($)** | **480** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 25078992 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Susan Chiarello |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 02-MAR-2016 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 15:15:04 |
| **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 | Issue Fee Payment (PTO-85B) | 160302_5981_81503_A_Isse_Fee_Transmittal_GJG.pdf | 330503<br>8dccb69bc1591aa23f71e3b85468d0e68bc75ea4 | no | 1 |

| | |
|---|---|
| **Warnings:** | |
| **Information:** | |



IPR2025-01353; IPR2025-01355
CSHL EX2002

| | | | | | |
|---|---|---|---|---|---|
| 2 | Fee Worksheet (SB06) | fee-info.pdf | 30632<br>86902ac8eeb16d8d6b86cc5ad54c63f63722dcaf | no | 2 |

| | |
|---|---|
| **Warnings:** | |
| **Information:** | |
| **Total Files Size (in bytes):** | 361135 |

**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.**

# PART B - FEE(S) TRANSMITTAL

**Complete and send this form, together with applicable fee(s), to: Mail** Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
**or Fax** (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

23432 3590 12/02/2015
COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

**Certificate of Mailing or Transmission**
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 10/21/2011 | James Hicks | 5981/81503-A/GJG/GC | 4469 |

TITLE OF INVENTION: VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

| APPLN. TYPE | ENTITY STATUS | ISSUE FEE DUE | PUBLICATION FEE DUE | PREV. PAID ISSUE FEE | TOTAL FEES DUE | DATE DUE |
|---|---|---|---|---|---|---|
| nonprovisional | SMALL | $480 | $0 | $0 | $480 | 03/02/2016 |

| EXAMINER | ART UNIT | CLASS-SUBCLASS |
|---|---|---|
| LU, FRANK WEI MIN | 1634 | 435-006110 |

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 Gary J. Gershik
2 Cooper & Dunham LLP
3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: Cold Spring Harbor Laboratory

(B) RESIDENCE: (CITY and STATE OR COUNTRY): Cold Spring Harbor, New York

Please check the appropriate assignee category or categories (will not be printed on the patent): ☐ Individual ☑ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
- ☑ Issue Fee
- ☐ Publication Fee (No small entity discount permitted)
- ☐ Advance Order - # of Copies

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ☐ A check is enclosed.
- ☐ Payment by credit card. Form PTO-2038 is attached.
- ☑ The director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number 03-3125 (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ☐ Applicant certifying micro entity status. See 37 CFR 1.29
  NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
- ☐ Applicant asserting small entity status. See 37 CFR 1.27
  NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
- ☐ Applicant changing to regular undiscounted fee status.
  NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature: Gary J. Gershik
Typed or printed name: Gary J. Gershik
Date: March 2, 2016
Registration No.: 39,992





IPR2025-01353; IPR2025-01355
800 of 819
CSHL EX2002

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 13278333 |
| **Filing Date:** | 21-Oct-2011 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | Gary J. Gershik/Susan Chiarello |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |

Filed as Small Entity

**Filing Fees for Utility under 35 USC 111(a)**

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| Utility Appl Issue Fee | 2501 | 1 | 480 | 480 |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Extension-of-Time:** | | | | |
| **Miscellaneous:** | | | | |
| | | | **Total in USD ($)** | **480** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 25080781 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Susan Chiarello |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 02-MAR-2016 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 16:10:28 |
| **Application Type:** | Utility under 35 USC 111(a) |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | Deposit Account |
| Payment was successfully received in RAM | $480 |
| RAM confirmation Number | 2404 |
| Deposit Account | 033125 |
| Authorized User | GERSHIK, GARY J |

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



IPR2025-01353; IPR2025-01355
CSHL EX2002

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Issue Fee Payment (PTO-85B) | 160302_5981_81503_A_Isse_Fee_Transmittal_GJG.pdf | 330503<br>8dccb69bc1591aa23f71e3b85468d0e68bc75ea4 | no | 1 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Fee Worksheet (SB06) | fee-info.pdf | 30641<br>799a0fc289a676bfa8ccadc3cfc94a4aabef241f | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 361144 | | |

**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.**

Change(s) applied to document,
/J.P.L./
12/9/2015

**AMENDMENTS TO THE SPECIFICATION**

Please amend the paragraph that runs from lines ~~27-29~~ 26-28 on page 13 of the subject application as filed, as follows:

