Notice of Allowability	Application No. 15/471,785	Applicant(s) AMORESE et al.	
	Examiner Angela Bertagna	Art Unit 1637	AIA 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 RCE filed on 6/15/18.
   ☐ 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 See Continuation Sheet . 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 areply 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. _____.
5. ☑ Examiner's Amendment/Comment
6. ☐ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/ANGELA M. BERTAGNA/
Primary Examiner, Art Unit 1637

Continuation of 3. The allowed claim(s) is/are: 51-63,66-68 and 70-75

DETAILED CORRESPONDENCE

Notice of Pre-AIA or AIA Status

The present application is being examined under the pre-AIA first to invent provisions.

Continued Examination Under 37 CFR 1.114

A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after allowance or after an Office action under *Ex Parte Quayle*, 25 USPQ 74, 453 O.G. 213 (Comm'r Pat. 1935). Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, prosecution in this application has been reopened pursuant to 37 CFR 1.114. Applicant's submission filed on June 15, 2018 has been entered.

Information Disclosure Statement

The Information Disclosure Statement filed on June 15, 2018 has been considered.

Allowable Subject Matter

Claims 51-63, 66-68, and 70-75 are allowed per the examiner's amendment set forth in the Notice of Allowability mailed on April 10, 2018.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela Bertagna whose telephone number is (571)272-8291. The examiner can normally be reached on 8-5, M-F.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on 571-272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/ANGELA M. BERTAGNA/
Primary Examiner, Art Unit 1637

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

12754 7590 12/26/2017
Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
CHUNDURU, SURYAPRABHA	
ART UNIT	PAPER NUMBER
1637	

DATE MAILED: 12/26/2017

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
14/390,012	10/01/2014	Christopher Armour	25115-759.831	8522

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NEGATIVE SELECTION OF NON-DESIRED NUCLEIC ACID SEQUENCES

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$960	$0	$0	$960	03/26/2018

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: Maintenance fees are due in utility patents issuing on applications filed on or after Dec. 12, 1980. It is patentee's responsibility to ensure timely payment of maintenance fees when due. More information is available at www.uspto.gov/PatentMaintenanceFees.

PTOL-85 (Rev. 02/11)

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)

12754 7590 12/26/2017
Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

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.
14/390,012	10/01/2014	Christopher Armour	25115-759.831	8522

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR NEGATIVE SELECTION OF NON-DESIRED NUCLEIC ACID SEQUENCES

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$960	$0	$0	$960	03/26/2018

EXAMINER	ART UNIT	CLASS-SUBCLASS
CHUNDURU, SURYAPRABHA	1637	435-006120

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.
14/390,012	10/01/2014	Christopher Armour	25115-759.831	8522

12754 7590 12/26/2017

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
CHUNDURU, SURYAPRABHA	
ART UNIT	PAPER NUMBER
1637	

DATE MAILED: 12/26/2017

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.

Notice of Allowability	Application No. 14/390,012	Applicant(s) Armour et al.	
	Examiner SURYAPRABHA CHUNDURU	Art Unit 1637	AIA 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 11/20/2017 .
   ☐ 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 151-202 . 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 areply 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 11/21/2017
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____ .
5. ☐ Examiner's Amendment/Comment
6. ☑ Examiner's Statement of Reasons for Allowance
7. ☑ Other Bib data sheet .

/SURYAPRABHA CHUNDURU/
Primary Examiner, Art Unit 1637

Notice of Pre-AIA or AIA Status

The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

1. The Applicants' response to the office action filed on November 20, 2017 has been considered and acknowledged.

2. The following is an examiner's statement of reasons for allowance:

Claims 151-202 are allowed.

The present invention is drawn to a method of reducing an abundance ratio of non-desired to desired nucleic acid sequence from a single stranded nucleic acid library and creating a strand retained library of nucleic acids. The closest prior art Jones et al. and Kurn et al. did not teach selectively cleaving the partial double-stranded nucleic acid at a cleavage site in the double-stranded adaptor and use of a partial duplex primer to create a strand retained library of nucleic acids as required by the instant claims. Thus the instant invention was not taught or obvious over the closest prior art.

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."

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SURYAPRABHA CHUNDURU whose telephone number is (571)272-0783. The examiner can normally be reached on 8.00am-4.30pm.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on 571-272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/SURYAPRABHA CHUNDURU/
Primary Examiner, Art Unit 1637

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.
14/990,339	01/07/2016	Doug Amorese	NUGN-762/02US 34079/28	7734

21710 7590 01/03/2018

BROWN RUDNICK LLP
ONE FINANCIAL CENTER
BOSTON, MASSACHUSETTS 02111

EXAMINER	
STRZELECKA, TERESA E	

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
01/03/2018	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

ip@brownrudnick.com
usactions@brownrudnick.com

Office Action Summary	Application No. 14/990,339	Applicant(s) Amorese et al.	
	Examiner TERESA E STRZELECKA	Art Unit 1637	AIA 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 11/17/17 .

☐ 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) 1-9 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) 1-9 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 _____ 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 11/17/17.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date _____.

4) ☐ Other: _____.

DETAILED CORRESPONDENCE

Notice of Pre-AIA or AIA Status

1. The present application is being examined under the pre-AIA first to invent provisions.

2. This office action is in response to an amendment filed November 17, 2017. Claims 1-8 were previously pending. Applicant amended claims 1 and 8 and added new claim 9. Claims 1-9 are under consideration.

3. Applicant's amendments overcame the rejection of claims 1-8 under 35 U.S.C. 112, second paragraph. All other previously presented rejections are maintained for reasons given in the "Response to Arguments" below.

4. This office action contains new grounds for rejection for claim 9.

Information Disclosure Statement

5. The information disclosure statement (IDS) submitted on November 17, 2017 was filed after the mailing date of the non-final office action on August 30, 2017. The submission is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

Response to Arguments

6. Applicant's arguments filed November 17, 2017 have been fully considered but they are not persuasive.

A) Regarding the rejection of claims 1-8 under 35 U.S.C. 101, Applicant argues the following:

"The amended claims are not directed to a judicial exception.

Each independent claim requires creating a nucleic acid library, sequencing nucleic acid in the library using a primer, and denaturing one or more nucleic acid fragments in the library. Those steps alone require the use of laboratory instruments and equipment. Nothing about that process that can reasonably be said to be an "abstract idea".

Moreover, the independent claims go further and require additional laboratory steps. The claims require annealing another primer to the nucleic acids in the library and sequencing the nucleic acids with that primer as well. Those laboratory steps just cannot reasonably be said to be an abstract idea. Instead, the claims are directed to a statutory process.

Even if the amended claims were directed to a judicial exception, which they are not, the claims are still directed to patentable subject matter at least because they recite "significantly more" than a law of nature, natural phenomenon, or abstract idea.

A claim that encompasses a judicial exception is patent-eligible where it adds a specific limitation that is not well-understood, routine, and conventional in the field. *Mayo, 132 S. Ct. at 1298-99.* In determining whether a claim amounts to significantly more than a judicial exception, the claim elements must be considered both individually and as an ordered combination. *Alice, 134 S. Ct. at 2357.*

The use of a second primer that binds to a conserved sequence within the target, after sequencing the target with the first primer, to sequentially sequence the target and thus phase the sequence reads inherently by means of those "on instrument" process steps was not something well-understood, routine, or conventional in the field. On the contrary, it was entirely unknown that a template could be re-sequenced using a second primer that binds a conserved

region in the template, and determining the phasing of loci was often impossible. *See specification, paragraph [0002].*

Moreover, the claimed methods are constrained by the annealing of the second primer to the conserved region in the template. Because the primer must be complementary to a region in the template, *a priori* knowledge of the template sequence is required. Annealing a primer to a region in the template was not routine or well-known because technologies for next-generation sequencing predominantly use primers that bind to sequencing adaptors. Such primers need only be complementary to synthetic sequences in the adaptors and need not have any complementarity to sequences in the template itself. Here, the second primer interacts directly with genomic or other input sequences in the template. If the interaction does not occur, no data is obtained from the second sequencing read. Consequently, the steps of "relating", "distinguishing", "mapping", and "aligning" are only possible by performing a step that was not well-understood, routine, or conventional. The claims therefore provide "significantly more" than any judicial exception."

Response

However, as explained in the rejection, the steps of library construction and sequencing are data gathering steps for the abstract step of "compiling data" (claim 1 and 8) or aligning the sequence reads (claim 9). Therefore, since compiling data and aligning sequence reads are abstract ideas, and the steps of library construction and sequencing are not novel as evidenced by Shendure et al., the claims are directed to a judicial exception without "significantly more".

The rejection is maintained.

B) Regarding the rejection of claims 1-8 under 35 U.S.C. 102(b) as anticipated by Shendure et. Al., Applicant argues the following:

"The claims are not anticipated by Shendure.

Each independent claim requires sequencing a template with a first primer and a second primer that hybridizes to a naturally-occurring, conserved region in the subject genome.

Shendure only discloses primers that bind to synthetic sequences in adaptor oligonucleotides and does not teach or suggest the use of a primer that is complementary to a naturally-occurring genomic sequence.

Shendure reports cell-free construction of a DNA library and ligation-based sequencing of the library fragments. *Shendure, abstract.* Shendure's sequencing method requires parallel sequencing reactions using four oligonucleotide "anchor" primers that are complementary to sequences from adaptor oligonucleotides used to make the library. *Shendure, Supplementary FIG. 7, page 10; and Supplemental Note 6; page 32.* Each library fragment is subjected to four separate reaction cycles in which a single anchor primer is repeatedly annealed to a template and then stripped from it. *Shendure, Supplementary FIG. 7, page 10; and Supplemental Note 6; page 32.* Shendure does not disclose sequential sequencing of templates using two different primers.

In addition, Shendure does not disclose the use of a primer that is complementary to a naturally-occurring, conserved sequence within a template.

Shendure does not anticipate the amended claims because it does not disclose all of the elements of the claimed methods. In particular, Shendure does not disclose sequential sequencing of nucleic acid templates using two different primers, nor does it disclose sequencing using a primer that is complementary to a conserved region in the template.

Applicant believes that the rejections have been addressed and overcome."

Response

However, Applicant did not define the term "conserved region within a library", therefore, since the library contains universal primer sequences, these are the "conserved regions", and thus Shendure et al. anticipates the instant claims.

The rejection is maintained.

Claim Rejections - 35 USC § 101

7. 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.

8. Claims 1-9 are rejected under 35 U.S.C. 101 because the claimed invention is directed to a judicial exception (i.e., a law of nature, a natural phenomenon, or an abstract idea) without significantly more.

Claim 1-8 are directed to relating two nucleic acid sequences (claims 1-7), distinguishing between two closely related nucleic acid sequences (claim 8), or sequencing (claim 9) therefore, they are inherently directed to an abstract ideas, since to "relate" sequences to each other one has to compare them, which involves a mental process. Further, claim 9 requires alignment of first

and second sequence reads to map the sequence reads to the template strand, which is also an abstract method step. The claims do not include additional elements that are sufficient to amount to significantly more than the judicial exception because the method steps of creating nucleic acid libraries and sequencing such libraries were well known and routine in the art at the time of the invention, as evidenced by the cited below reference of Shendure et al.

In conclusion, the claims are not patent-eligible under 35 U.S.C. 101.

Claim Interpretation

9. Applicant defined the term "directional library" in paragraph [0034] as follows: "[0034] The term "strand specific" or "directional", as used herein, refers to the ability to differentiate in a double-stranded polynucleotide between the original template strand and the strand that is complementary to the original template strand."

10. Applicant did not define the term "conserved region", therefore it is interpreted as any sequence.

11. Applicant did not define the term "variable region", therefore it is interpreted as any sequence.

12. Applicant did not define the term "reduced complexity", therefore any library is considered to anticipate the term.

13. Applicant did not define the term "target capture", therefore any method of retrieving the target is considered to anticipate this term.

Claim Rejections - 35 USC § 102

14. 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.

15. Claims 1-9 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Shendure et al. (Science, vol. 309, pp. 1728-1732, and supplementary material, pp. 1-41, 2005; previously cited).

Claims 1, 8 and 9 are considered together in claim 8, since it is specie of claim 1.

Regarding claims 1 and 8, Shendure et al. teach a method for distinguishing between two closely related nucleic acid sequences, the method comprising (Abstract):

a. creating a strand-oriented nucleic acid library with closely related nucleic acid sequences as inserts (Fig. 1; supplement pages 13-23);

b. sequencing the strand-oriented library with an oligonucleotide primer (Fig. 1; Fig. S7; supplement page 32)

c. denaturing the first strands of the nucleic acid fragments in the library (Fig. S7; supplement page 32);

d. annealing a new oligonucleotide primer that is complementary to a conserved region or conserved regions within the nucleic acid fragments in the nucleic acid library (Fig. S7; supplement page 32);

e. sequencing the nucleic acid library with the new oligonucleotide primer (Fig. S7; supplement page 32); and

f. compiling data from first and second sequencing reads to map reads originating from the same nucleic acid fragments (page 1730, first and last paragraph; page 1731, first and fourth paragraph; Fig. 3; Table 2).

Regarding claims 2 and 3, Shendure et al. teach creating a library from the full E. coli genome, and detecting sequence variations, therefore they inherently teach amplicons originating from conserved regions adjacent to variable regions (supplement pages 13-23; Fig. 3; Table 2).

Regarding claims 4 and 5, Shendure et al. teach aligning multiple variable regions to differentiate related transcripts of related microorganisms (Fig. 3; Table 2).

Regarding claims 6 and 7, Shendure et al. teach reduced complexity libraries obtained by target capture on emulsion PCR beads (Fig. 1; supplement page 24, 25).

16. No claims are allowed.

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. Any inquiry concerning this communication or earlier communications from the examiner should be directed to TERESA E STRZELECKA whose telephone number is (571)272-0789. The examiner can normally be reached on M-F 9-7.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on 571-272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/TERESA E STRZELECKA/
Primary Examiner, Art Unit 1637
December 28, 2017

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.
13/750,768	01/25/2013	Doug Amorese	25115-755.201	1162

12754 7590 02/05/2014

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
KAUP, SAHANA S	

ART UNIT	PAPER NUMBER
1639	

MAIL DATE	DELIVERY MODE
02/05/2014	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. 13/750,768	Applicant(s) AMORESE ET AL.	
	Examiner SAHANA KAUP	Art Unit 1639	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 *20 May 2013*.

☐ 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) *1-3,6,9-11,14,40-42,45-49 and 51-54* 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) *1-3,6, 9-11,14, 40-42, 45-49 and 51-54* 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 _____ 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 *4 documents*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

Priority

This application filed on 25 January 2013 claims priority to US provisional 61/591,241 filed on 26 January 2012.

Status of the Application

Claims 1-3, 6, 9, 11, 14, 40-42, 45-49 and 51-54 are pending and under examination.

Claim Rejections - 35 USC § 102

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

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.

Claims 1-3, 6, 9, 11, 14, 40-42, 45-48 and 51-53 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Borodina et al.(Methods Enzymol. 2011;500:79-98, published January 1, 2011).

Borodina et al. teach: A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a. fragmenting nucleic acids in a sample comprising the nucleic acids, thereby generating nucleic acid fragments;

b. appending a first adaptor sequence to the nucleic acid fragments, thereby generating nucleic acid fragments comprising the first adaptor sequence;

c. annealing one or more oligonucleotides to the nucleic acid fragments comprising the first adaptor sequence, whereby wherein each of the one or more oligonucleotides comprises a 3' portion that is complementary to a nucleic acid sequence of interest present in one or more of the nucleic acid fragments, and a 5' portion comprising a second adaptor sequence;

d. extending the one or more oligonucleotides annealed to the nucleic acid fragments comprising the first adaptor sequence with a polymerase thereby generating one or more oligonucleotide extension products comprising sequence complementary to the first adaptor sequence at a first end, the nucleic acid sequence of interest, and the second adaptor sequence at a second end; and

e. amplifying the one or more oligonucleotide extension products using a first primer that is complementary to directed against the complement of the first adaptor sequence and a second primer that is directed against the complement of the second adaptor sequence to enrich for the one or more oligonucleotide extension products comprising the nucleic acid sequence of interest nucleic acid fragments containing the

first adaptor and the second adaptor sequence at each end (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 1**.

Borodina et al. teach: The method of claim 1, further comprising an additional step of sequencing the one or more oligonucleotide extension products following amplification (e.g. as indicated in ssRNA-SEQ library preparation section, pg. 92-96 and Figure 5.2, pg. 85) as set forth by instant **claim 2**.

Borodina et al. teach: The method of claim 1, wherein the nucleic acid sequence of interest comprises genomic DNA, RNA, or cDNA (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-84 and Figure 5.1, pg. 84) as set forth by instant **claim 3**.

Borodina et al. teach: The method of claim 52, further comprising denaturing the nucleic acid fragments prior to step c, thereby generating single-stranded nucleic acid fragments comprising the first adaptor sequence at the 5'end (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 6**.

Borodina et al. teach: The method of claim 1, wherein the first adaptor sequence and the second adaptor sequence are distinct from each other (e.g. as indicated in section 2.4.7, pg. 95-96) as set forth by instant **claim 9**.

Borodina et al. teach: The method of claim 1, wherein step b is performed by ligation (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 11**.

Borodina et al. teach: The method of claim 1, wherein the polymerase is a DNA polymerase (e.g. as indicated in 2.3 cDNA synthesis section, pg. 90)as set forth by instant **claim 14**.

Borodina et al. teach: A method for enriching for a nucleic acid sequences sequence of interest from a library comprising nucleic acid inserts with a first adaptor sequence on a first end and a second adaptor sequence on a second end, the method comprising:

a. denaturing the nucleic acid inserts, thereby generating a library of single stranded nucleic acid inserts;

b. annealing one or more oligonucleotides to the single stranded nucleic acid inserts, wherein each of the one or more oligonucleotides comprises a 3' portion that is complementary to a target nucleic acid sequence of interest present in one or more of the nucleic acid inserts, and a 5' portion comprising a third adaptor sequence;

c. extending the one or more oligonucleotides with a polymerase thereby generating one or more oligonucleotide extension products comprising sequence complementary to the first adaptor sequence at the first end, the nucleic acid sequence of interest, and the third adaptor sequence at the second end; and

d. amplifying the one or more oligonucleotide extension products using a first primer that is directed against the complement of the first adaptor sequence and a second primer that is complementary to directed against the complement of the third adaptor sequence to enrich for the one or more oligonucleotide extension products

comprising the nucleic acid sequence of interest (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 40**.

Borodina et al. teach: The method of claim 40, further comprising an additional step of sequencing the amplified one or more oligonucleotide extension products from step d (e.g. as indicated in ssRNA-SEQ library preparation section, pg. 92-96 and Figure 5.2, pg. 85) as set forth by instant **claim 41**.