--**Figure 5. ROMA CGH of sKBR3 compared with profiles of 3 independent single cell isolates derived from sparse DNA sequencing. (A) Batch DNA. (B, C, D) Single Cell.**--

Please amend the paragraph that runs from lines 3-4 on page 14 of the subject application as filed, as follows:

--**Figure 7(A-C). Flowchart for genomic profiling by representational sequence counting**--

Please amend the paragraph on line 8 of page 14 of the subject application as filed, as follows:

--**Figure 9(A and B). Nick ligation protocol**--

Please amend the paragraph on line 15 of page 14 of the subject application as filed, as follows:

--**Figure 10(A and B). Schema for targeted sequencing of regional markers**--


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Form PTO-1449 (Substitute) U.S. Department of Commerce Patent and Trademark Office | Application Number | 13/278,333 |
|---|---|---|
| | Filing Date | October 21, 2011 |
| | First Named Inventor | James Hicks |
| **INFORMATION DISCLOSURE STATEMENT** (Use several sheets if necessary) | Art Unit | 1634 |
| | Examiner Name | Frank W. M. Lu |
| | Attorney Docket No. | 5981/81503-A/GJG/REB |

**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 |
|---|---|---|---|---|
| /F.L./ | 1 | 6,569,617 | ~~05-27-2003~~ 5/2003 | Wigler et al. |
| | 2 | 7,531,307 | 05-12-2009 | Wigler et al. |
| | 3 | 8,273,871 | 09-25-2012 | Hannon et al. |
| | 4 | 8,554,488 | 10-08-2013 | Wigler et al. |
| | 5 | 2005-0032095 | 02-10-2005 | Wigler et al. |
| | 6 | 2005-0266444 | 12-01-2005 | Wigler et al. |
| | 7 | 2007-0207481 | 09-06-2007 | Wigler et al. |
| | 8 | 2010-0227768 | 09-10-2010 | Wigler et al. |
| | 9 | 2012-0149593 | 06-14-2012 | Hicks et al. |
| | 10 | 2012-0328607 | 12-27-2012 | Hicks et al. |
| | 11 | 2013-0179999 | 07-11-2013 | Hannon et al. |
| | 12 | 2014-0025307 | 01-23-2014 | Wigler et al. |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

Change(s) applied to document, /M.A.H./ 12/18/2015

**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 | T[6] |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |

| EXAMINER SIGNATURE | /Frank Lu/ | DATE CONSIDERED | 04/19/2014 |
|---|---|---|---|

*EXAMINER: Initial if citation 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 of 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.

UNITED STATES PATENT AND TRADEMARK OFFICE

**UNITED STATES DEPARTMENT OF COMMERCE**
**U.S. Patent and Trademark Office**
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

23432 e 2016-06-30

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

**Paper No.**

| | | | |
|---|---|---|---|
| **Application No.:** | **13/278,333** | **Date Mailed:** | **2016-06-30**<br>**PAPER** |
| First Named Inventor: | James Hicks | Examiner: | LU, FRANK WEI MIN |
| Attorney Docket No.: | 5981/81503-A/GJG/GC | Art Unit: | 1634 |
| Confirmation No.: | 4469 | Filing Date: | 10/21/2011 |

**Please find attached an Office communication concerning this application or proceeding.**

**Commissioner for Patents**

PTO-90c (Rev.08-06)



IPR2025-01353; IPR2025-01355
CSHL EX2002

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

June 30, 2016

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

**DECISION ON PETITION**

In re Application of
HICKS ET AL.
Application No. 13/278,333
Filed: 21 October, 2011
Attorney Docket No. 5981/81503-A/GJG/GC

***ACCEPTANCE OF COLOR DRAWINGS***

This is a decision on the Petition to Accept Color Drawings under 37 C.F.R 1.84 (a) (2), received in the United States Patent and Trademark Office (USPTO) on October 21, 2011.

The petition is **GRANTED**.

A grantable petition under 37 C.F.R. 1.84(a) (2) must meet the following requirements:

1. Payment of the fee set forth under 37 C.F.R. 1.17(h),
2. Submission of three (3) sets of color drawings, or one (1) set of color drawings if filed by EFS-Web, and
3. The first paragraph of the brief description of the drawings in the Specification contains the required language pertaining to colored drawings.

A review of U.S. Patent and Trademark records indicates that each of the above three requirements have been satisfied. Accordingly, the petition under 37 C.F.R. 1.84(a)(2) is GRANTED.