Borodina et al. teach: The method of claim 40, wherein the nucleic acid sequence of interest comprises genomic DNA, RNA or cDNA (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 42**.

Borodina et al. teach: The method of claim 40, wherein the first adaptor and the second adaptor are common to each nucleic acid fragment as set forth by instant **claim 45** and the method of claim 40, wherein the third adaptor sequence is common to the one or more oligonucleotides as set forth by instant **claim 46** (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84; section 2.4.7, pg. 95-96).

Borodina et al. teach: The method of claim 40, wherein the first adaptor and the second adaptor are distinct from each other as set forth by instant **claim 47** and the method of claim 40, wherein the first adaptor and the second adaptor are the same as set forth by instant **claim 48** (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84; section 2.4.7, pg. 95-96).

Borodina et al. teach: The method of claim 1, wherein each of the one or more oligonucleotides comprises a 3' portion that is complementary to a different nucleic acid sequence of interest in one or more of the nucleic acid fragments, and a 5' portion comprising the second adapter sequence (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84; section 2.4.7, pg. 95-96) as set forth by instant **claim 51**.

Borodina et al. teach: The method of claim 1, wherein the nucleic acid fragments are double stranded (e.g. as indicated in 2.1.1 section, pg. 84-86) as set forth by instant **claim 52**.

Borodina et al. teach: The method of claim 1, wherein the first adaptor sequence is appended to a 5' end of each of the nucleic acid fragments (e.g. as indicated in ssRNA-SEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 53**.

Claim Rejections - 35 USC § 103

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been

obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

Claims 10, 49 and 54 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Borodina et al. in view of Vaidyanathan et al.(US20110104785).

Borodina et al. teach adaptors for nucleic acid amplification and sequencing. However, Borodina et al. do not teach barcodes or restriction site sequences within the adaptor.

Vaidyanathan et al. teach: The method of claim 1, wherein the first adaptor sequence and/or the second adaptor sequence further comprise barcode sequence as set forth by instant **claim 10** and the method of claim 40, wherein the first adaptor, the second adaptor and/or the third adaptor sequence further comprise barcode sequence as set forth by instant **claim 49**(e.g. as indicated in 0072, pg. 6-7).

Vaidyanathan et al. teach: The method of claim 1, wherein the first adaptor sequence comprises a restriction and/or cleavage site for a nucleic acid modifying enzyme (e.g. as indicated in 0072, pg. 6-7)as set forth by instant **claim 54**.

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to combine the above citations to develop the instant invention because both teach adaptors for RNA sequencing. In addition, Vaidyanathan et al. teach tagged adaptors with the features of identifiers(i.e. barcodes) or restriction site sequences. Therefore, it would be obvious for a person of ordinary skill in the art to develop the instant invention of method for enriching for a nucleic acid sequence of interest in a sample, as taught by Borodina et al., in which the adaptors have features such as identifiers(i.e. barcodes) and restriction site sequences. One of ordinary skill in the art would have been motivated to combine these citations and would have had a reasonable expectation of success in doing so because Vaidyanathan et al. teach modified adaptors are suitable for RNA sequencing(e.g. as indicated para 0058, pg. 6).

Conclusion

No claims are allowable.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SAHANA KAUP whose telephone number is (571)272-6897. The examiner can normally be reached on Monday - Friday 7:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, HEATHER CALAMITA can be reached on (571) 272-2876. The fax phone

number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/S. K./
Examiner, Art Unit 1639

/Suryaprabha Chunduru/
Primary Examiner, Art Unit 1637

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.
14/920,254	10/22/2015	Chris Armour	25115-749.301	9848

12754 7590 03/15/2018

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CALIFORNIA 94304
UNITED STATES OF AMERICA

EXAMINER	
WEILER, KAREN S	

ART UNIT	PAPER NUMBER
1639	

NOTIFICATION DATE	DELIVERY MODE
03/15/2018	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	**Application No.** 14/920,254	**Applicant(s)** Armour, Chris	
	Examiner Karen S Weiler	**Art Unit** 1639	**AIA 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 11-07-2017.

☐ 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) 64-82 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) 64-82 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 _____ 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

Status of Claims

1. This action is in reply to the Applicant's amendments and request for reconsideration on the merits, filed on 11-07-2017.

Claims 64-82 of the instant application are under examination under the pre-AIA first to invent provisions.

Claim Interpretation

2. The adaptors of **claim 64** are described as having a ligation strand and a non-ligation strand. These descriptors reflect an intended use of the adaptors, and not a structural feature. A review of the instant Specification did not find a definition for a ligation strand and a non-ligation strand that might require a particular structure, and the descriptors are not terms of art. As the ability of two nucleic acids to be ligated to each other depends upon structural features of each nucleic acid when considered as part of the pair, each nucleic acid is not individually limited in structure. As long as a DNA strand does not possess a structural feature that prevents ligation, it is reasonably considered as a ligatable strand. As well, a non-ligation strand could have any structure.

3. The adaptors of **claim 64** are described as having a ligation strand that is "configured to ligate to a 5' end of a double-stranded nucleic acid". The descriptor "configured to" has the potential to limit the structure of a particular product, such as an adaptor, if the associated functionality requires a particular structure. However, in the instant case, the functionality of ligation to a 5' end of a double-stranded nucleic acid does not require any particular structure. As described above for the descriptor "ligation strand", the ability of two nucleic acids to be ligated to each other depends upon structural features of each nucleic acid in the context of said pair of nucleic acids. Accordingly, as long as the ligation strand does not possess a structural feature at its 3' end that prevents ligation to a 5' end of a double-stranded nucleic acid, it is reasonably considered as being configured to ligate to a 5' end of a double-stranded nucleic acid.

Claim Rejections - 35 USC § 101 - Maintained

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.

4. Claims 64-66, 68-71 and 74-82 are rejected under 35 U.S.C. 101 because the claimed invention is directed to a judicial exception (i.e. a law of nature, a natural phenomenon, or an abstract idea) without significantly more. **Claims 64-66, 68-71 and 74-82** are directed to a kit comprising reagents, which are products of nature. Based upon an analysis with respect to the claims as a whole, the claims do not include additional elements that are sufficient to amount to significantly more than the judicial exception, as explained below.

Every claimed invention must be examined to determine whether the claimed invention complies with 35 U.S.C. 101, particularly whether the claimed invention falls within a 35 U.S.C. 101 judicial exception of non-patentable subject matter (e.g. an abstract idea, law of nature, natural phenomenon, natural product etc.). Phenomena of nature, though just discovered, natural products, mental processes, and abstract intellectual concepts are not patentable, as they are the basic tools of scientific and technological work. See MPEP 2106. As per the "2014 Interim Eligibility Guidance" (Federal Register Vol. 79, No. 241, available 12-16-2014), claims drawn to a process, machine, manufacture or composition of matter are further analyzed according to a two-part process to determine if 1) the claim(s) is/are directed to a judicially recognized exception and, if so, 2) if the claim(s) recite(s) additional elements that amount to *significantly more* than the judicial exception.

Part A) The instant invention of **claim 64** is drawn to a kit comprising one or more primers, reverse transcriptase, glycosylase and two adaptors configured to ligate to a 5' end of a double-stranded nucleic acid, wherein a first adaptor does not comprise a modified dNTP in a ligation strand and a second adaptor does comprise a modified dNTP in a ligation strand. Thus, the kit encompasses natural products. The primers and enzymes are not disclosed as possessing markedly different characteristics from their counterparts in nature. Regarding the claimed primers, it is noted that RNA and DNA molecules within cells can serve as primers. The two adaptors comprise DNA molecules that have features found in nature. DNA molecules found within cells can be ligated to a 5' end of a double-stranded nucleic acid.

DNA molecules found within cells include nucleotides that are modified, via damage or via enzymatic activity (e.g. nucleotide methylation), in nature. The packaging of the nucleic acids (e.g. primers and adaptors) and enzymes (e.g. reverse transcriptase and glycosylase) in the form of a kit does not result in a difference in the structure or characteristics of any individual kit component. Thus, the kit of claim 64 is considered to be directed to "product of nature" exceptions.

Part B) There are no elements of claim 64 that are in addition to the judicial exception (i.e. products of nature) that might add significantly more. Accordingly, the kit is not considered to be patent eligible subject matter.

The dependent claims were similarly analyzed for elements in addition to the judicial exception that, separately or in combination, might add significantly more. Dependent **claim 65** is directed to written instructions, i.e. nonfunctional descriptive material. Instructions are well-understood, routine and conventional in the field, and thus do not add significantly more to the claim. **Claim 66** recites a modified dNTP. As noted above, modified dNTPs are found in cells (e.g. 8-oxo-dGTP), and thus encompass products of nature. Dependent **claims 68-70 and 80** generally recite reagents useful for amplification, sequencing, such as massively parallel sequencing, and end repair. A variety of natural products, such as salts and nucleotides, for example, are encompassed by these claims. Therefore, the claims do not require a non-natural element that might add significantly more. **Claim 71** recites a polyamine, which is a natural product (i.e. judicial exception). Dependent **claims 74-77** recite enzymes that are found in nature, and are therefore judicial exceptions. The limitation of **claim 78** that a primer of the kit comprises a random hybridizing sequence does not distinguish it from its counterpart in nature, in terms of structure, function or other properties. As well, the deoxyribonucleotides of **claim 79** are found in nature. The feature of the first adaptor of **claim 81**, of comprising a sequence complementary to a sequencing primer, does not limit the structure (sequence) of the adaptor, and thereby confer a markedly different structural characteristic. As any sequence can function as a sequencing primer, the complement of a sequencing primer is any sequence. Similarly, the barcode sequence of **claim 82** is not particularly defined. As any sequence can function as a barcode sequence, which is an intended use, an adaptor comprising a barcode does not have markedly different characteristics from a natural product. Therefore, a review of

the dependent claims failed to identify an additional element that may be enough to qualify as significantly more, or further limitations of a product that confer upon it markedly different structural characteristics.

In summary, claims 64-66, 68-71 and 74-82 do not recite something significantly more than the judicial exception and, accordingly, are not patent-eligible.

Claim Rejections - 35 USC § 103 – New Grounds

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. Claims 64-70 and 73-80 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Soldatov et al. (PCT document WO 2011/003630 A1, published 01-13-2011, effectively filed 07-06-2009; as cited in the IDS) in view of Makarov (US PGPub 2011/0269194 A1, published 11-03-2011, effectively filed 04-20-2010).

With regard to instant **claim 64**, Soldatov et al. teach a kit comprising: (a) one or more primers; (b) a reverse transcription enzyme; (c) a glycosylase; and (d)/(e) adaptors comprising a ligation strand configured to ligate to a 5' end of a double stranded nucleic acid and a non-ligation strand (e.g. p. 42, lines 22-34 and p. 17, line 36).

Soldatov et al. teach instant **claim 65**, further comprising written instructions for use of the kit (e.g. absent a new and unobvious functional relationship between the instructions and kit, the limitation is not accorded patentable weight, as per MPEP 2111.05).

Soldatov et al. teach instant **claim 66**, further comprising at least one modified dNTP (e.g. p. 42, lines 22-34).

Soldatov et al. teach instant **claim 67**, wherein the modified dNTP comprises dUTP (e.g. p. 17, line 36).

Soldatov et al. teach instant **claims 68-70**, further comprising reagents for amplification, for sequencing, wherein the sequencing comprises massively parallel sequencing (e.g. DNA polymerases and/or dNTPs; e.g. p. 42, lines 25-29). .

Soldatov et al. teach instant **claim 74**, wherein the glycosylase comprises uracil DNA glycosylase (UDG) (a.k.a. UNG; e.g. p. 17, lines 36-37 and Figure 5).

Soldatov et al. teach instant **claims 75 and 76**, further comprising an endonuclease, wherein the endonuclease comprises an apurinic/apyrimidinic endonuclease (APE) (e.g. p. 42, lines 22-34 and Figure 2A-3).

Soldatov et al. teach instant **claim 77**, further comprising one or more polymerases (e.g. p. 42, lines 22-34).

Soldatov et al. teach instant **claim 78**, wherein at least one of the one or more primers comprises a primer comprising a random hybridizing sequence (e.g. p. 2, lines 1-5 and p. 57, lines 25-26; see also p. 13, lines 19-35).

Soldatov et al. teach instant **claim 79**, further comprising dATP, dCTP, dGTP, dTTP, or any mixture thereof (e.g. p. 42, lines 22-34).

Soldatov et al. teach instant **claim 80**, further comprising reagents for end repair (e.g. implicitly taught by end-repair or end-polishing; e.g. Fig 2A and Figure 5).

Soldatov et al. do not explicitly teach a second adaptor having a modified dNTP in the ligation strand, as per instant claim 64.

With regard to instant **claims 64 and 73**, Makarov teaches a first adaptor that does not have a modified dNTP in the ligation strand and a second adaptor that has a modified dNTP in the ligation strand, wherein the modified dNTP is dUTP (e.g. Figure 2 and 5, and paragraphs [0119] and [0134]-[0135]).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time of the Applicant's invention to modify the kit comprising polarity-specific adaptors of Soldatov et al. by using the polarity-specific adaptors of Makarov, thus obtaining Applicant's invention without any surprising or unexpected results. One would have been motivated to do so because Makarov teaches that the adaptors of his invention enable the isolation of a desired DNA strand (e.g. paragraph [0008]) and can be

used to select polynucleotides of a desired size (e.g. paragraph [0010]). One of ordinary skill in the art would have had a reasonable expectation of success in so modifying the invention because Makarov teaches that the adaptors can be utilized in combination with whole transcriptome amplification (e.g. paragraphs [0015] and [0133]), which is a preferred embodiment taught by Soldatov et al. (e.g. p. 22, lines 36-38, and working examples 1-3 and 7-9).

6. Claims 64-70 and 73-82 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Soldatov et al. (PCT document WO 2011/003630 A1, published 01-13-2011, effectively filed 07-06-2009; as cited in the IDS) in view of Makarov (US PGPub 2011/0269194 A1, published 11-03-2011, effectively filed 04-20-2010), as applied to claims 64-70 and 73-80, in further view of Lennon et al. (Genome Biology 11(2):R15, published 02-05-2011; as cited in the IDS).

The teachings of Soldatov et al. in view of Makarov are described above.

Soldatov et al. in view of Makarov are silent regarding a barcode sequence and a binding site for a sequencing primer within an adaptor, as per instant claims 81 and 82.

Regarding instant **claims 81 and 82**, Lennon et al. teach wherein the first adaptor comprises a sequence complementary to a sequencing primer and wherein the first and/or second adaptor comprise a barcode sequence (e.g. adaptors A and B comprise a site for sequencing primer binding and adaptor A comprises a barcode sequence; e.g. Figure 2).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time of the Applicants' invention to modify the adaptors of Soldatov et al. in view of Makarov to include a barcode sequence and sequencing primer binding sites, as taught by Lennon et al., thus obtaining the Applicant's invention without any surprising or unexpected results. One would have been motivated to incorporate a sequencing primer binding site into a first adaptor because sequencing primer binding sites are necessary for performing DNA sequencing. Importantly, Soldatov et al. teach kits for carrying out the methods of their invention (e.g. p. 41, line 15 though p. 44, line 40) which are generally directed to constructing directional libraries (e.g. abstract and p. 5, lines 4-22) and the molecular analysis of said libraries via, e.g., sequencing (e.g. embodiments 9-15, Figure 5 and working examples 1-3 and 7-9). One would have been motivated to incorporate a barcode sequence into a first or second adaptor in order to

enable the processing of multiple samples together in a pool, which increases efficiency. Lennon et al. teach that a molecular barcode functions to link a sequencing read to its library source (e.g. p. 5, 1st column, 1st paragraph), and that inserting a molecular barcode sequence into each construct during library construction is an efficient and flexible way to support sample multiplexing, which is cost-efficient (e.g. p. 5, 2nd column, 2nd paragraph). One of ordinary skill in the art would have had a reasonable expectation of success in so modifying the method since oligonucleotide (i.e. adaptor) synthesis is commercially available and inexpensive, and the incorporation of these short, functional sequences into an adaptor is routinely performed by the person of ordinary skill during adaptor design.

7. Claims 64-80 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Kurn et al. (US PGPub 2011/0224105 A1, published 09-15-2011, effectively filed 08-12-2009; as cited in the IDS) in view of Makarov (US PGPub 2011/0269194 A1, published 11-03-2011, effectively filed 04-20-2010).

With regard to instant **claim 64**, Kurn et al. teach a kit (e.g. paragraphs [0186]-[0199]) comprising: (a) one or more primers (e.g. paragraph [0188]); (b) a reverse transcription enzyme (e.g. paragraphs [0189] and [0109]); (c) a glycosylase (e.g. paragraph [0192]); and (d)/(e) adaptors comprising a ligation strand configured to ligate to a 5' end of a double stranded nucleic acid and a non-ligation strand (e.g. paragraph [0198]).

Kurn et al. teach instant **claim 65**, further comprising written instructions for use of the kit (e.g. paragraph [0199]).

Kurn et al. teach instant **claims 66 and 67**, further comprising at least one modified dNTP, wherein the modified dNTP comprises dUTP (e.g. paragraph [0186]).

Kurn et al. teach instant **claims 68-70**, further comprising reagents for amplification, for sequencing, wherein the sequencing comprises massively parallel sequencing (e.g. paragraphs [0186], [0191]).

Kurn et al. teach instant **claims 71 and 72**, further comprising a polyamine, wherein the polyamine comprises N, N-dimethylethylenediamine (DMED) (e.g. paragraphs [0193], [0010] and [0087]).

Kurrn et al. teach instant **claim 74**, wherein the glycosylase comprises uracil DNA glycosylase (UDG) (e.g. paragraph [0192]).

Kurn et al. teach instant **claims 75 and 76**, further comprising an endonuclease, wherein the endonuclease comprises an apurinic/apyrimidinic endonuclease (APE) (e.g. paragraph [0193]).

Kurn et al. teach instant **claim 77**, further comprising one or more polymerases (e.g. paragraph [0189]).

Kurn et al. teach instant **claim 78**, wherein at least one of the one or more primers comprises a primer comprising a random hybridizing sequence (e.g. [0188]).

Kurn et al. teach instant **claim 79**, further comprising dATP, dCTP, dGTP, dTTP, or any mixture thereof (e.g. paragraphs [0189], [0197]).

Kurn et al. teach instant **claim 80**, further comprising reagents for end repair (e.g. paragraph [0196]).

Kurn et al. do not teach adaptor molecules wherein one of the adaptors has a modified dNTP in the ligation strand, as per instant claim 64.

With regard to instant **claims 64 and 73**, Makarov teaches a first adaptor that does not have a modified dNTP in the ligation strand and a second adaptor that has a modified dNTP in the ligation strand, wherein the modified dNTP is dUTP (e.g. Figure 2 and 5, and paragraphs [0119] and [0134]-[0135]).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time of the Applicant's invention to modify the kit of Kurn et al. by including the adaptors of Makarov, thus obtaining Applicant's invention without any surprising or unexpected results. One would have been motivated to do so because the adaptors of Makarov are associated with several advantages, such as the ability to select polynucleotides of a desired size (e.g. paragraph [0010]) and to isolate a desired DNA strand of the polynucleotide (e.g. paragraph [0008]). The biotin label additionally facilitates purification of the adaptor-ligated polynucleotides from reagents of prior reaction steps, such as the end-repair and kinase steps. One of ordinary skill in the art would have had a reasonable expectation of success in so modifying the invention because Kurn et al. teach that their kit may contain any adaptor molecules known in the art (e.g. paragraph [0198]), and Makarov teaches that the adaptors of his invention can be utilized in combination with whole transcriptome amplification (e.g. paragraphs [0015] and [0133]), which is a use of the kit of Kurn et al. (e.g. paragraph [0036] and Figure 6).

8. Claims 64-82 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Kurn et al. (US PGPub 2011/0224105 A1, published 09-15-2011, effectively filed 08-12-2009; as cited in the IDS) in view of Makarov (US PGPub 2011/0269194 A1, published 11-03-2011, effectively filed 04-20-2010), as applied to claims 64-80, in further view of Lennon et al. (Genome Biology 11(2):R15, published 02-05-2011; as cited in the IDS).

The teachings of Kurn et al. in view of Makarov are described above.

Kurn et al. in view of Makarov are silent regarding a barcode sequence and a binding site for a sequencing primer within an adaptor, as per instant claims 81 and 82.

With regard to instant **claims 81 and 82**, Lennon et al. teach wherein the first adaptor comprises a sequence complementary to a sequencing primer and wherein the first and/or second adaptor comprise a barcode sequence (e.g. adaptors A and B comprise a site for sequencing primer binding and adaptor A comprises a barcode sequence; e.g. Figure 2).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time of the Applicants' invention to modify the adaptors of Kurn et al. in view of Makarov to include a barcode sequence and sequencing primer binding sites, as taught by Lennon et al., thus obtaining the Applicant's invention without any surprising or unexpected results. One would have been motivated to incorporate a sequencing primer binding site into a first adaptor because sequencing primer binding sites are necessary for performing DNA sequencing, which is a disclosed use of the kit taught by Kurn et al. (e.g. paragraphs [0056], [0048], [0199] and Figure 6). One would have been motivated to incorporate a barcode sequence into a first or second adaptor in order to enable the processing of multiple samples together in a pool, which increases efficiency. Lennon et al. teach that a molecular barcode functions to link a sequencing read to its library source (e.g. p. 5, 1st column, 1st paragraph), and that inserting a molecular barcode sequence into each construct during library construction is an efficient and flexible way to support sample multiplexing, which is cost-efficient (e.g. p. 5, 2nd column, 2nd paragraph). One of ordinary skill in the art would have had a reasonable expectation of success in so modifying the method since oligonucleotide (i.e. adaptor) synthesis is commercially available and inexpensive, and the

introduction of these short, functional sequences into an adaptor is routinely performed by the person of ordinary skill during adaptor design.

Response to Arguments

9. RE: 35 U.S.C. § 101 rejection of claims 64-66, 68-71 and 74-80

Applicant submits that amended claim 64 is eligible for patenting because "the first adaptor and second adaptor as claimed in amended independent claim 64 are not natural products" and "the two specific adaptors not conventional, well-understood, and routine, and make the claims significantly more than a judicial exception". This argument was carefully considered, but was not found to be persuasive. The analysis of amended claim 64 for patent eligibility is described above, and leads to the conclusion that the claimed adaptors do not exhibit markedly different structural characteristics from their natural counterparts. See also the Claim Interpretation section.

10. RE: 35 U.S.C. § 103 rejection of claims 64-70 and 73-80 (Soldatov et al. in view of Hendrickson)

This rejection is withdrawn in view of the Applicants' claim amendments.

11. RE: 35 U.S.C. § 103 rejection of claims 64-80 (Kurn et al. in view of Hendrickson)

This rejection is withdrawn in view of the Applicants' claim amendments.

Conclusion

12. Claims 64-82 are rejected.

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.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Karen S Weiler whose telephone number is (571)270-5322. The examiner can normally be reached on M-F, ~8:00 - ~ 4:30 EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Calamita can be reached on ((571) 272-2876. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/Karen S. Weiler/
Primary Examiner, Art Unit 1639

Karen S. Weiler
Primary Examiner
Art Unit 1639

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.
14/995,882	01/14/2016	Chris Armour	25115-749.302	6916

12754 7590 04/04/2016
Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER
WEILER, KAREN S

ART UNIT	PAPER NUMBER
1639	

NOTIFICATION DATE	DELIVERY MODE
04/04/2016	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 14/995,882	Applicant(s) ARMOUR, CHRIS	
	Examiner KAREN S. WEILER	Art Unit 1639	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 *01-14-2016*.

☐ 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) *1-28* 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) *1-28* 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 _____ 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

Claims 1-28 of the instant application are under examination under the pre-AIA first to invent provisions.

Priority

1. Applicant's claim for the benefit of a prior-filed application under 35 U.S.C. 119(e) or under 35 U.S.C. 120, 121, 365(c), or 386(c) is acknowledged. Applicant has not complied with one or more conditions for receiving the benefit of an earlier filing date under 35 U.S.C. 119(e) or under 35 U.S.C. 120 as follows:

The later-filed application must be an application for a patent for an invention which is also disclosed in the prior application (the parent or original nonprovisional application or provisional application). The disclosure of the invention in the parent application and in the later-filed application must be sufficient to comply with the requirements of 35 U.S.C. 112(a) or the first paragraph of pre-AIA 35 U.S.C. 112, except for the best mode requirement. See *Transco Products, Inc. v. Performance Contracting, Inc.*, 38 F.3d 551, 32 USPQ2d 1077 (Fed. Cir. 1994)

The disclosure of the prior-filed applications, Application Nos. 13/643056 and 14/920254, fail to provide adequate support or enablement in the manner provided by 35 U.S.C. 112(a) or pre-AIA 35 U.S.C. 112, first paragraph for one or more claims of this application. A comparison of the scope of instant **claim 10** and the scope of the Specification of the present and parental applications reveals that they are not commensurate. More specifically, the Specification at paragraph [0079] discloses that a barcode sequence may be part of an adaptor sequence. However, instant claim 10 is more broad in scope, as it recites a barcode within the amplified product of step (f) and thus includes additional embodiments e.g. a barcode within the amplification primers.

Claim Rejections - 35 USC § 103

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1-28 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Soldatov et al. (PCT publication WO 2011/003630 A1, published 01-13-2011; as cited in the IDS) in view of Hendrickson (USPGPub 2012/0238738 A1, published 09-20-2012), as evidenced by Levesque-Sergerie et al. (BMC Molecular Biology 8:93, published 10-22-2007) for claim 22.

With regard to instant **claim 1**, Soldatov et al. teach method for construction of a directional cDNA library (e.g. method for differentiation of polynucleotide strands; e.g. title), the method comprising:

a. reverse transcribing an RNA in an RNA sample to generate a first strand cDNA (e.g. a first cDNA strand is generated by using random primers for annealing to said at least one RNA strand and an enzyme with Reverse Transcriptase activity; e.g. p. 23, lines 3-5);

b. generating a second strand cDNA from the first strand cDNA, wherein the generating the second strand cDNA comprises incorporating a modified dNTP during the generating the second strand cDNA, thereby generating a double-stranded cDNA (e.g. second strand synthesis is performed using at least one modified dNTP and the first cDNA strand as a template; e.g. p. 23, lines 12-14);

c. performing end repair on the double-stranded cDNA (e.g. end polishing; e.g. Figure 2);

d. ligating a polarity-specific adaptor to the double-stranded cDNA thereby generating a ligation product (e.g. p. 23, lines 14-15);

e. selectively cleaving the second strand cDNA in the ligation product (e.g. the second cDNA strand containing the modified dNTP is degraded; e.g. p. 23, lines 16-18); and

f. amplifying the first strand resulting from step (e), thereby generating a directional cDNA library comprising the first strand cDNA (e.g. the remaining intact first cDNA strand is amplified; e.g. p. 23, lines 21-23).

Soldatov et al. teach instant **claim 2**, wherein the RNA in the RNA sample comprises RNA fragments (e.g. polyA RNA was chemically fragmented prior to cDNA synthesis; e.g. p. 54, line 28).

Soldatov et al. teach instant **claims 3 and 4**, wherein the reverse transcribing the RNA sample is performed using a primer comprising a random sequence, wherein the primer is a random hexamer primer (e.g. a first cDNA strand is generated by using random primers; e.g. p 23, line 3; e.g. first strand cDNA synthesis using a tagged random hexamer primer; e.g. p. 2, lines 3-5; e.g. $(dN)_6$; e.g. Figure 5).

Soldatov et al. teach instant **claim 5**, wherein the modified dNTP comprises dUTP (e.g. p. 17, lines 32-33).

Soldatov et al. teach instant **claim 6**, wherein the method further comprises purifying the double-stranded cDNA prior to step (c) (e.g. different purification methods may be used for removal of modified and unmodified dNTPs from the solution; e.g. p. 16, lines 10-11; see also p. 50, lines 24-25).

Soldatov et al. teach instant **claim 7**, wherein the end repair comprises adding a dA nucleotide to the double-stranded cDNA (e.g. A-tailing; e.g. Figure 5).

Soldatov et al. teach instant **claim 8**, wherein the polarity specific adaptor comprises a dT overhang (e.g. Figure 5).

Soldatov et al. teach instant **claim 9**, wherein the amplifying the first strand cDNA is performed using one or more tailed primers (e.g. PCR #1 primer in Figure 5).

Soldatov et al. teach instant **claim 11**, wherein the amplification comprises polymerase chain reaction (PCR) (e.g. p. 23, lines 21-23).

Soldatov et al. teach instant **claim 12**, further comprising sequencing the amplified product of step (f) (e.g. p. 28, line 3).

Soldatov et al. teach instant **claims 13, 27 and 28**, wherein the sequencing comprises massively parallel sequencing, wherein the massively parallel sequencing comprises bridge amplification, wherein the massively parallel sequencing comprises use of labeled reversible terminators (e.g. library preparation and sequencing using the Illumina sequencing platform; e.g. p. 50, line 33 through p. 51, line 7).

Soldatov et al. teach instant **claims 14, 15, 16 and 17**, wherein step (e) comprises selectively cleaving a base portion of the modified dNTP, thereby forming an abasic site, wherein the selectively cleaving the base portion of the modified dNTP comprises use of an enzyme, wherein the enzyme

comprises a glycosylase, wherein the glycosylase is uracil DNA glycosylase (UDG) (e.g. specific degradation of the modified dNTP, dUTP, by uracil-N-glycosylase to result in an abasic site; e.g. p. 17, lines 36-39).

Soldatov et al. teach instant **claim 18**, wherein step (e) comprises cleaving a phosphodiester backbone at the abasic site (e.g. at temperatures of over 50˚C in slightly alkaline conditions nicks occur in the DNA strand at the abasic positions; e.g. p. 18, lines 1-2).

Soldatov et al. teach instant **claim 22**, wherein the reverse transcribing is performed using a reverse transcriptase, wherein the reverse transcriptase is moloney murine leukemia virus (M-MLV) reverse transcriptase (e.g. Superscript III; e.g. p. 50, line 2, wherein Levesque-Sergerie et al. teach that Superscript III is a version of M-MLV; e.g. p. 4, Table 1).

Soldatov et al. teach instant **claims 24, 25 and 26**, further comprising depleting non-desired RNA from the RNA sample, wherein the non-desired RNA comprises ribosomal RNA (rRNA), wherein the RNA in the RNA sample comprises poly(A)+ RNA (e.g. isolating total RNA and purifying the polyA+ fraction using the Dynabead mRNA purification kit; e.g. p. 49, lines 20-31, wherein it is common knowledge to one of ordinary skill in the art that total RNA includes rRNA).

Soldatov et al. teach polarity-specific adaptors such a Y-shaped adaptors having both double-stranded and single-stranded regions (e.g. Figure 2A-1), but do not teach a stem-loop structure, as per instant claim 1.

With regard to instant **claim 1**, Hendrickson teaches a stem-loop adaptor, wherein the stem-loop adaptor comprises a modified dNTP (e.g. paragraphs [0009]-[0010] and Figures 1 and 2) and cleaving the portion of the stem-loop adaptor that has the modified dNTP (e.g. paragraph [0012] and Figure 3).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time of the Applicant's invention to modify the method of Soldatov et al. for differentiation of polynucleotide strands by substituting stem-loop adaptors for Y-shaped adaptors, as taught by Hendrickson, thus obtaining Applicant's invention without any surprising or unexpected results. One would have been motivated to do so because Hendrickson teaches the advantages of a stem-loop adaptor over a Y-shaped adaptor, including the ability to remove unligated adaptors and unligated polynucleotides (following the ligation

reaction step) using a simple exonuclease treatment (e.g. paragraph [0008]). Indeed, Hendrickson notes that enzymatic degradation of non-ligated adapters without damaging the ligated adapters and subsequent denaturation of the degradative enzymes in a single step provides significant advantages in efficiency and cost (e.g. paragraph [0060]). One of ordinary skill in the art would have had a reasonable expectation of success in so modifying the invention since cleavage of the stem-loop adaptor following ligation effectively generates the Y-shaped adaptor ligation product of Soldatov et al. (e.g. see Figure 3 of Hendrickson). As well, Hendrickson teaches that the (stem-) loop adapters can be utilized under any condition in which sequencing or amplification or both is desirable (e.g. paragraph [0037]), which is consistent with the aims of Sodatov et al.

Hendrickson et al. teach instant **claims 14-21**, wherein step (e) comprises selectively cleaving a base portion of the modified dNTP, thereby forming an abasic site, wherein the selectively cleaving the base portion of the modified dNTP comprises use of an enzyme, wherein the enzyme comprises a glycosylase, wherein the glycosylase is uracil DNA glycosylase (UDG), wherein step (e) comprises cleaving a phosphodiester backbone at the abasic site, wherein the cleaving the phosphodiester backbone at the abasic site comprises use of an enzyme, wherein the enzyme comprises an endonuclease, wherein step (e) comprises use of a glycosylase and an endonuclease (e.g. cleavage at the modified nucleotide dU may be achieved using uracil DNA glycosylase and endonuclease VIII; e.g. paragraph [0040]).

Soldatov et al. do not teach inclusion of barcode sequences within the polarity-specific adaptors, as per instant claim 10.

Hendrickson teaches instant **claim 10**, wherein the amplified product of step (f) comprises a barcode sequence (e.g. paragraph [0004] and claim 7).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time of the Applicant's invention to modify the invention of Soldatov et al. for differentiation of polynucleotide strands by including a barcode sequence in the adaptor as taught by Hendrickson, thus obtaining Applicant's invention without unexpected or surprising results. One would have been motivated to do so because the use of barcodes to tag or identify samples (or experiments) enables sample multiplexing that takes

advantage of the large capacity of present high-throughput sequencing methodologies, thus improving efficiency and lowering cost. One of ordinary skill in the art would have had a reasonable expectation of success in so modifying the method because the small size of barcodes (e.g. paragraph [0004]) enables them to easily be incorporated into oligonucleotide adaptor design.

Soldatov et al. teach the use of Illumina's DNA sample kit for preparation of the sequencing library (e.g. p. 50, lines 33-37), but are silent with regard to the precise steps and thus the purification of ligation products as per claim 23.

Hendrickson teaches instant **claim 23**, further comprising purifying the ligation product of step (d) (e.g. purification with Agencourt® Ampure® magnetic beads; e.g. paragraph [0072]).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time of the Applicant's invention to modify the invention of Sodatov et al. for differentiation of polynucleotide strands by including a purification step following adaptor ligation, thus obtaining Applicant's invention without unexpected or surprising results. One would have been motivated to do so in order to remove contaminants such as enzymes which could interfere with downstream reaction steps. One of ordinary skill in the art would have had a reasonable expectation of success in so modifying the method since kits for nucleic acid purification at high recovery, such as Beckman Coulter's Agencourt® Ampure® purification system, are widely available and easy to use.

Double Patenting

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory double patenting rejection is appropriate where the claims at issue are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645

(Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); In re Vogel, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the reference application or patent either is shown to be commonly owned with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement. See MPEP § 717.02 for applications subject to examination under the first inventor to file provisions of the AIA as explained in MPEP § 2159. See MPEP §§ 706.02(l)(1) - 706.02(l)(3) for applications not subject to examination under the first inventor to file provisions of the AIA. A terminal disclaimer must be signed in compliance with 37 CFR 1.321(b).

The USPTO Internet website contains terminal disclaimer forms which may be used. Please visit www.uspto.gov/forms/. The filing date of the application in which the form is filed determines what form (e.g., PTO/SB/25, PTO/SB/26, PTO/AIA/25, or PTO/AIA/26) should be used. A web-based eTerminal Disclaimer may be filled out completely online using web-screens. An eTerminal Disclaimer that meets all requirements is auto-processed and approved immediately upon submission. For more information about eTerminal Disclaimers, refer to http://www.uspto.gov/patents/process/file/efs/guidance/eTD-info-I.jsp.

3. Claims 1-5, 11-22 and 24-26 are rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1, 2, 4-8, 10, 11, 17, 19, 21, 22, and 27-29 of U.S. Patent No. 9,206,418.

Although the claims at issue are not identical, they are not patentably distinct from each other because the elements of instant claims 1-5, 11-22 and 24-26 are to be found in patent claims 1, 2, 4-8, 10, 11, 17, 19, 21, 22, and 27-29. Although patent claim 1 has additional limitations with respect to instant claim 1, in this situation the species renders the genus obvious. It is noted that the claims of patent 9,206,418 do not recite the limitation of instant claim 1 that the adaptor has a stem-loop structure. However, upon reading patent claim 1, step (d) which recites ligating adaptors, the person of ordinary skill in the art would consult the specification of the patent to determine the scope of the term "adaptor" and would find at column 22, lines 56-66 that suitable adaptor molecules include stem-loop nucleic acids and thus that stem-loop adaptors were specifically envisioned. As well, upon reading the claim 1 of patent

9,206,418, the person of ordinary skill would consult the patent Specification to determine the scope of "an RNA sample" and would find at column 11, lines 40-51 that an RNA sample comprising mRNA (i.e. polyA+ RNA) and comprising RNA fragments, as per instant claims 2 and 26, was explicitly envisioned. Similarly, patent claim 1 recites "reverse transcribing", wherein column 16, lines 6-8 reveals that reverse transcription is specifically envisioned as using Moloney murine leukemia virus (M-MLV) reverse transcriptase, as per instant claim 22, and mediated by random hexamer primers (e.g. column 5, lines 38-40), as per instant claims 3 and 4. Furthermore, upon reading claim 27 of patent 9,206,418, which is directed to reducing or depleting non-desired nucleic acid sequences from the directional cDNA library, one of ordinary skill in the art would consult the Specification of the patent to determine the scope of "reducing or depleting non-desired nucleic acid sequences from the directional cDNA library" and would find at column 10, lines 23-32 that the depletion of non-desired nucleic acid from the starting nucleic acid population, as per instant claim 24, was intended.