Telephone inquiries relating to this decision may be directed to the undersigned in the Office of Data Management at 571-272-4200.

/NIOMI P. FARMER/
Office of Data Management
Publications Branch

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 NO. | ISSUE DATE | PATENT NO. | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 13/278,333 | 08/02/2016 | 9404156 | 5981/81503-A/GJG/GC | 4469 |

23432 7590 07/13/2016

COOPER & DUNHAM, LLP
30 Rockefeller Plaza
20th Floor
NEW YORK, NY 10112

# ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

**Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)**
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 731 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

James Hicks, Locust Valley, NY;
Nicholas Navin, Houston, TX;
Jennifer Troge, Oyster Bay, NY;
Zihua Wang, Huntington, NY;
Michael Wigler, Cold Spring Harbor, NY;

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 USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage and facilitate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

DocCode – PATENT.GRANT

# SCORE Placeholder Sheet for IFW Content

Application Number: **13278333** Issue Date: **08/02/2016**

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

- **Utility Patent Grant with Color/Grayscale Drawings**

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: May 13, 2015


IPR2025-01353; IPR2025-01355
CSHL EX2002

Docket No. 81503-A/GJG/AC

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Patentees : James Hicks et al.

Patent No. : 9,404,156

Issued : August 2, 2016

U.S. Serial No. : 13/278,333

Filed : October 21, 2011

For : VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION

Conf. No. : 4469

30 Rockefeller Plaza
20th Floor
New York, New York 10112
June 19, 2017

**Attn: Certificate of Corrections Branch**

**BY EFS-WEB**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

**REQUEST FOR A CERTIFICATE OF CORRECTION UNDER 37 C.F.R. §1.323**

Patentees hereby request that the Director of the U.S. Patent and Trademark Office issue a Certificate of Correction under 37 C.F.R. §1.323 in connection with above-identified U.S. Patent No. 9,404,156.

Patentees attach hereto as **Exhibit A**, Form PTO/SB/44 (PTO-1050) indicating the correction to be made in the subject patent. The correction to be made in the above-identified patent is as follows:

On the first page of the specification, column 1, third paragraph, lines 16-19, the government support clause should be revised as follows:

~~The~~This invention ~~disclosed herein~~ was ~~supported by the~~made with government support under grant ~~No.~~ W81XWH-09-1-0591 ~~from the Department of Defense~~awarded by Army Medical Research Materiel Command. ~~Accordingly, the U.S. Government~~The government has certain rights ~~to this~~in the invention.


IPR2025-01353; IPR2025-01355
CSHL EX2002

This change is being made to comply with 37 C.F.R. §401.14(f)(4) and 35 U.S.C. §202(c)(6).

37 C.F.R. §1.323 provides that the Office may issue a certificate of correction under the conditions specified in 35 U.S.C. §255 at the request of the patentee or the patentee's assignee upon payment of the fee set forth in 37 C.F.R. §1.20(a). 35 U.S.C. §255 recites:

> Whenever a mistake of a clerical or typographical nature, or of minor character, which was not the fault of the Patent and Trademark Office, appears in a patent and a showing has been made that such mistake occurred in good faith, the Director may, upon payment of the required fee, issue a certificate of correction, if the correction does not involve such changes in the patent as would constitute new matter or would require reexamination. Such patent, together with the certificate, shall have the same effect and operation in law on the trial of actions for causes thereafter arising as if the same had been originally issued in such corrected form.

Patentees maintain that the mistake to be corrected is of a clerical or typographical nature, or of minor character, and that this mistake occurred in good faith. In addition, Patentees maintain that the correction does not involve any change in the patent that would constitute new matter or require re-examination. The fee required under 37 C.F.R. §1.20(a) for filing of this certificate of correction is ONE HUNDRED DOLLARS ($100.00) and the undersigned hereby authorizes charging this amount to Deposit Account Number 03-3125.

On behalf of the Patentees, the undersigned respectfully requests that the Director issue a Certificate of Correction under 37 C.F.R. §1.323 which provides for the correction of applicants' mistake upon payment of the fee set forth in 37 C.F.R. §1.20(a), provided that such a mistake occurred in good faith and the correction does not involve such changes in the patent as would constitute new matter or would require reexamination.