Additionally, instant claims 6-10, 23, 27 and 28 are rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1, 2, 4-8, 10, 11, 17, 19, 21, 22, and 27-29 of U.S. Patent No. 9,206,418 in view of Soldatov et al. (for claims 6, 9, 27 and 28) and in further view of Hendrickson (for claims 7, 8, 10 and 23).

As described above, Soldatov et al. teach the limitation of instant claim 6, purifying the double-stranded cDNA (e.g. e.g. p. 16, lines 10-11). It would have been *prima facie* obvious to one of ordinary skill in the art to modify the invention of patent 9,206,418 to include a step of cDNA purification because the presence of extraneous reagents, including enzymes, can interfere with subsequent reaction steps. Soldatov et al. also teach the limitations of instant claims 7 and 8, directed to ligation of the adaptor to the cDNAs by the method of "TA cloning" (i.e. the ligation of two fragments having complementary T and A single-base overhangs; e.g. Figure 5). It would have been *prima facie* obvious to one of ordinary skill in the art to modify the invention of patent 9,206,418 by performing the ligation step using "TA cloning", because the single base overhang dramatically increases ligation efficiency with respect to blunt end cloning. Additionally, Soldatov et al. teach the use of tailed primers during amplification of the selected cDNA strand (e.g. Figure 5), as per instant claim 9. One of ordinary skill in the art would have found it to

be *prima facie* obvious to modify the invention of patent 9,206,418 such that a tailed primer was used in order to prepare the amplification products for attachment to a sequencing flow cell. Furthermore, whereas patent claim 22 is directed to next generation sequencing methods, a particular sequencing method such as Illumina sequencing, which comprises bridge amplification and the use of labeled reversible terminators, as recited in instant claims 27 and 28, is not specifically claimed. However, Soldatov et al. exemplify the Illumina sequencing platform (e.g. working example 1), and Illumina sequencing would have been obvious to try by the skilled artisan given its widespread use in the scientific research community. Finally, as described above (see 103 rejection), it would have been *prima facie* obvious to one of ordinary skill in the art to modify the invention of patent 9,206,418 to incorporate the limitations of instant claims 10 and 23 in view of the teachings of Hendrickson.

Thus, all of the limitations of instant claims 1-5, 11-22 and 24-26 are recited in claims 1, 2, 4-8, 10, 11, 17, 19, 21, 22, and 27-29 of the issued patent, in view of the disclosure of the patent which defines claim terminology. Furthermore, instant claims 1-28 and patent claims 1, 2, 4-8, 10, 11, 17, 19, 21, 22, and 27-29 are not patentably distinct from each other because the limitations of instant claims 6-10, 23, 27 and 28 would have been *prima facie* obvious over claims 1, 2, 4-8, 10, 11, 17, 19, 21, 22, and 27-29 of patent 9,206,418 in view of Soldatov et al. and Hendrickson, before the effective filing date of the claimed invention.

Conclusion

4. Claims 1-28 are rejected.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to KAREN S. WEILER whose telephone number is (571)270-5322. The examiner can normally be reached on Monday through Friday, 7:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Calamita can be reached on (571)272-2876. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/KAREN S WEILER/
Examiner, Art Unit 1639

Notice of References Cited	Application/Control No. 14/995,882	Applicant(s)/Patent Under Reexamination ARMOUR, CHRIS	
	Examiner KAREN S. WEILER	Art Unit 1639	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-2012/0238738 A1	09-2012	Hendrickson; Cynthia	C12Q1/6813	536/23.1
	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	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Levesque-Sergerie et al. (2007) "Detection limits of several commercial reverse transcriptase enzymes: impact on the low- and high-abundance transcript levels assessed by quantitative RT-PCR" BMC Molecular Biology 8:93
	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.

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.
14/390,012	10/01/2014	Christopher Armour	25115-759.831	8522

12754 7590 04/07/2016

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
CHUNDURU, SURYAPRABHA	

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
04/07/2016	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 14/390,012	Applicant(s) ARMOUR ET AL.	
	Examiner Suryaprabha Chunduru	Art Unit 1637	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 *3/24/16*.

☐ 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) *151-202* 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) *151-202* 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 *10/01/14* 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 *2/5/15; 2/18/15; 3/2/15; 3/10/15;3/18/15; 4/10/15; 3/23/16* .

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

Status

1. Claims 151-172 and new claims 173-202 are pending. Claims 1-150 were cancelled.

Priority

2. This application filed on October 01, 2014 is a 371 of PCT/US13/32606 filed on March 15, 2013, which claims benefit of US 61/661,293 filed on June 18, 2012.

Objection to the Specification

3. The disclosure is objected to the following informalities:

The disclosure is objected to because it contains an embedded hyperlink and/or other form of browser-executable code (see at least page 80-81, para 00218, 00222 and 00227). Applicant is required to delete the embedded hyperlink and/or other form of browser-executable code; references to websites should be limited to the top-level domain name without any prefix such as http:// or other browser-executable code. See MPEP § 608.01.

Claim Rejections - 35 USC § 112

4. 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.

A. Claims 151-167 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. Claim 151 step d) recites 'amplifying the desired nucleic acid sequences, thereby reducing the ratio of non-desired to desired nucleic acid sequence'. The meets and bounds of the claim 151 are unclear and indefinite because it is not clear whether the step a)-c) are required for amplifying desired nucleic acid sequences or not required. Specifically because the step a)-c) are directed to non-desired nucleic acid sequences and are not required for amplification of desired nucleic acids from the single stranded library.

B. Claims 169-171 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. Claim 169 recites 'the method of *any one* of claims 168'. The meets and bounds of the claim 169 are unclear and indefinite because it is not clear which other claims the instant claim referring to.

Claim Rejections - 35 USC § 102

5. 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.

A. Claims 151-159, 161-162, 173-179, 195 and 199-202 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Jones et al. (US 2006/0068415).

Jones et al. teach a method of claim 151, 173, reducing an abundance ratio of non-desired to desired nucleic acid sequence from a single-stranded nucleic acid library or depleting or reducing non-desired nucleic acid sequences from a nucleic acid library comprising:

a) annealing a sequence-specific oligonucleotide probe to a non-desired nucleic acid sequence in the single-stranded nucleic acid library, wherein the single-stranded nucleic acid library comprises single stranded nucleic acid fragments comprising an adaptor sequence on one end or both the ends of the single-stranded nucleic acid sequence (see at least page 9-10, para 0080-0090, page 11, para 0092-0097, page 6, para 0053-0057);

b) extending the oligonucleotide probe to produce a partial double-stranded nucleic acid (see at least page 9-10, para 0080-0090, page 11, para 0092-0097);

c) selectively cleaving the partial double-stranded nucleic acid at a cleavage site in the adaptor (see at least page 9-10, para 0086-0090, page 11, para 0093-0097); and

d) amplifying the desired nucleic acid sequences, thereby reducing the abundance ratio of non-desired to desired nucleic acid from the single-stranded nucleic acid library (see at least page 9-10, para 0080-0090, page 11, para 0093-0097).

With reference to claims 152-159, 161, 173-179, Jones et al. teach that the single stranded nucleic acid fragments are DNA fragments and comprise an adapter sequence at both ends; amplifying comprises a primer specific for adaptor sequence comprises cleavage site; selective cleavage site is capable of recognition by a cleaving agent, which is specific for a double-stranded nucleic acid sequence in the partial double-stranded nucleic acid, which is a

restriction endonuclease (see page 9-10, para 0080-0090, page 11, para 0093-0097, page 5-6, para 0050-0057, 0046).

With reference to claims 162, 195, 199-202, Jones et al. teach that the method further comprises performing PCR with a set pf primers hybridizable to a sequence of the adapter, thereby amplifying a desired nucleic acid and generating unidirectional library, wherein adapter comprises a barcode and purifying the DNA fragments before and after amplification (see at least page 9-12, para 0084-0097). For all the above the claims are anticipated.

B. Claims 151-155, 158-162, 173-174, 176-185, 190, 195, 199 and 201-202 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Fan et al. (US 2008/0241831).

Fan et al. teach a method of claim 151, 173, reducing an abundance ratio of non-desired to desired nucleic acid sequence from a single-stranded nucleic acid library or depleting or reducing non-desired nucleic acid sequences from a nucleic acid library comprising:

a) annealing a sequence-specific oligonucleotide probe to a non-desired nucleic acid sequence in the single-stranded nucleic acid library, wherein the single-stranded nucleic acid library comprises single stranded nucleic acid fragments comprising an adaptor (linker) sequence on one end or both the ends of the single-stranded nucleic acid sequence (see at least page 3-4, para 0026-0035, page 6-7, para 0050-0052, page 13, para 0102, page 15, example 1, indicating modifying the stating nucleic acid molecules to add adaptor sequences having universal priming sequences and hybridizing to a probe nucleic acid sequence);

b) extending the oligonucleotide probe to produce a partial double-stranded nucleic acid (see at least page at least page 3-4, para 0032-0035, page 6-7, para 0050-0052, page 15, example 1, indicating extension of hybridized probe);

c) selectively cleaving (endonuclease) the partial double-stranded nucleic acid at a cleavage site in the adaptor (see at least page at least page 3-4, para 0026-0035, page 6-7, para 0050-0053, page 15, example 1); and

d) amplifying the desired nucleic acid sequences, thereby reducing the abundance ratio of non-desired to desired nucleic acid from the single-stranded nucleic acid library (see at least page at least page 3-4, para 0026-0038, page 7, para 0054, page 15, example 1).

With reference to claims 152-155, 158-159, 173-174, 176-185, Fan et al. teach that the single stranded nucleic acid fragments are DNA fragments and comprise an adapter sequence at both ends; amplifying comprises a primer specific for adaptor sequence comprises cleavage site; selective cleavage site is capable of recognition by a cleaving agent, which is specific for a double-stranded nucleic acid sequence in the partial double-stranded nucleic acid, wherein the sample is a RNA and the method comprises reverse transcribing total RNA into cDNA and RNA comprises rRNA and RNA sample comprises less than 20ng of nucleic acid (see page 3-4, para 0026-0038, page 6-7, para 0050-0053, page 5-6, para 0045-0049).

With reference to claims 160-162, 190, Fan et al. teach that the non-desired nucleic acid sequence (non-coding RNA, rRNA) comprises ribosomal RNA ; amplifying comprises a DNA polymerase (see page 5, para 0043).

With reference to claims 195, 199, 201-202, Fan et al. teach that the adapter sequence comprises a barcode and a set of primers comprise a barcode and purifying the DNA fragments before and after amplification (see page 4, para 0037, page 7, para 0055). For all the above the claims are anticipated.

Claim Rejections - 35 USC § 103

6. The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 151-202 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Fan et al. (US 2008/0241831) in view of Kurn et al. (US 2011/0224105) and Raymond et al. (2010/0029511).

Fan et al. teach a method for reducing non-targeted nucleic acids in a population of nucleic acids as discussed above in section 5B. However, Fan et al. did not teach non-canonical nucleotide comprising nucleic acids, creating basic site, cleaving abasic site, reverse transcription of RNA using dUTP, random primers, partially duplex primer and massively parallel sequencing.

Kurn et al. teach a method of generating a population of nucleic acids comprising modifying nucleic acids to comprise non-canonical nucleotides (uracil or inosine), creating abasic sites during reverse transcription PCR by using dUTP and random primers and cleaving the abasic sites using a cleavage nuclease (see at least page 1-3, para 0005-0023, page 8-10, para 0070-0079). Kurn et al. also teach massively parallel sequencing of amplification nucleic acids (see at least page 17-19, para 0112-0121).

Raymond et al. teach a method for selectively amplifying a target population (mRNA) within a large non-target population of nucleic acids (rRNA, mitochondrial rRNA, globin mRNA) using target specific oligonucleotide primers comprising random nucleotide sequences and partial duplex, analyzing informative RNA gene expression profiling (see at least page 5-7, para 0055-0069).

It would have been prima facie obvious to a person of ordinary skill in the art at the time the invention was made, to modify the method of Fan et al. to incorporate non-canonical nucleotides and reverse transcription of RNA with random primers as taught by Kurn et al. and selective amplification of targeted nucleic acids from a population of nucleic acids comprising large non-target nucleic acids to develop an improved method. The ordinary person skilled in the art would have motivated to combine the references because the ordinary person would have a reasonable expectation of success that the combination of references would result in enhancing the sensitivity of the method. Kurn et al. explicitly taught use of non-canonical nucleotides that would reduce the non-specific priming and enhance the sensitivity of specific amplification and Raymond et al. explicitly taught depleting non-targeted nucleic acids from a sample using oligonucleotide primers comprising non-random sequences and selectively amplifying target nucleic acids to analyse informative RNA gene expression profiling and such a modification of the method is obvious over the cited prior art.

Double Patenting

7. The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible

harassment by multiple assignees. A nonstatutory double patenting rejection is appropriate where the claims at issue are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the reference application or patent either is shown to be commonly owned with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement. A terminal disclaimer must be signed in compliance with 37 CFR 1.321(b).

The USPTO internet Web site contains terminal disclaimer forms which may be used. Please visit http://www.uspto.gov/forms/. The filing date of the application will determine what form should be used. A web-based eTerminal Disclaimer may be filled out completely online using web-screens. An eTerminal Disclaimer that meets all requirements is auto-processed and approved immediately upon submission. For more information about eTerminal Disclaimers, refer to http://www.uspto.gov/patents/process/file/efs/guidance/eTD-info-I.jsp.

Claims 151-169, 173-202 are provisionally rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1,3, 6, 9-10, 14, 40, 42, 46-49, 79-80, 82, 93 of

copending Application No. 13/750,768. Although the claims at issue are not identical, they are not patentably distinct from each other because the method steps of the claims in the co-pending application encompass the method steps as claimed in the instant claims. Specifically the method steps for depleting or reducing non-desired nucleic acids from a single stranded library of the instant claims are within the scope of the claims in the co-pending application. Thus the claims in the co-pending application encompass the instant claims and are related as genus and species, and are coextensive in scope.

The courts have stated that a genus is obvious in view of the teachings of a species. see Slayter, 276 F.2d 408, 411, 125 USPQ 345, 347 (CCPA 1960); and In re Gosteli, 872 F.2d 1008, 10 USPQ2d 1614 (Fed.Cir. 1989). Therefore the instantly claimed method is obvious over the claims in the co-pending application. This is a provisional nonstatutory double patenting rejection because the patentably indistinct claims have not in fact been patented.

Conclusion

No claim is allowable.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Suryaprabha Chunduru whose telephone number is 571-272-0783. The examiner can normally be reached on 8.30A.M. - 4.30P.M, Mon - Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on 571-272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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).

/Suryaprabha Chunduru/
Primary Examiner, Art Unit 1637

Notice of References Cited	Application/Control No. 14/390,012	Applicant(s)/Patent Under Reexamination ARMOUR ET AL.	
	Examiner Suryaprabha Chunduru	Art Unit 1637	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-2006/0068415 A1	03-2006	Jones; Keith W.	C12Q1/6809	435/6.18
*	B	US-2008/0241831 A1	10-2008	Fan; Jian-Bing	C12Q1/6844	435/6.16
*	C	US-2011/0224105 A1	09-2011	Kurn; Nurith	C12P19/34	506/26
*	D	US-2010/0029511 A1	02-2010	Raymond; Christopher K.	C12Q1/686	506/26
	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	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	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.

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.
13/750,768	01/25/2013	Doug Amorese	25115-755.201	1162

12754 7590 07/15/2015
Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
BERTAGNA, ANGELA MARIE	

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
07/15/2015	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 13/750,768	Applicant(s) AMORESE ET AL.	
	Examiner Angela M. Bertagna	Art Unit 1637	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 *3/16/15*.

☐ 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) *1-3,6,9,10,14,40-42,45-49 and 51-62* 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) *1-3,6,9,10,14,40-42,45-49 and 51-62* is/are rejected.

8)☒ Claim(s) *55* 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 _____ 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

Notice of Pre-AIA or AIA Status

1. The present application is being examined under the pre-AIA first to invent provisions.

Continued Examination Under 37 CFR 1.114

2. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on January 21, 2015 has been entered.

Claims 1-3, 6, 9, 10, 14, 40-42, 45-49, and 51-62 are pending and under examination.

Response to Arguments

3. Applicant's arguments filed on January 21, 2015 have been fully considered.

Rejection of claims 1-3, 6, 9, 11, 14, and 51-53 under pre-AIA 35 U.S.C. 102(b) as being anticipated by Borodina

Applicant's arguments on pages 7-8 have been fully considered. Applicant argues that Borodina fails to teach steps (c) and (e) of claim 1. This argument was persuasive. The rejection has been withdrawn. The rejection has also been withdrawn because Borodina fails to teach step (b) of claim 1. The reference does not disclose ligation of the first adaptor to the nucleic acid fragments. Instead, the first adaptor is ligated to the cDNA generated from the RNA fragments (see the reproduction of Figure 5.1 of Borodina on page 8 of Applicant's arguments).

Rejection of claims 10, 54, and 55 under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Borodina in view of Vaidyanathan

Applicant argues that the rejection should be withdrawn because Vaidyanathan fails to remedy the deficiencies of Borodina with respect to claim 1 (page 12). This argument was persuasive. The rejection has been withdrawn.

Rejection of claims 40-42, 45-49, and 52 under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Costa in view of Kurn

Applicant argues that the references fail to teach or suggest step (b) of claim 40 (pages 9-12). This argument was persuasive. The rejection has been withdrawn.

Double Patenting

4. Applicant is advised that should claim 45 be found allowable, claim 48 will be objected to under 37 CFR 1.75 as being a substantial duplicate thereof. When two claims in an application are duplicates or else are so close in content that they both cover the same thing, despite a slight difference in wording, it is proper after allowing one claim to object to the other as being a substantial duplicate of the allowed claim. See MPEP § 706.03(k).

Claim Objections

5. Claim 55 is objected to because "comprise" should be replaced with "comprises".

Claim Rejections - 35 USC § 112

6. The following is a quotation of 35 U.S.C. 112(d):

> (d) REFERENCE IN DEPENDENT FORMS.—Subject to subsection (e), a claim in dependent form shall contain a reference to a claim previously set forth and then specify a further limitation of the subject matter claimed. A claim in dependent form shall be construed to incorporate by reference all the limitations of the claim to which it refers.