No fee, other than the enclosed $100.00 fee for a Certificate of Correction, is deemed necessary in connection with the filing of this Request. However, if any additional fee is required, authorization is hereby given to charge the amount of such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Gary J. Gershik
Registration No. 39,992
Attorney for Patentees
Cooper & Dunham LLP
30 Rockefeller Plaza
20th Floor
New York, New York 10112
(212) 278-0400

**Certificate of Transmission**

I hereby certify that this correspondence is being transmitted via the Electronic Filing System (EFS) to the U.S. Patent and Trademark Office on the date below.

AbbieChung 6/19/2017
Abbie Chung Date

PTO/SB/44 (09-07)
Approved for use through 01/31/2020. OMB 0651-0033
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.
(Also Form PTO-1050)

# UNITED STATES PATENT AND TRADEMARK OFFICE
## CERTIFICATE OF CORRECTION



PATENT NO. : 9,404,156

APPLICATION NO.: 13/278,333

ISSUE DATE : August 2, 2016

INVENTOR(S) : James Hicks, Nicholas Navin, Jennifer Troge, Zihua Wang, and Michael Wigler

It is certified that an error appears or errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 1, lines 16-19 should read as follows:

This invention was made with government support under grant W81XWH-09-1-0591 awarded by Army Medical Research Materiel Command. The government has certain rights in the invention.

MAILING ADDRESS OF SENDER (Please do not use Customer Number below):

Gary J. Gershik, Esq.
Cooper & Dunham LLP
30 Rockefeller Plaza, New York, NY 10112

This collection of information is required by 37 CFR 1.322, 1.323, and 1.324. 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 1.0 hour 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: Attention Certificate of Corrections Branch, 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.*



Patentees: James Hicks et al.
Patent No.: 9,404,156
Issued: August 2, 2016
**Exhibit A**

IPR2025-01353, IPR2025-01355
CSHL EX2002

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 13278333 |
| **Filing Date:** | 21-Oct-2011 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Filer:** | Gary J. Gershik/Darren Haber |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |

Filed as Large Entity

**Filing Fees for Utility under 35 USC 111(a)**

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| Certificate of correction | 1811 | 1 | 100 | 100 |


IPR2025-01353; IPR2025-01355
CSHL EX2002

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Extension-of-Time:** | | | | |
| **Miscellaneous:** | | | | |
| | | | **Total in USD ($)** | **100** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 29537100 |
| **Application Number:** | 13278333 |
| **International Application Number:** | |
| **Confirmation Number:** | 4469 |
| **Title of Invention:** | VARIETAL COUNTING OF NUCLEIC ACIDS FOR OBTAINING GENOMIC COPY NUMBER INFORMATION |
| **First Named Inventor/Applicant Name:** | James Hicks |
| **Customer Number:** | 23432 |
| **Filer:** | Gary J. Gershik/Darren Haber |
| **Filer Authorized By:** | Gary J. Gershik |
| **Attorney Docket Number:** | 5981/81503-A/GJG/GC |
| **Receipt Date:** | 19-JUN-2017 |
| **Filing Date:** | 21-OCT-2011 |
| **Time Stamp:** | 14:46:15 |
| **Application Type:** | Utility under 35 USC 111(a) |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | DA |
| Payment was successfully received in RAM | $100 |
| RAM confirmation Number | 062017INTEFSW00001194033125 |
| Deposit Account | |
| Authorized User | |

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


IPR2025-01353; IPR2025-01355
CSHL EX2002

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Request for Certificate of Correction | 170619_81503-A_Request_for_Certificate_of_Correction_AC.pdf | 347557<br>f8d8e2fddae85f30a2a5cee267d1f8db85ac070a | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Fee Worksheet (SB06) | fee-info.pdf | 30507<br>014055cfa1affbeec381dfb94c2e5abbecd8d31b | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 378064 | | |

**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.**

UNITED STATES PATENT AND TRADEMARK OFFICE

# CERTIFICATE OF CORRECTION



PATENT NO. : 9,404,156 B2
APPLICATION NO. : 13/278333
DATED : August 2, 2016
INVENTOR(S) : James Hicks et al.

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

In the Specification

Column 1, Lines 16-19 should read as follows:
This invention was made with government support under grant W81XWH-09-1-0591 awarded by Army Medical Research Materiel Command. The government has certain rights in the invention.

Signed and Sealed this
Eighteenth Day of July, 2017

Joseph Matal

Joseph Matal
*Performing the Functions and Duties of the*
*Under Secretary of Commerce for Intellectual Property and*
*Director of the United States Patent and Trademark Office*



IPR2025-01353; IPR2025-01355
CSHL EX2002