The following is a quotation of 35 U.S.C. 112 (pre-AIA), fourth paragraph:

> Subject to the [fifth paragraph of 35 U.S.C. 112 (pre-AIA)], a claim in dependent form shall contain a reference to a claim previously set forth and then specify a further limitation of the subject matter claimed. A claim in dependent form shall be construed to incorporate by reference all the limitations of the claim to which it refers.

Claims 53 and 59 are rejected under 35 U.S.C. 112(d) or pre-AIA 35 U.S.C. 112, 4th paragraph, as being of improper dependent form for failing to further limit the subject matter of the claim upon which it depends, or for failing to include all the limitations of the claim upon which it depends.

Claim 53 depends from claim 1 and states that the first adaptor sequence is ligated to a 5' end of the nucleic acid fragments. This recitation is not further limiting because steps (c)-(e) of claim 1 already require ligation of the first adaptor to a 5' end of the nucleic acid fragments.

Claim 59 is not further limiting for the same reason.

Applicant may cancel the claim(s), amend the claim(s) to place the claim(s) in proper dependent form, rewrite the claim(s) in independent form, or present a sufficient showing that the dependent claim(s) complies with the statutory requirements.

Claim Rejections - 35 USC § 102

7. 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.

8. Claims 1-3, 6, 9, 10, 14, 40-42, 45-49, 51-54, and 57-62 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Jones et al. (US 2003/0232348 A1).

Regarding claims 1, 14, and 53, Jones teaches a method for enriching for a nucleic acid sequence of interest in a sample that comprises the following steps (see Figure 3 and paragraph 94; see also Example 1 on page 14, paragraphs 129-132):

(a) fragmenting nucleic acids in the sample, wherein the nucleic acid fragments comprise the sequence of interest (paragraphs 94 and 129);

(b) ligating a first adaptor to a 5' end of the nucleic acid fragments (Figure 3 and paragraphs 94 and 129);

(c) annealing one or more oligonucleotides in solution to the nucleic acid sequence of interest in the nucleic acid fragments comprising the ligated first adaptor, wherein the oligonucleotides comprise a 3' portion that is complementary to the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence (Figure 3, paragraph 94, and paragraph 129);

(d) extending said annealed oligonucleotides using a DNA polymerase to obtain extension products comprising sequence complementary to the ligated first adaptor, sequence complementary to the nucleic acid of interest, and the second adaptor (Figure 3, paragraph 94, and paragraph 129); and

(e) amplifying the extension products using a primer that anneals to the complement of the first adaptor sequence and a primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest (Figure 3 and paragraphs 130-131).

Regarding claim 2, Jones teaches that the method further comprises sequencing the enriched nucleic acid sequence of interest (paragraph 132).

Regarding claim 3, Jones teaches that the nucleic acid sequence of interest comprises genomic DNA (paragraph 129). The reference also teaches that the nucleic acid sequence of interest may comprise cDNA or RNA (paragraph 47).

Regarding claims 6 and 52, Jones teaches that the nucleic acid fragments are double-stranded and that method comprises denaturing the nucleic acid fragments before the step of annealing the oligonucleotides (step (c) above), thereby generating single-stranded nucleic acid fragments comprising the first adaptor sequence and the nucleic acid sequence of interest (Figure 3 and paragraph 129).

Regarding claim 9, Jones teaches that the first and second adaptor sequence are distinct (Figure 3, paragraph 94, and paragraphs 129-131, where the A2-A3 adaptor ligated to the fragments is distinct from the A1 adaptor contained in the capture probe oligonucleotide).

Regarding claims 10 and 54, Jones teaches that the second adaptor sequence may comprise a barcode sequence (paragraphs 71-72) and also that the first adaptor may include a restriction site (paragraph 56).

Regarding claim 51, in the method of Jones the oligonucleotides comprise a 3' portion that is complementary to a different nucleic acid sequence of interest in the nucleic acid fragments and a 5' portion comprising the second adapter sequence (paragraph 129; see also Figure 3).

Regarding claims 57-61, Jones teaches that the first adaptor sequence is in a long strand of a partial duplex adaptor comprising a long strand and a short strand, the long strand is ligated

to a 5' end of the nucleic acid fragments, and the short strand is ligated to a 3' end of the nucleic acid fragments (see Figure 3 and paragraph 94).

Regarding claim 62, Jones teaches that the first adaptor sequence is in a ligation strand of a first adaptor, wherein the first adaptor comprises the ligation strand and a non-ligation strand, wherein the ligation strand ligates to a 5' end of the nucleic acid fragments (paragraph 59).

Regarding claim 40, Jones teaches a method for enriching a nucleic acid sequence of interest from a library (see Figure 3, paragraph 94, and Example 3 at paragraphs 129-132), wherein the library comprises nucleic acid inserts comprising the nucleic acid sequence of interest (Figure 3 and paragraphs 94 and 129), wherein the nucleic acid inserts have a first adaptor sequence attached on a first end and a second adaptor sequence attached on a second end (Figure 3 and paragraphs 94 and 129, where the fragments comprising the A2-A3 adaptor at each end are taught), the method comprising:

(a) denaturing the nucleic acid inserts, thereby generating a library of single-stranded nucleic acids with the first and second adaptor sequences flanking the nucleic acid sequence of interest (Figure 3 and paragraph 129);

(b) annealing one or more oligonucleotides in solution to the nucleic acid sequence of interest in the single-stranded nucleic acids, wherein the oligonucleotides comprise a 3' portion that is complementary to the nucleic acid sequence of interest and a 5' portion comprising a third adaptor sequence that is distinct from the first and second adaptor sequences (Figure 3, paragraph 94, and paragraph 129, where the A1-containing oligonucleotides are annealed);

(c) extending said annealed oligonucleotides using a polymerase to obtain extension products comprising sequence complementary to the first adaptor, sequence complementary to

the nucleic acid of interest, and the third adaptor (Figure 3, paragraph 94, and paragraph 129); and

(d) amplifying the extension products using a primer that anneals to the complement of the first adaptor sequence and a primer that anneals to the complement of the third adaptor sequence to enrich for the nucleic acid sequence of interest (Figure 3 and paragraphs 130-131).

Regarding claim 41, Jones teaches that the method further comprises sequencing the enriched nucleic acid sequence of interest (paragraph 132).

Regarding claim 42, Jones teaches that the nucleic acid sequence of interest comprises genomic DNA (paragraph 129). The reference also teaches that the nucleic acid sequence of interest may comprise cDNA or RNA (paragraph 47).

Regarding claims 45 and 48, Jones teaches that the first and second adaptor sequence are the same among each of the inserts in the library (Figure 3 and paragraphs 94 and 129).

Regarding claim 46, Jones teaches that the third adaptor sequence is the same among the one or more oligonucleotides (Figure 3 and paragraphs 94 and 129).

Regarding claim 47, Jones teaches that the first and second adaptor sequence may be distinct from each other (Figure 3 and also paragraph 58).

Regarding claim 49, Jones teaches that the third adaptor sequence may comprise a barcode sequence (paragraphs 71-72).

9. Claims 1-3, 6, 9, 10, 14, 40-42, 45-49, 51-62 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Jones et al. (US 2005/0142577 A1).

Regarding claims 1, 2, 14, and 53, Jones teaches a method for enriching a nucleic acid sequence of interest in a sample comprising nucleic acids that comprises the following steps (see, for example, Figures 3-4 and paragraphs 86-90, 96, 99, and 100): (a) fragmenting nucleic acids in the sample to generate fragments comprising the sequence of interest, (b) ligating a first adaptor to a 5' end of the nucleic acid fragments, (c) annealing one or more oligonucleotides in solution to the nucleic acid sequence of interest in the fragments, wherein said oligonucleotides comprise a 3' portion that is complementary to the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence, (d) extending said annealed oligonucleotides using a DNA polymerase to obtain extension products comprising sequence complementary to the ligated first adaptor, sequence complementary to the nucleic acid of interest, and the second adaptor, (e) amplifying the extension products using a primer that anneals to the complement of the first adaptor sequence and a primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest, and (f) sequencing the enriched nucleic acid of interest.

Regarding claim 3, Jones teaches that the nucleic acid sequence of interest comprises genomic DNA (paragraph 100). The reference also teaches that the nucleic acid sequence of interest may comprise RNA (paragraphs 43-44).

Regarding claims 6 and 52, Jones teaches that the nucleic acid fragments are double-stranded and that method comprises denaturing the nucleic acid fragments before the step of annealing the oligonucleotides (step (c) above), thereby generating single-stranded nucleic acid fragments comprising the first adaptor sequence and the nucleic acid sequence of interest (Figure 3 and paragraphs 87 and 100).

Regarding claim 9, Jones teaches that the first and second adaptor sequence are distinct (Figure 3, paragraph 87, and paragraph 100).

Regarding claims 10 and 54, Jones teaches that the second adaptor sequence may comprise a barcode sequence (paragraphs 58 and 61) and also that the first adaptor may include a restriction site (paragraph 49).

Regarding claim 51, in the method of Jones, the oligonucleotides comprise a 3' portion that is complementary to a different nucleic acid sequence of interest in the nucleic acid fragments and a 5' portion comprising the second adapter sequence (paragraphs 87, 99, and 100; see also Figure 3).

Regarding claim 55, Jones teaches that the 3' portion of the oligonucleotides comprises a random sequence (see Figure 4, paragraph 100 and paragraphs 88-90).

Regarding claims 57-61, Jones teaches that the first adaptor sequence is in a long strand of a partial duplex adaptor comprising a long strand and a short strand, the long strand is ligated to a 5' end of the nucleic acid fragments, and the short strand is ligated to a 3' end of the nucleic acid fragments (see Figure 3 and paragraph 87).

Regarding claim 62, Jones teaches that the first adaptor sequence is in a ligation strand of a first adaptor, wherein the first adaptor comprises the ligation strand and a non-ligation strand, wherein the ligation strand ligates to a 5' end of the nucleic acid fragments (paragraph 50).

Regarding claim 40, Jones teaches a method for enriching a nucleic acid sequence of interest from a library (see Figure 3, paragraphs 86-90, 99, and 100), wherein the library comprises nucleic acid inserts comprising the nucleic acid sequence of interest (Figure 3 and paragraphs 86 and 100), wherein the nucleic acid inserts have a first adaptor sequence attached

on a first end and a second adaptor sequence attached on a second end (Figure 3 and paragraphs 86 and 100), the method comprising:

(a) denaturing the nucleic acid inserts, thereby generating a library of single-stranded nucleic acids with the first and second adaptor sequences flanking the nucleic acid sequence of interest (Figure 3 and paragraphs 86 and 100);

(b) annealing one or more oligonucleotides in solution to the nucleic acid sequence of interest in the single-stranded nucleic acids, wherein the oligonucleotides comprise a 3' portion that is complementary to the nucleic acid sequence of interest and a 5' portion comprising a third adaptor sequence that is distinct from the first and second adaptor sequences (Figure 3, paragraph 86, and paragraph 100, where the oligonucleotides containing an A1 or degenerate sequence are annealed);

(c) extending said annealed oligonucleotides using a polymerase to obtain extension products comprising sequence complementary to the first adaptor, sequence complementary to the nucleic acid of interest, and the third adaptor (Figure 3, paragraph 86, and paragraph 100); and

(d) amplifying the extension products using a primer that anneals to the complement of the first adaptor sequence and a primer that anneals to the complement of the third adaptor sequence to enrich for the nucleic acid sequence of interest (Figure 3 and paragraphs 86 and 100).

Regarding claim 41, Jones teaches that the method further comprises sequencing the enriched nucleic acid sequence of interest (paragraph 100).

Regarding claim 42, Jones teaches that the nucleic acid sequence of interest comprises genomic DNA (paragraph 100). The reference also teaches that the nucleic acid sequence of interest may comprise RNA (paragraphs 43-44).

Regarding claims 45 and 48, Jones teaches that the first and second adaptor sequence are the same among each of the inserts in the library (Figure 3 and paragraphs 86 and 100).

Regarding claim 46, Jones teaches that the third adaptor sequence is the same among the one or more oligonucleotides (Figure 3 and paragraphs 86 and 100).

Regarding claim 47, Jones teaches that the first and second adaptor sequence may be distinct from each other (Figure 3 and also paragraph 53).

Regarding claim 49, Jones teaches that the third adaptor sequence may comprise a barcode sequence (paragraph 58).

Regarding claim 56, Jones teaches that the 3' portion of the oligonucleotides comprises a random sequence (paragraph 100 and paragraphs 88-90).

Claim Rejections - 35 USC § 103

10. The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made

absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

11. Claims 55 and 56 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Jones et al. (US 2003/0232348 A1; Jones 1 below) in view of Jones et al. (US 2005/0142577 A1; Jones 2 below).

As discussed above, Jones 1 teaches the method of claims 1-3, 6, 9, 10, 14, 40-42, 45-49, 51-54, and 57-62.

Jones 1 does not teach that the 3' portion of the one or more oligonucleotides comprises a random sequence as required by claims 55 and 56.

However, Jones 2 teaches that incorporating a degenerate sequence into the 3' portion of oligonucleotides used for the same purpose as the oligonucleotides of Jones 1 is advantageous because the degenerate sequence "may improve efficiency and reduce the cost of locus-specific genotyping applications" (paragraph 88). The reference further teaches that the degenerate sequence may include a random sequence (paragraphs 89-90).

It would have been *prima facie* obvious for one of ordinary skill in the art at the time of the invention to practice the method of Jones 1 using oligonucleotides having a 3'portion comprising a random sequence. The ordinary artisan would have been motivated to do so to obtain a more efficient and less expensive method as taught by Jones 2 and would have had a

reasonable expectation of success since the oligonucleotides of Jones 2 were used for the same purpose as the oligonucleotides of Jones 1. Thus, the methods of claims 55 and 56 are *prima facie* obvious.

Conclusion

12. No claims are currently allowable.

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

(1) Rabinowitz et al. (US 2011/0288780 A1) is cited for its disclosure on page 22, paragraphs 213-222 of a method comprising (i) forming a library of DNA fragments with universal adaptors on both ends, (ii) amplifying the library with universal primers, (iii) performing multiplexed target-specific amplification, (iv) barcoding the resulting amplification products, and (v) sequencing the barcoded amplification products.

(2) Rabinowitz et al. (US 2012/0270212 A1) is cited for its disclosure in Figure 6 and paragraphs 316 and 322 of a method comprising (i) forming a library of DNA fragments with adaptors on both ends, (ii) amplifying the library using an adaptor-specific primer and a target-specific primer, and (iii) further amplifying the resulting amplification products.

(3) Link et al. (US 2012/0309002 A1) is cited for its disclosure in Figures 3-5 of a method comprising (i) ligation of adaptors to both strands of a double-stranded nucleic acid, (ii) hybridization and extension of an oligonucleotide comprising a 3' target-specific portion and a 5' adaptor sequence, and amplification of the resulting extension products using a first primer that

targets an adaptor present in step (i) and a second primer that targets the adaptor added in step (ii).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela M. Bertagna whose telephone number is (571)272-8291. The examiner can normally be reached on Monday-Friday, 9-5.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on (571)272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/Angela M. Bertagna/
Primary Examiner, Art Unit 1637

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.
15/471,785	03/28/2017	Doug AMORESE	25115-755.301	8042

12754 7590 08/18/2017
Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER
BERTAGNA, ANGELA MARIE

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
08/18/2017	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 15/471,785	Applicant(s) AMORESE ET AL.	
	Examiner Angela M. Bertagna	Art Unit 1637	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 *7/10/17*.

☐ 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) *51-75* 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) *51-75* is/are rejected.

8)☒ Claim(s) *51 and 60* 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 *3/28/17* 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

Notice of Pre-AIA or AIA Status

1. The present application is being examined under the pre-AIA first to invent provisions.

Preliminary Amendment

2. Applicant's submission of a preliminary amendment on July 10, 2017 is acknowledged. Claims 51-75 are pending.

It is noted that declaration evidence concerning the method recited in the instant claims 51-75 was presented in the parent application. Since such evidence does not automatically become part of the instant application, Applicant is advised to submit a copy the declaration evidence and any supporting arguments if consideration of said evidence is desired.

Information Disclosure Statement

3. The information disclosure statement filed on July 10, 2017 fails to comply with the provisions of 37 CFR 1.97, 1.98 and MPEP § 609 because the copending US applications listed in citations 317-332 do not include the inventor(s). See 37 CFR 1.98. These references have not been considered.

As well, foreign patent documents 69, 74, 81, 82, 97, 102, 104, and 109 and non-patent literature documents 9, 34, 161, 175, 256, and 362 have not been considered because the copy required by 37 CFR 1.98(a)(2) has not been provided in the instant case or prior-filed Application Serial No. 13/750,768.

Applicant is advised that the date of any re-submission of any item of information contained in this information disclosure statement or the submission of any missing element(s) will be the date of submission for purposes of determining compliance with the requirements based on the time of filing the statement, including all certification requirements for statements under 37 CFR 1.97(e). See MPEP § 609.05(a).

Specification

4. The specification is objected to because the continuity information in the first paragraph should be updated to state that prior-filed Application Serial No. 13/750,768 has been issued as U.S. Patent No. 9,650,628.

Claim Objections

5. Claim 51 is objected to because it contains a typographical error in line 1 of step (a) where "partial duplex adaptor ," is recited.

Claim 60 is objected to because of the following informalities: Amending the concluding clause to state that each barcode is unique is suggested since doing so would be more succinct.

Claim Rejections - 35 USC § 112

6. 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.

Claims 51-64 and 66-75 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.

Claim 51 is indefinite because the metes and bounds of "appended to" in step (a) are not clear. The specification states that ligation is encompassed by the term, but it is not clear what else is encompassed by the term "appended to". Neither the specification nor the claims gives any additional description as to what is encompassed by "appended to" other than ligation. It is unclear, for example, if "appended to" also encompasses partial duplex adaptors that are merely hybridized to the nucleic acid fragment. Claims 60, 64, and 66-68 are also indefinite for the same reason (i.e., the claims recite "appended to", the metes and bounds of which are unclear).

Claims 52-59, 61-63, and 69-75 are also indefinite since they depend from claim 51.

Claim Rejections - 35 USC § 103

7. The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to

point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

9. Claims 51, 52, 54-67, 69-71, and 75 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Jones et al. (US 2005/0142577 A1; IDS reference) in view of Tucker et al. (The American Journal of Human Genetics 2009; **85:** 142-154; IDS reference).

Regarding claims 51, 63-67, 69, and 75, Jones teaches a method for enriching a nucleic acid sequence of interest in a sample comprising nucleic acids that comprises the following steps (see, for example, Figures 3-4 and paragraphs 86-90, 96, 99, and 100): (a) fragmenting nucleic acids in the sample to generate fragments comprising the sequence of interest; (b) ligating a partial duplex adaptor to each 5' and each 3' end of the nucleic acid fragments, wherein the partial duplex adaptor comprises a first long strand containing a first adaptor sequence and second shorter strand; (c) annealing one or more oligonucleotides in solution to the nucleic acid sequence of interest in the fragments, wherein said oligonucleotides comprise a 3' portion with at least ten bases that are complementary to the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence; (d) extending said annealed oligonucleotides using a DNA polymerase, thereby obtaining extension products comprising sequence complementary to the ligated first adaptor, sequence complementary to the nucleic acid of interest, and the second adaptor; (e) amplifying the extension products using a primer that anneals to the complement of the first adaptor sequence and a primer that anneals to the complement of the second adaptor

sequence to enrich for the nucleic acid sequence of interest; and (f) sequencing the enriched nucleic acid of interest.

Regarding claims 52 and 54, Jones teaches that the nucleic acid sequence of interest comprises genomic DNA (paragraph 100) or RNA (paragraphs 43-44).

Regarding claim 55, Jones teaches that the first and second adaptor sequences are distinct from each other (Figure 3, paragraph 87, and paragraph 100).

Regarding claim 56, Jones teaches using a plurality of oligonucleotide to practice the capture and extension steps, and the plurality of oligonucleotides comprises a 3' portion that is complementary to a different nucleic acid sequence of interest in the nucleic acid fragments and a 5' portion comprising the same second adapter sequence (paragraphs 87, 99, and 100; see also Figure 3).

Regarding claim 57, Jones teaches that the second adaptor sequence may comprise a barcode sequence (paragraphs 58 and 61).

Regarding claim 60, as noted above, Jones teaches using a plurality of different oligonucleotides, each designed to capture a different nucleic acid sequence of interest and containing the same 5' adaptor (paras. 87, 99, and 100; see also Figure 3). Jones also teaches that the second adaptor may include a barcode sequence (paras. 58 & 61). Therefore, Jones teaches a plurality of different oligonucleotides, each designed to capture a different nucleic acid sequence of interest and containing the same 5' adaptor and a unique barcode.

Regarding claims 61 and 62, Jones teaches that the nucleic acid fragments are double-stranded and that method comprises denaturing the nucleic acid fragments before the step of annealing the oligonucleotides (step (c) above), thereby generating single-stranded nucleic acid

fragments comprising the first adaptor sequence and the nucleic acid sequence of interest (Figure 3 and paragraphs 87 and 100).

Jones does not teach massively parallel sequencing as required by independent claim 51. The reference teaches that nucleic acids amplified by the disclosed method may be analyzed by array hybridization or any other method known in the art to obtain sequence information about a region of interest (paragraph 79-80), but massively parallel sequencing is not disclosed.

Tucker describes methods of obtaining sequence information about a region of interest in a nucleic acid, and specifically describes a number of massively parallel sequencing methods (abstract, pages 142-146, and Figures 2-5). The reference states that "Massively parallel sequencing will allow simultaneous screening for mutations in hundreds of loci in genetically heterogeneous disorders, whole-genome screening for novel mutations, and sequence-based detection of novel pathogens that cause human disease. In addition, massively parallel sequencing will permit clinical application of our expanding knowledge of pharmacogenetics, cancer genetics, epigenetics, and complex traits" (page 142, column 2). The reference further teaches that massively parallel sequencing "has reduced the cost and increased the throughput of genomic sequencing by more than three orders of magnitude...and may also replace arrays and Sanger sequencing in clinical applications where they are currently being used" (abstract).

Further regarding claims 51, 64-67, 70, 71, and 75, Tucker describes a massively parallel sequencing method developed by Illumina (Fig. 2). The method comprises bridge amplification and the use of four labeled reversible terminators (Fig. 2). The Illumina method also comprises attaching an adaptor to each end of a nucleic acid fragment of interest and using the adaptors to

hybridize the nucleic acid of interest to flow cell sequences (Fig. 2). Each of the two adaptors is also complementary to a sequencing primer (Fig. 2).

It would have been *prima facie* obvious for one of ordinary skill in the art at the time of the invention to analyze the amplification products generated by the method of Jones using one of the massively parallel sequencing methods described in Tucker. The ordinary artisan would have been motivated to do so since Tucker taught that such methods offered very high-throughput sequencing (abstract and page 142) and as well as the ability to accurately detect minor alleles (page 144, column 2). Additional motivation is provided in the abstract of Tucker, which teaches that massively parallel sequencing methods may replace array methods. The ordinary artisan would have had a reasonable expectation of success in view of the guidance provided by Tucker concerning the use of massively parallel sequencing methods and also since Jones taught that the amplification products could be analyzed using any desired method. In using the Illumina sequencing method, for example, the ordinary artisan would have recognized that the 3'end of the amplification products should be complementary to a surface-immobilized sequence (i.e., as recited in claim 51) and, accordingly, would have been motivated to design the adaptors such that they are complementary to sequences on a support to be used in the Illumina sequencing method. The ordinary artisan would have had a reasonable expectation of success since the Illumina method has been publicly disclosed and commercialized.

Further regarding claims 58 and 59, the teachings of Jones concerning barcodes do not include specific lengths for said barcodes (see paras. 58 and 61 of Jones), but the reference teaches that the preferred length for the adaptors in which the barcodes are contained preferably ranges from at least five to no more than 60 bases in length (para. 49). The length range of at

least eight nucleotides overlaps with the ranges disclosed in Jones, and the specific length of 12 nucleotides lies within the ranges of Jones. As well, no evidence of unexpected results with respect to the claimed barcode lengths has been presented. Therefore, a *prima facie* case of obviousness exists with respect to the barcode lengths recited in claims 58 and 59 (MPEP 2144.05 I).

Thus, the methods of claims 51, 52, 54-67, 69-71, and 75 are *prima facie* obvious.

10. Claims 53, 68, 73, and 74 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Jones et al. (US 2005/0142577 A1; IDS reference) in view of Tucker et al. (The American Journal of Human Genetics 2009; **85:** 142-154; IDS reference) and further in view of Gormley et al. (US 7,741,463 B2; IDS reference).

As discussed above, the teachings of Jones in view of Tucker render obvious the methods of claims 51, 52, 54-67, 69-71, and 75.

Regarding claim 53, Jones does not teach that the nucleic acid fragment comprises cDNA. As a result, Jones also does not teach reverse transcription of RNA to generate the nucleic acid fragment of interest as required by claim 74. Tucker does not remedy these deficiencies in Jones.

Regarding claim 68 and 73, neither Jones nor Tucker teaches purifying the nucleic acid fragment appended to the partial duplex adaptor or end repair of nucleic acid fragments prior to appending said adaptor. Tucker does not remedy these deficiencies in Jones.

Gormley describes a method for preparing a library for sequencing that comprises fragmenting nucleic acids contained in a sample, purifying the fragments, end-repair of the

purified fragments, ligation of a partial duplex adaptor, and PCR (see the working example at columns 23-26 and especially column 23, lines 25-67, column 25, line 41 – column 26, line 20, and column 26, lines 25-62; see also Figs. 2a-2b). Gormley further teaches using sequencing the nucleic acids in the library using a massively parallel method comprising the steps disclosed in Figure 2 of Tucker (i.e., bridge amplification and the use of reversible terminators) (column 20, line 18 – column 22, line 20). Further regarding claims 53 and 74, Gormley teaches that libraries may be prepared from cDNA fragments (col. 6, ll. 5-27).

It would have been *prima facie* obvious for one of ordinary skill in the art at the time of the invention to further include purification and end repair steps between the fragmentation and adaptor ligation steps when practicing the method suggested by Jones in view of Tucker. The teachings of Gormley cited above provide motivation to do so since they indicate that these steps are a useful part of a method for preparing a library of nucleic acids for massively parallel sequencing. Gormley also provides a reasonable expectation of success since the teachings cited above contain a detailed description of how to purify nucleic acid fragments and subject the purified fragments to end repair. Thus, the methods of claims 68 and 73 are *prima facie* obvious.

It also would have been *prima facie* obvious for one of ordinary skill in the art at the time of the invention to practice the method suggested by Jones in view of Tucker using cDNA fragments of interest prepared from RNA by reverse transcription. Gormley provides motivation and a reasonable expectation of success since the teachings cited above indicate that cDNA libraries are a desirable template for massively parallel sequencing. Since it was routine and conventional to use reverse transcription to prepare cDNA, the teachings of Gormley cited above

also render obvious the method of claim 74. Thus, the methods of claims 53 and 74 are *prima facie* obvious.

11. Claim 72 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Jones et al. (US 2005/0142577 A1; IDS reference) in view of Tucker et al. (The American Journal of Human Genetics 2009; **85:** 142-154; IDS reference) and further in view of Rothberg et al. (Nature 2011; **475:** 348-352 + Supplementary Information; IDS reference).

As discussed above, the teachings of Jones in view of Tucker render obvious the methods of claims 51, 52, 54-67, 69-71, and 75.

Jones does not teach semiconductor sequences as recited in claim 72, and Tucker does not remedy this deficiency. Rothberg, however, describes massively parallel sequencing using a semiconductor device (Figure 1 and pages 348-350) and discloses that this sequencing method is highly robust, scalable, and advantageously does not require the use of optics or labeled nucleotides (abstract and page 348).

It would have been *prima facie* obvious for one of ordinary skill in the art at the time of the invention to select semiconductor sequencing as described by Rothberg as the massively parallel sequencing method when practicing the method of Jones in view of Tucker. Rothberg provides motivation to do so by stating that the disclosed semiconductor sequencing method is robust, highly scalable, and requires neither optical detection nor modified nucleotides (abstract and page 348). The ordinary artisan would have had a reasonable expectation of success in view of Rothberg's guidance concerning practice of the method (pages 349-350 and the Supplemental Methods). Thus, the method of claim 72 is *prima facie* obvious.

Double Patenting

12. The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory double patenting rejection is appropriate where the conflicting claims are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on nonstatutory double patenting provided the reference application or patent either is shown to be commonly owned with the examined application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement. See MPEP § 717.02 for applications subject to examination under the first inventor to file provisions of the AIA as explained in MPEP § 2159. See MPEP §§ 706.02(l)(1) - 706.02(l)(3) for applications not subject to examination under the first inventor to file provisions of the AIA. A terminal disclaimer must be signed in compliance with 37 CFR 1.321(b).

The USPTO Internet website contains terminal disclaimer forms which may be used. Please visit www.uspto.gov/patent/patents-forms. The filing date of the application in which the form is filed determines what form (e.g., PTO/SB/25, PTO/SB/26, PTO/AIA/25, or PTO/AIA/26) should be used. A web-based eTerminal Disclaimer may be filled out completely online using web-screens. An eTerminal Disclaimer that meets all requirements is auto-processed and approved immediately upon submission. For more information about eTerminal Disclaimers, refer to www.uspto.gov/patents/process/file/efs/guidance/eTD-info-I.jsp.

13. Claims 51-57, 60-72, 74, and 75 are rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1-22 of U.S. Patent No. 9,650,628 B2. Although the claims at issue are not identical, they are not patentably distinct from each other because the patented claims teach or suggest all of the limitations contained in the instant claims. Each of the instant claims is discussed in relation to the claims of the '628 patent in more detail below.

The instant claim 51 is drawn to a method for sequencing an enriched nucleic acid of interest. The method comprises: (a) obtaining a nucleic acid fragment appended to a partial duplex adaptor containing a first adaptor sequence and having a long first strand and shorter second strand; (b) annealing one or more oligonucleotides to the nucleic acid sequence of interest, wherein said oligonucleotides have a 3' portion at least ten nucleotides in length that binds to the nucleic acid sequence of interest and a 5’ tail portion that comprises a second adaptor and is not complementary to the nucleic acid sequence of interest; (c) extending the oligonucleotides of step (b) to form extension products having a 5’ second adaptor sequence, a sequence complementary to the nucleic acid sequence of interest, and a sequence complementary

to the first sequence in the partial duplex adaptor; (d) amplifying the extension products of step (c) using a first primer that anneals to a complement of the first adaptor sequence and a second primer that anneals to a complement of the second adaptor sequence to enrich for the sequence of interest and produce an amplification product with a 3' end that is complementary to a sequence on a surface; (e) hybridizing the amplification products to said sequence on a surface, and (f) sequencing the enriched nucleic acid of interest on a massively parallel sequencing platform.

Claims 1 and 12-14 of the ‘628 patent are also drawn to a method for sequencing an enriched nucleic acid of interest that includes all of the elements of the instant claim 51. Claim 1 of the ‘628 patent includes the elements of steps (b)-(f) of the instant claim 1. Claim 1 of the '628 patent does not include the structural features of the adaptor set forth in the instant claim 51, but these elements are disclosed in claims 12-14 of the '628 patent. Therefore, the method set forth in the instant claim 51 clearly overlaps in scope with and is an obvious variant of the methods set forth in the '628 patent.

The limitations of the instant claims 52-54 are recited in claim 2 of the '628 patent.

The limitations of the instant claims 55-57 are recited in claims 3, 4, and 7 of the ‘628 patent.

The limitations of the instant claim 60 are suggested by claims 4 and 7 of the '628 patent. Claim 7 teaches the use of a plurality of different oligonucleotides, each designed to capture a different nucleic acid sequence of interest and containing the same 5' adaptor. Claim 4 teaches that the second adaptor may include a barcode sequence. These claims in combination suggest a plurality of different oligonucleotides, each designed to capture a different nucleic acid sequence of interest and containing the same 5' adaptor and a unique barcode.

The limitations of the instant claims 61 and 62 are recited in claims 8-9 of the '628 patent.

The limitations of the instant claim 63 are recited in claims 5 and 6 of the '628 patent.

The limitations of the instant claims 64-68 are recited in claims 1, 15, and 16 of the '628 patent.

The limitations of the instant claim 69 are recited in claim 12 of the ‘628 patent.

The limitations of the instant claims 70-72 are recited in claims 17-19 of the '628 patent.

The instant claim 74 requires using a nucleic acid fragment obtained by reverse transcription of RNA. This is clearly suggested by claim 2 of the ‘628 patent, which teaches using cDNA fragments.

The instant claim 75 requires the partial duplex adaptor to be appended to each end of the nucleic acid fragment of interest. This is clearly suggested by claims 1, 15, and 16 of the '628 patent, which state that the adaptor is ligated to an end of the fragment. This language in the claims of the '628 patent encompasses ligation to one or both ends. Therefore, the ordinary artisan would have been motivated to select either option with a reasonable expectation of success.

Thus, the instant claims 51-57, 60-72, 74, and 75 are not patentably distinct from claims 1-22 the ‘628 patent.

14. Claims 58 and 59 are rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1-22 of U.S. Patent No. 9,650,628 B2 in view of Jones et al. (US 2005/0142577 A1; IDS reference).

As discussed above, the instant claims 51-57, 60-72, 74, and 75 are not patentably distinct from claims 1-22 the ‘628 patent.

Regarding the instant claims 58 and 59, the claims of the ‘628 patent do not include any specific lengths for the barcodes.

The teachings of Jones are discussed in greater detail in the obviousness rejections set forth above. The reference describes a method that is quite similar to the method recited in the instant claims and the claims of the ‘628 patent and discloses that adaptors useful in said method preferably range from at least five to no more than 60 bases in length (para. 49). Jones additionally teaches that said adaptors may include barcodes (paras. 58 and 61).

It would have been *prima facie* obvious for the ordinary artisan practicing the method set forth in the claims of the '628 patent to select a barcode length of 12 nucleotides as recited in the instant claim 59 or a barcode length of at least eight nucleotides as recited in the instant claim 58. As discussed above, Jones teaches a similar method in which the preferred length for the adaptors is disclosed. The claimed length range of at least eight nucleotides overlaps with the ranges disclosed in Jones, and the specific length of 12 nucleotides lies within the ranges of Jones. As well, no evidence of unexpected results with respect to the claimed barcode lengths has been presented. Therefore, a *prima facie* case of obviousness exists with respect to the barcode lengths recited in the instant claims 58 and 59 (MPEP 2144.05 I), and these claims are not patentably distinct from the claims of the ‘628 patent in view of Jones.

15. Claim 73 is rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1-22 of U.S. Patent No. 9,650,628 B2 in view of Gormley et al. (US 7,741,463 B2; IDS reference)

As discussed above, the instant claims 51-57, 60-72, 74, and 75 are not patentably distinct from claims 1-22 the ‘628 patent.

End repair as recited in the instant claim 73 is not present in the claims of the '628 patent.

Gormley describes a method for preparing a library for sequencing that comprises fragmenting nucleic acids contained in a sample, purifying the fragments, end-repair of the purified fragments, ligation of a partial duplex adaptor, and PCR (see the working example at columns 23-26 and especially column 23, lines 25-67, column 25, line 41 – column 26, line 20, and column 26, lines 25-62; see also Figs. 2a-2b). Gormley further teaches using sequencing the nucleic acids in the library using a massively parallel method comprising bridge amplification and the use of reversible terminators as recited in claims 17 and 18 of the ‘628 patent (column 20, line 18 – column 22, line 20).

It would have been *prima facie* obvious for the ordinary artisan practicing the method recited in the claims of the '628 patent to include an end repair step between the fragmentation and adaptor ligation steps. The teachings of Gormley cited above provide motivation to do so since they indicate that this step is a useful part of a method for preparing a library of nucleic acids for massively parallel sequencing. Gormley also provides a reasonable expectation of success since the teachings cited above contain a detailed description of how to purify nucleic acid fragments and subject the purified fragments to end repair. Thus, the instant claim 73 is not patentably distinct from the claims of the '628 patent in view of Gormley.

Conclusion

16. No claims are currently allowable.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela M. Bertagna whose telephone number is (571)272-8291. The examiner can normally be reached on Monday-Friday, 9-5.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on (571)272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/Angela M. Bertagna/
Primary Examiner, Art Unit 1637

Notice of References Cited	Application/Control No. 15/471,785	Applicant(s)/Patent Under Reexamination AMORESE ET AL.	
	Examiner Angela M. Bertagna	Art Unit 1637	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-9,650,628 B2	05-2017	Amorese; Doug	C12N15/1068	1/1
	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	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	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.

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.
14/836,936	08/26/2015	Doug Amorese	25115-755.202	2434

12754 7590 08/22/2017

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER
BERTAGNA, ANGELA MARIE

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
08/22/2017	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 14/836,936	Applicant(s) AMORESE ET AL.	
	Examiner Angela M. Bertagna	Art Unit 1637	AIA (First Inventor to File) Status Yes

-- 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 *2/16/16*.

☐ 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) *1-23* 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) *1-23* is/are rejected.

8)☒ Claim(s) *10,13 and 14* 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 *8/26/15* 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

Notice of Pre-AIA or AIA Status

1. The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

Preliminary Amendment

2. Applicant's submission of a preliminary amendment on February 16, 2016 is acknowledged. Claims 1-23 are pending.

Information Disclosure Statement

3. The information disclosure statement filed on July 10, 2017 fails to comply with the provisions of 37 CFR 1.97, 1.98 and MPEP § 609 because the copending US applications listed in citations 317-332 do not include the inventor(s). See 37 CFR 1.98. These references have not been considered.

As well, foreign patent documents 69, 74, 81, 82, 97, 102, 104, and 109 and non-patent literature documents 9, 34, 161, 175, 256, and 362 in the IDS filed on July 10, 2017 have not been considered because the copy required by 37 CFR 1.98(a)(2) has not been provided in the instant case or prior-filed Application Serial No. 13/750,768. The references lined through on the

IDS filed on November 9, 2015 have also not been considered because the copy required by 37 CFR 1.98(a)(2) has not been provided.

Applicant is advised that the date of any re-submission of any item of information contained in this information disclosure statement or the submission of any missing element(s) will be the date of submission for purposes of determining compliance with the requirements based on the time of filing the statement, including all certification requirements for statements under 37 CFR 1.97(e). See MPEP § 609.05(a).

Specification

4. The disclosure is objected to because it contains an embedded hyperlink and/or other form of browser-executable code on pages 41, 65, and 69. Applicant is required to delete the embedded hyperlink and/or other form of browser-executable code. See MPEP 608.01.

Drawings

5. The drawings are objected to as failing to comply with 37 CFR 1.84(p)(4) because several reference characters have been used to designate the same part.

In Figure 10, reference characters "1003" and "1005" have been used to designate both "fragments with an adaptor ligated at the 5' end" (p. 69, para. 249) and "an oligonucleotide extension product" (p. 69, para. 249). Also, reference characters "1007" and "1009" have been used to designate both "primer B" (p. 69, para. 249) and "an oligonucleotide extension product" (p. 69, para. 249).

In Figure 11, reference character "1103" has been used to designate "the 5' end", "nucleic acid fragments", "a polymerase", "an oligonucleotide extension product", and "a first end" (p. 69, para. 250).

In Figure 12, reference character "1205" is used to designate "a sequence complementary to the transposon sequence at the 3' end" and "a polymerase". Also, reference character "1207" is used to designate "a first end" and "an oligonucleotide extension product" (p. 70, para. 251).

In Figure 14, two different reference characters are used for the same part. Therefore, the figure does not comply with 37 CFR 1.84(p)(4). Specifically, reference characters "1403" and "1405" are each used to designate "a first oligonucleotide extension product" (p. 70, para. 253).

Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. 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). If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Duplicate Claims

6. Applicant is advised that should claim 16 be found allowable, claim 18 will be objected to under 37 CFR 1.75 as being a substantial duplicate thereof. When two claims in an application are duplicates or else are so close in content that they both cover the same thing, despite a slight

difference in wording, it is proper after allowing one claim to object to the other as being a substantial duplicate of the allowed claim. See MPEP 706.03(k).

Claim Objections

7. Claim 10 is objected to because "a 3' end" in line 3 should be replaced with "the 3' end" since the oligonucleotide extension product has only one 3' end.

Claim 13 is objected to because the word "a" should be inserted before "barcode sequence" in line 2.

Claim 14 is objected to because the word "a" should be inserted before "sequence that can anneal" in line 2 and also before "complement" in line 3.

Claim Rejections - 35 USC § 112

8. 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.

Claims 1 and 3-23 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.

Claim 1 is indefinite because the relationship between the nucleic acid fragment recited in step (a) and the sample comprising nucleic acids recited in the preamble is not clearly stated. As written, the claim does not require the fragment to be generated from or present in the sample

containing nucleic acids, but this does not appear to be Applicant's intention. Amending the claim to introduce the fragmentation step recited in claim 2 would address the issue.

Claims 3-23 are also indefinite since they depend from claim 1.

Claim 7 is further indefinite because denaturing prior to the extension step (i.e., step b) of claim 1) does not necessarily render a double-stranded nucleic acid single-stranded prior to the annealing step (i.e., step a) of claim 1) as required by the subsequently recited "wherein" clause. Amending claim 7 to replace "step b)" with "step a)" would address the issue.

Claim 9 is further indefinite because the requirements of "designed to anneal to the nucleic acid sequence of interest" are not clear. This language encompasses any oligonucleotide meeting the requirements set forth in step a) of claim 1 because oligonucleotides with a 3' portion containing a sequence complementary to the nucleic acid sequence of interest are necessarily "designed to anneal to the nucleic acid sequence of interest". This is the case regardless of whether the complementary sequence is or includes a random portion. As a result, it is not clear what is additionally required by the language in claim 9.

Claim 20 is further indefinite because the intended meaning of the claim is not clear. The claim states that "the nucleic acid sequence of interest comprises a gene". This indicates that the nucleic acid sequence of interest is quite large (i.e., the size of an entire gene), but it is not clear that Applicant intends the claim to be so narrow. It is not clear, for example, that Applicant did not intend the claim to additionally encompass nucleic acid sequences of interest that are simply present in or part of a gene. Based on the disclosure as a whole, it appears that the latter was intended. Clarification is requested.

Claim Rejections – 35 USC § 112

9. The following is a quotation of 35 U.S.C. 112(d):

> (d) REFERENCE IN DEPENDENT FORMS.—Subject to subsection (e), a claim in dependent form shall contain a reference to a claim previously set forth and then specify a further limitation of the subject matter claimed. A claim in dependent form shall be construed to incorporate by reference all the limitations of the claim to which it refers.

The following is a quotation of pre-AIA 35 U.S.C. 112, fourth paragraph:

> Subject to the following paragraph [i.e., the fifth paragraph of pre-AIA 35 U.S.C. 112], a claim in dependent form shall contain a reference to a claim previously set forth and then specify a further limitation of the subject matter claimed. A claim in dependent form shall be construed to incorporate by reference all the limitations of the claim to which it refers.

Claims 9, 10, and 14 are rejected under 35 U.S.C. 112(d) or pre-AIA 35 U.S.C. 112, 4th paragraph, as being of improper dependent form for failing to further limit the subject matter of the claim upon which it depends, or for failing to include all the limitations of the claim upon which it depends.

Claim 9 depends from claim 1 and states that "the 3' portion of the oligonucleotide comprises a sequence designed to anneal to the nucleic acid sequence of interest". This is not further limiting because it is already required by claim 1. Since step a) of claim 1 requires the 3' portion of the oligonucleotide to contain a sequence that is complementary to the sequence of interest, it is inherent that said sequence was "designed to anneal to the nucleic acid sequence of interest" as recited in claim 9. There are two ways to obtain the sequence complementary to the sequence of interest recited in claim 1. One way is to incorporate a random sequence. The other is to use a non-random sequence complementary to the sequence of interest. Each of these oligonucleotides, however, is designed to anneal to the nucleic acid sequence of interest. Therefore, the claim is not further limiting.

Claim 10 depends from claim 1 and requires ligating the sequence complementary to the second adaptor sequence to a 3' end of the oligonucleotide extension product. This is not further limiting because it is already required by claim 1. Steps a) and b) of claim 1 require extension of an oligonucleotide with a 5' first adaptor sequence to generate an extension product having a 5' first adaptor sequence followed by a 3' sequence complementary to the nucleic acid sequence of interest. Step c) then requires ligating a sequence complementary to the second adaptor sequence at a second end. This limitation as well as the limitations in step d) concerning the amplification primers clearly requires ligation to the 3' end of the oligonucleotide extension product.

Claim 14 depends from claim 1 and requires the first and/or second adaptor to include a sequence that can anneal to a sequencing primer or a complement of a sequence that can anneal to a sequencing primer. This is not further limiting because it is an inherent property of any first and second adaptor. Any oligonucleotide adaptor sequence is necessarily capable of hybridizing to a complementary oligonucleotide, and the requirement for the complementary oligonucleotide to be a sequencing primer does not confer any additional structural requirements on the adaptor. Therefore, claim 14 is not further limiting since any adaptor contains a sequence that can anneal to a complementary sequencing primer or a sequence that is complementary to a sequence that can anneal to a complementary sequencing primer.

Applicant may cancel the claim(s), amend the claim(s) to place the claim(s) in proper dependent form, or present a sufficient showing that the dependent claim(s) complies with the statutory requirements.

Claim Rejections - 35 USC § 102

10. The following is a quotation of the appropriate paragraphs of 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 –
>
> (a)(1) the claimed invention was patented, described in a printed publication, or in public use, on sale or otherwise available to the public before the effective filing date of the claimed invention.

11. Claims 1-5, 8-14, and 16-18 are rejected under 35 U.S.C. 102(a)(1) as being anticipated by Jones et al. (US 2005/0142577 A1; IDS reference).

Regarding claims 1 and 10, Jones teaches a method for enriching a nucleic acid sequence of interest in a sample comprising nucleic acids that includes the following steps (see Fig. 4 and its description in para. 12):

(a) annealing an oligonucleotide to the sequence of interest, wherein the oligonucleotide comprises a 3' portion complementary to the sequence of interest and a 5' portion comprising a first adaptor sequence (Fig. 4, where the oligonucleotide containing a 5' T7 portion corresponds to the claimed oligonucleotide);

(b) extending the annealed oligonucleotide to produce an extension product comprising the first adaptor at the 5' end followed by a sequence complementary to the nucleic acid sequence of interest;

(c) ligating a sequence complementary to a second adaptor sequence to the 3' end of the extension product of step (b) to produce a ligation product comprising the first adaptor at one end, the sequence complementary to the second adaptor at the other end, and the nucleic acid complementary to the sequence of interest therebetween; and

(d) amplifying the ligation product using a primer than anneals to the complement of the first adaptor sequence and a primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest.

Further regarding claim 1 and also regarding claim 2, Jones teaches fragmenting nucleic acids in the sample and annealing the oligonucleotide of step (a) to a fragment containing the nucleic acid sequence of interest (para. 58).

Further regarding claim 1, Jones also teaches using a polymerase for the extension step (para. 77).

Regarding claim 3, Jones teaches practicing the method using genomic DNA (para. 64) and further teaches that the genomic DNA may be from a human chromosome (paras. 54-55). This type of genomic DNA is double-stranded. Jones additionally teaches a fragmentation step before the extension step (para. 58). Therefore, Jones teaches nucleic acid fragments that are double-stranded before the oligonucleotide annealing step.

Regarding claims 4 and 5, Jones teaches that the method may be practiced using genomic DNA (para. 64).

Regarding claim 8, the 3' portion of the oligonucleotide of Jones contains a random sequence (Fig. 4).

Regarding claim 9, the 3' portion of the oligonucleotide of Jones is designed to anneal to the sequence of interest (see Fig. 4; see also paras. 12 and 88-91). Also, as discussed above, it is inherent that oligonucleotides containing the features set forth in the instant claim 1 are designed to anneal to the nucleic acid sequence of interest.

Regarding claim 11, Jones teaches that the adaptors may be double-stranded (para. 49).

Regarding claim 12, Jones teaches that the first and second adaptor sequences are distinct (Fig. 4).

Regarding claim 13, Jones teaches that the first adaptor may include a barcode sequence (para. 58).

Regarding claim 14, the adaptors disclosed in Figure 4 of Jones necessarily include a sequence that is capable of annealing to a sequencing primer or a sequence complementary to a sequence capable of annealing to a sequencing primer. As discussed above, this property is inherent to all adaptors since they necessarily contain sequences capable of annealing to complementary sequences or the complements of said sequences. The requirement for the complementary sequence to be a sequencing primer does not require the adaptors to contain any additional structural features.

Regarding claims 16 and 18, Jones discloses using the method to enrich a plurality of sequences of interest (Fig. 4, where three loci are shown).

Regarding claim 17, Jones teaches uses PCR for the amplification in step (d) (paras. 25, 84, and 100).

Claim Rejections - 35 USC § 103

12. The following is a quotation of 35 U.S.C. 103 which forms the basis for all obviousness rejections set forth in this Office action:

> A patent for a claimed invention may not be obtained, notwithstanding that the claimed invention is not identically disclosed as set forth in section 102, if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains. Patentability shall not be negated by the manner in which the invention was made.

13. This application currently names joint inventors. In considering patentability of the claims the examiner presumes that the subject matter of the various claims was commonly owned as of the effective filing date of the claimed invention(s) absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and effective filing dates of each claim that was not commonly owned as of the effective filing date of the later invention in order for the examiner to consider the applicability of 35 U.S.C. 102(b)(2)(C) for any potential 35 U.S.C. 102(a)(2) prior art against the later invention.

14. Claims 6, 7, 15, 19, 22, and 23 are rejected under 35 U.S.C. 103 as being unpatentable over Jones et al. (US 2005/0142577 A1; IDS reference) in view of Gormley et al. (US 7,741,463 B2; IDS reference).

As discussed above, Jones teaches the method of claims 1-5, 8-14, and 16-18.

Regarding claim 6, Jones imposes no particular limitations on the nucleic acid sequence to be enriched, but the reference fails to disclose enriching a cDNA of interest.

Regarding claim 7, Jones does not specify that double-stranded nucleic acid fragments are denatured before the oligonucleotide annealing step.

Regarding claim 15, Jones teaches practicing the method using DNA fragments (para. 64) as well as amplification by PCR, which uses a DNA polymerase (paras. 25, 84, and 100), but the reference does not specify that the polymerase used in the oligonucleotide extension step is a DNA polymerase.

Regarding claims 19, 22, and 23, Jones teaches obtaining sequence information concerning the nucleic acids enriched by the disclosed method (para. 100), but the reference fails to disclose sequencing the enriched nucleic acid(s) of interest.

Gormley describes a method for preparing a library of nucleic acids for subsequent analysis (abstract, Fig. 2, and col. 3, ll. 44-67), which may include amplification and sequencing (col. 18, ll. 57-65 and col. 20, ll. 30-41). Gormley is similar to Jones in that the reference relates to a method for preparing a collection of nucleic acid fragments of interest for subsequent analysis. The method of Gormley is also similar to Jones in that the library is prepared by a process comprising nucleic acid fragmentation, adaptor ligation, and amplification (Fig. 2).

Regarding claim 6, Gormley teaches that libraries may be generated from genomic DNA or cDNA (col. 6, ll. 5-27).

Regarding claim 7, Gormley describes denaturing double-stranded genomic DNA fragments prior to adaptor prior to primer annealing and extension (col. 26, ll. 20-56). Gormley also describes denaturing double-stranded genomic DNA fragments prior to adaptor annealing (col. 25, ll. 5-40).

Regarding claim 15, Gormley teaches using a DNA polymerase to extend a DNA oligonucleotide that is annealed to a complementary DNA molecule (col. 26, ll. 20-56).

Regarding claims 19, 22, and 23, Gormley teaches that libraries prepared by the disclosed method may be sequenced (e.g., using a method that comprises the use of reversible dye terminators) (col. 20, ll. 30-47 and col. 21, l. 57 - col. 22, l. 20).

Prior to the effective filing date of the claimed invention, it would have been *prima facie* obvious for the ordinary artisan to practice the method of Jones using cDNA as the nucleic acid

of interest. As noted above, Jones does not restrict the nucleic acid of interest in any way, and Gormley teaches that cDNA is a suitable molecule for library construction. In view of these teachings in the prior art, the ordinary artisan would have had motivation and a reasonable expectation of success in using cDNA as the nucleic acid of interest when practicing the method of Jones. See also MPEP 2144.06 and 2144.07, which respectively indicate that it is *prima facie* obvious to substitute art-recognized equivalents and to select a known material based on its suitability for the intended purpose in the absence of unexpected results. In this case, the teachings of Gormley indicate that genomic DNA and cDNA are art-recognized equivalents known to be useful as the starting material for library construction. The teachings of Gormley also indicate that cDNA is a molecule suitable for library preparation. Further, no evidence of unexpected results has been presented with respect to the use of cDNA as the nucleic acid of interest. Thus, the method of claim 6 is *prima facie* obvious.

Prior to the effective filing date of the claimed invention, it also would have been *prima facie* obvious for the ordinary artisan to denature double-stranded genomic DNA fragments prior to the oligonucleotide annealing step when practicing the method of Jones. Gormley provides motivation and a reasonable expectation of success by disclosing denaturation of double-stranded genomic DNA fragments prior to annealing oligonucleotides (col. 25, ll. 5-40 and col. 26, ll. 20-55). The ordinary artisan would have recognized from these teachings of Gormley that denaturation should also be included in the method of Jones before the oligonucleotide annealing step when the nucleic acid fragments are double-stranded, and accordingly, would have been motivated to include this step with a reasonable expectation of success. Thus, the method of claim 7 is *prima facie* obvious.

Prior to the effective filing date of the claimed invention, it also would have been *prima facie* obvious for the ordinary artisan to use a DNA polymerase for the oligonucleotide extension step in the method of Jones. Gormley provides motivation and a reasonable expectation of success by disclosing the use of DNA polymerase for the extension of a DNA oligonucleotide annealed to a complementary DNA sequence (col. 26, ll. 20-55). Based on these teachings of Gormley, the ordinary artisan would have recognized that DNA polymerase would be suitable for the extension step of Jones in which DNA oligonucleotides annealed to a complementary DNA sequence of interest are extended, and accordingly, would have been motivated to select this known material based on its suitability for the intended purpose with a reasonable expectation of success. As discussed in MPEP 2144.07, it is *prima facie* obvious to select a known material based on its suitability for the intended purpose in the absence of unexpected results. In this case, the teachings of Gormley indicate that DNA polymerase is suitable for use in the oligonucleotide extension step of Jones, and no evidence of unexpected results has been presented. Thus, the method of claim 15 is *prima facie* obvious.

Finally, prior to the effective filing date of the claimed invention, it would have been *prima facie* obvious to sequence the amplification products generated by the method of Jones using the method comprising the use of reversible dye terminators described by Gormley. The ordinary artisan would have been motivated to sequence the amplification products of Jones to obtain sequence information in addition to that obtainable using the array-based methods that are the preferred embodiment of Jones and would have had a reasonable expectation of success since Gormley taught that similar libraries of nucleic acids could be sequenced. The ordinary artisan also would have been motivated to use the sequencing method comprising the use of reversible

dye terminators described by Gormley since it allows one to obtain sequence information in real-time and in a single reaction vessel. The ordinary artisan would have had a reasonable expectation of success in using this sequencing method since it is described by Gormley as suitable for sequencing libraries of nucleic acids (col. 21, l. 57 – col. 22, l. 20), and Gormley provides sufficient disclosure concerning its implementation.

Thus, the methods of claims 19, 22, and 23 are *prima facie* obvious.

15. Claims 20 and 21 are rejected under 35 U.S.C. 103 as being unpatentable over Jones et al. (US 2005/0142577 A1; IDS reference) in view of Jang et al. (Journal of Human Genetics 2012; 57: 212-215).

As discussed above, Jones teaches the method of claims 1-5, 8-14, and 16-18.

Regarding claims 20 and 21, Jones does not impose any particular limitations on the nucleic acid sequence of interest, but the reference does not teach that the sequence comprises a gene, and specifically, a cancer gene.

Jang, however, teaches that whole gene sequencing of the *BRCA1* and *BRCA2* genes is desirable for identifying novel mutations in these genes that may be useful for diagnosis of breast cancer in different populations (abstract; p. 212, col. 2; pp. 214-215).

Prior to the effective filing date of the claimed invention, it would have been *prima facie* obvious for the ordinary artisan to practice the method of Jones using at least one of the cancer genes disclosed in Jang as the nucleic acid sequence of interest. Jang provides motivation to do so by disclosing that whole gene sequencing is useful for identifying novel mutations in cancer-associated genes that may be useful for diagnosis of breast cancer in different populations

(abstract; p. 212, col. 2; pp. 214-215). The ordinary artisan would have had a reasonable expectation of success since Jones does not limit the disclosed method to any particular type of nucleic acid sequence of interest. Thus, the methods of claims 20 and 21 are *prima facie* obvious.

Conclusion

16. No claims are currently allowable.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela M. Bertagna whose telephone number is (571)272-8291. The examiner can normally be reached on Monday-Friday, 9-5.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on (571)272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/Angela M. Bertagna/
Primary Examiner, Art Unit 1637

Notice of References Cited	Application/Control No. 14/836,936	Applicant(s)/Patent Under Reexamination AMORESE ET AL.	
	Examiner Angela M. Bertagna	Art Unit 1637	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US 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	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Jang et al. Journal of Human Genetics 2012; 57: 212-215.
	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.

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.
14/990,339	01/07/2016	Doug Amorese	NUGN-762/02US 34079/28	7734

21710 7590 08/29/2018

BROWN RUDNICK LLP
ONE FINANCIAL CENTER
BOSTON, MA 02111

EXAMINER	
STRZELECKA, TERESA E	

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
08/29/2018	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

ip@brownrudnick.com
usactions@brownrudnick.com

Office Action Summary	**Application No.** 14/990,339	**Applicant(s)** Amorese et al.	
	Examiner TERESA E STRZELECKA	**Art Unit** 1637	**AIA 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 3/30/18.

☐ 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) 1-10 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) 1-10 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 _____ 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 CORRESPONDENCE

Notice of Pre-AIA or AIA Status

1. The present application is being examined under the pre-AIA first to invent provisions.

Continued Examination Under 37 CFR 1.114

2. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on March 30, 2018 has been entered.

3. Claims 1-9 were previously pending. Applicant amended claims 1, 8 and 9 and added new claim 10. Claims 1-10 are under consideration.

4. Applicant's amendments overcame the rejection of claims 1-8 under 35 U.S.C. 102(b) as anticipated by Shendure et al. All other previously presented rejections are maintained for reasons given in the "Response to Arguments" below.

5. This office action contains new grounds for rejection.

Response to Arguments

6. Applicant's arguments filed March 30, 2018 have been fully considered but they are not persuasive.

Regarding the rejection of claims 1-8 under 35 U.S.C. 101, Applicant argues the following:

"The amended claims are not directed to a judicial exception.

Each of the currently amended independent claims is directed to a statutory process that includes the tangible steps of creating a nucleic acid library, sequencing the library with a

first primer, denaturing the nucleic acid fragment, and annealing a second primer complementary to a conserved region of the chromosome. All of these steps are physically performed using laboratory instruments and equipment. None of these laboratory steps can reasonably be said to be an abstract idea. Therefore, the claimed process does not fall within a judicial exception.

Even if the amended claims were directed to a judicial exception, which they are not, the claims are still directed to patentable subject matter at least because they recite "significantly more" than a law of nature, natural phenomenon, or abstract idea.

The claimed process adds "significantly more" than any judicial exception because the claimed steps are embodied in physically-performed steps using man-made primer sequences, and because the claims recite specific and meaningful limitations to the steps. For example, two different primers are used, in which the second primer is complementary to a conserved region within a chromosome.

Thus the invention includes the insight that providing a sequencing primer, and a second sequencing primer that is complementary to a conserved gene addresses and solves prior art problems with haplotype phasing. Each of the independent claims not only require first and second sequencing primers, but those claims also require that the second primer hybridized to a conserved region such as, for example, 16S rRNA as described in Example 1 of the original application. Because the sequencing steps use primers that hybridize to evolutionarily conserved sequences, the resultant sequence reads can be assigned to their respective haplotypes and by sequentially using those second primers in turn on the library after the first, the claimed sequential sequencing methods are useful for haplotype phasing. These are human-made processes requiring human-made oligonucleotides of a particular structure (complementary to a conserved region in a chromosome) to control the process.

At least those features of the claimed methods satisfy the requirements of "significantly more" than any judicial exception, and the independent claims recite patent eligible subject matter.

The Office Action on page 4 states that the steps of library construction and sequencing are abstract ideas because these are data gathering steps for compiling data and aligning sequence reads. Applicant disagrees. The step of library construction involves physically processing and transforming a genomic sequence into a format conducive to the sequencing steps. Further, sequencing involves chemical processes to, for example, bind a first primer, denature the nucleic acid fragment, and bind the second primer. None of these physically-performed steps are mere "data collection" and these steps are not abstract ideas. There is no case law or evidence on record supporting that the steps of a method may be parsed and those characterized as relating to data collection may be disregarded as abstract ideas.

The Office Action on page 6 states that claim 1 is directed to relating two nucleic acid sequences, and that independent claims 1, 8, and 9 are all inherently directed to a mental process because relating sequences to each other requires the mental step of comparing.

Applicant disagrees. Amended claim 1—as amended—is not directed to relating two nucleic acid sequences; rather, it is directed to a method for haplotyping that involves at least the physical steps of sequencing with a first primer and a second primer complementary to a conserved region within the chromosome, prior to "mapping first and second sequencing reads to the chromosome." A person of ordinary skill would understand that the step of "mapping" the reads is not a mental process."

Response

However, each of the claims 1, 8 and 9 ends in an abstract step, which, in principle, can be performed using a pen and paper. Step f of claim 1 reads "mapping first and second sequencing reads to the chromosome"; step f of claim 8 reads "compiling data from first and second sequencing reads to map reads originating from a single insert", and the last step of claim 9 reads "aligning the first sequence read and the second sequence read to thereby map the first sequence read and the second sequence read to the template strand". Claim 10 is drawn to the

method of claim 9, further comprising determining whether two or more nucleic acid sequences are located on the same template strand, thereby phasing a haplotype of the chromosome.

Therefore, each of these steps requires mapping reads to a chromosome sequence, i.e., aligning string of letters with each other, which is a mental process. As stated by MPEP 2106.04(a)(2).III.A:

"Another example is *University of Utah Research Foundation v. Ambry Genetics,* 774 F.3d 755, 113 USPQ2d 1241 (Fed. Cir. 2014). The patentee in *Ambry Genetics* claimed methods of screening a human's genome for an altered BRCA gene, comprising comparing the sequence of the human's BRCA gene with the sequence of the wild-type gene, and identifying any differences that arise. 774 F.3d at 763-764, 113 USPQ2d at 1246. The Federal Circuit determined that these claims were directed to the concept of "comparing BRCA sequences and determining the existence of alterations", which was an "abstract mental process". *Id."*

The instant claims are entirely analogous to the above case. Further, the final abstract method steps are recited at a high level of generality.

Finally, the additional steps do not amount to significantly more, as evidenced by the references cited in prior art rejection of the claims.

The rejection is maintained.

Claim Rejections - 35 USC § 101

7. 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.

8. Claims 1-10 are rejected under 35 U.S.C. 101 because the claimed invention is directed to a judicial exception (i.e., a law of nature, a natural phenomenon, or an abstract idea) without significantly more.

Claim 1-8 are directed to relating two nucleic acid sequences (claims 1-7), distinguishing between two closely related nucleic acid sequences (claim 8), or sequencing (claim 9) therefore, they are inherently directed to an abstract ideas, since to "relate" sequences to each other one has to compare them, which involves a mental process.

Step f of claim 1 reads "mapping first and second sequencing reads to the chromosome"; step f of claim 8 reads "compiling data from first and second sequencing reads to map reads originating from a single insert", and the last step of claim 9 reads "aligning the first sequence read and the second sequence read to thereby map the first sequence read and the second sequence read to the template strand". Claim 10, is drawn to the method of claim 9, further comprising determining whether two or more nucleic acid sequences are located on the same template strand, thereby phasing a haplotype of the chromosome.

Therefore, each of these steps requires mapping reads to a chromosome sequence, i.e., aligning string of letters with each other, which is a mental process. As to claim 10, haplotype is just another string of letters. As stated by MPEP 2106.04(a)(2).III.A:

"Another example is *University of Utah Research Foundation v. Ambry Genetics,* 774 F.3d 755, 113 USPQ2d 1241 (Fed. Cir. 2014). The patentee in *Ambry Genetics* claimed methods of screening a human's genome for an altered BRCA gene, comprising comparing the sequence of the human's BRCA gene with the sequence of the wild-type gene, and identifying any differences that arise. 774 F.3d at 763-764, 113 USPQ2d at 1246. The Federal Circuit

determined that these claims were directed to the concept of "comparing BRCA sequences and determining the existence of alterations", which was an "abstract mental process". *Id."*

The instant claims are entirely analogous to the above case. Further, the final method steps are recited at a high level of generality.

The claims do not include additional elements that are sufficient to amount to significantly more than the judicial exception because the method steps of creating nucleic acid libraries and sequencing such libraries were well known and routine in the art at the time of the invention, as evidenced by the cited below reference of Baldwin et al.

In conclusion, the claims are not patent-eligible under 35 U.S.C. 101.

Claim Interpretation

9. Applicant did not define the term "nucleic acid library", therefore any collection of nucleic acids is considered to anticipate this term.

10. Applicant did not define the term "strand-oriented" with respect to a nucleic acid library. Applicant defined the term "strand specific" or "directional" in paragraph [0034] as follows: "[0034] The term "strand specific" or "directional", as used herein, refers to the ability to differentiate in a double-stranded polynucleotide between the original template strand and the strand that is complementary to the original template strand."

11. Applicant did not define the terms "conserved region" and "variable region", therefore any nucleic acid sequence is considered to anticipate these terms.

12. Applicant did not define the term "target capture", therefore any selection method of the target sequence is considered to anticipate this term.

13. Applicant defined the term "haplotype" in paragraph [0023] as follows:

"[0023] As used herein, the term "haplotype" refers to the pattern of alleles within each individual chromosome. Alternatively, haplotype may refer to a set of single-nucleotide polymorphisms (SNPs) that are linked or present together on a single chromosome. The term haplotype may be used to refer to as few as two alleles or SNPs that are linked or present together on a single chromosome."

14. Regarding claims 1, 8 and 9, the claims do not require that the second primer anneals to the nucleic acid library in the same reaction as the first primer, therefore the claims encompass annealing two primers in two different sequencing reactions.

Claim Rejections - 35 USC § 102

15. In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

16. 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.

17. Claim 1-10 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Baldwin et al. (BMC Microbiology, vol. 9:223, pp. 1-9, 2009) as evidenced by BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, pp. 1-72, 2002).

Regarding claims 1, 8 and 9, Baldwin et al. teach a method comprising:

a. creating a strand-oriented nucleic acid library comprising nucleic acid from a chromosome (page 7, paragraphs 7-10, where the library contains amplified chromosomal fragments; since the fragments are amplified using primers designed for the two strands, the library is strand-oriented);

b. sequencing the strand-oriented library with an oligonucleotide primer (page 7, last paragraph; Table 1);

c. denaturing strands of the nucleic acid fragments in the library (page 7, last paragraph; as evidenced by BigDye Terminator v3.1 Cycle Sequencing Kit, the sequencing process uses template strand denaturation in each reaction with a single primer, page 31 (3-5));

d. annealing a new oligonucleotide primer that is complementary to a conserved region or conserved regions within the chromosome (Table I; page 7, last paragraph; as evidenced by BigDye Terminator v3.1 Cycle Sequencing Kit, the sequencing process uses template strand denaturation in each reaction with a single primer, page 31 (3-5));

e. sequencing the nucleic acid library with the new oligonucleotide primer (page 7, last paragraph); and

f. mapping first and second sequencing reads to the chromosome (page 8, first paragraph).

Regarding claims 2 and 3, Baldwin et al. teach amplifying conserved regions which are adjacent to variable regions (page 2, fifth paragraph).

Regarding claims 4 and 5, Baldwin et al. teach differentiation of related transcripts which enables differentiation of related microorganisms (page 2, paragraphs 4-6; page 3; page 4, first and second paragraph).

Regarding claim 6, Baldwin et al. teach libraries of PCR amplicons (page 7, 10th paragraph), therefore they inherently teach reduced complexity libraries.

Regarding claim 7, Baldwin et al. teach libraries of PCR amplicons (page 7, 10th paragraph), therefore they inherently teach capture of specific sequences.

Regarding claims 8-10, Baldwin et al. teach determining whether two or more nucleic acid sequences are located on the same template strand (page 8, second paragraph; page 4, paragraphs 3-5).

18. No claims are allowed.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to TERESA E STRZELECKA whose telephone number is (571)272-0789. The examiner can normally be reached on M-F 9-7.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on 571-272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/TERESA E STRZELECKA/
Primary Examiner, Art Unit 1637
August 23, 2018

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.
14/990,339	01/07/2016	Doug Amorese	25115-762.301	7734

12754 7590 08/30/2017

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER
STRZELECKA, TERESA E

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
08/30/2017	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 14/990,339	Applicant(s) AMORESE ET AL.	
	Examiner TERESA STRZELECKA	Art Unit 1637	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 _____.

☐ 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) 1-8 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) 1-8 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 1/7/16 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 2/2/16;2/9/16;2/29/16;3/23/16;5/2/16;5/11/16;5/23/16;6/14/16;8/22/16;12/1/16.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

1. The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

2. Claims 1-8 are pending and under consideration.

Information Disclosure Statement

3. The information disclosure statements (IDSs) submitted on February 2, 2016; February 9, 2016; February 29, 2016; March 23, 2016; May 2, 2016; May 11, 2016; May 23, 2016; June 14, 2016; August 22, 2016 and December 1, 2016 are in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statements are being considered by the examiner.

Claim Rejections - 35 USC § 112

4. 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.

5. Claims 1-8 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.

A) Claims 1-7 are indefinite in claim 1. Claim 1 is indefinite over the recitation of "(i.e. directional) nucleic acid library", because it is not clear whether "directional" is a limitation of the claim.

B) Claims 1-7 are indefinite in claim 1. Claim 1 recites the limitation "the first strands of the nucleic acid fragments" in line 5. There is insufficient antecedent basis for this limitation in the claim. Neither the preamble nor the preceding steps refer to “a first strand” or “nucleic acid fragments”.

C) Claim 8 recites the limitation "the first strands of the nucleic acid fragments" in line 6. There is insufficient antecedent basis for this limitation in the claim. Neither the preamble nor the preceding steps refer to “a first strand” or “nucleic acid fragments”.

Claim Rejections - 35 USC § 101

6. 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.

7. Claims 1-8 are rejected under 35 U.S.C. 101 because the claimed invention is directed to a judicial exception (i.e., a law of nature, a natural phenomenon, or an abstract idea) without significantly more.

Claim 1-8 are directed to relating two nucleic acid sequences (claims 1-7) or distinguishing between two closely related nucleic acid sequences (claim 8), therefore, they are inherently directed to an abstract idea, since to “relate” sequences to each other one has to compare them, which involves a mental process. The claims do not include additional elements that are sufficient to amount to significantly more than the judicial exception because the method steps of creating nucleic acid libraries and sequencing such libraries were well known and routine in the art at the time of the invention, as evidenced by the cited below reference of Shendure et al.

In conclusion, the claims are not patent-eligible under 35 U.S.C. 101.

Claim Interpretation

8. Applicant defined the term "directional library" in paragraph [0034] as follows: "[0034] The term "strand specific" or "directional", as used herein, refers to the ability to differentiate in a double-stranded polynucleotide between the original template strand and the strand that is complementary to the original template strand."

9. Applicant did not define the term "conserved region", therefore it is interpreted as any sequence.

10. Applicant did not define the term "variable region", therefore it is interpreted as any sequence.

11. Applicant did not define the term "reduced complexity", therefore any library is considered to anticipate the term.

12. Applicant did not define the term "target capture", therefore any method of retrieving the target is considered to anticipate this term.

Claim Rejections - 35 USC § 102

13. 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.

14. Claims 1-8 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Shendure et al. (Science, vol. 309, pp. 1728-1732, and supplementary material, pp. 1-41, 2005).

Claims 1 and 8 are considered together in claim 8, since it is specie of claim 1.

Regarding claims 1 and 8, Shendure et al. teach a method for distinguishing between two closely related nucleic acid sequences, the method comprising (Abstract):

a. creating a strand-oriented nucleic acid library with closely related nucleic acid sequences as inserts (Fig. 1; supplement pages 13-23);

b. sequencing the strand-oriented library with an oligonucleotide primer (Fig. 1; Fig. S7; supplement page 32)

c. denaturing the first strands of the nucleic acid fragments in the library (Fig. S7; supplement page 32);

d. annealing a new oligonucleotide primer that is complementary to a conserved region or conserved regions within the nucleic acid fragments in the nucleic acid library (Fig. S7; supplement page 32);

e. sequencing the nucleic acid library with the new oligonucleotide primer (Fig. S7; supplement page 32); and

f. compiling data from first and second sequencing reads to map reads originating from the same nucleic acid fragments (page 1730, first and last paragraph; page 1731, first and fourth paragraph; Fig. 3; Table 2).

Regarding claims 2 and 3, Shendure et al. teach creating a library from the full E. coli genome, and detecting sequence variations, therefore they inherently teach amplicons originating from conserved regions adjacent to variable regions (supplement pages 13-23; Fig. 3; Table 2).

Regarding claims 4 and 5, Shendure et al. teach aligning multiple variable regions to differentiate related transcripts of related microorganisms (Fig. 3; Table 2).

Regarding claims 6 and 7, Shendure et al. teach reduced complexity libraries obtained by target capture on emulsion PCR beads (Fig. 1; supplement page 24, 25).

15. No claims are allowed.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to TERESA STRZELECKA whose telephone number is (571)272-0789. The examiner can normally be reached on M-F 9-5.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on 571-272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/TERESA STRZELECKA/
Primary Examiner, Art Unit 1637
August 24, 2016

Notice of References Cited	Application/Control No. 14/990,339	Applicant(s)/Patent Under Reexamination AMORESE ET AL.	
	Examiner TERESA STRZELECKA	Art Unit 1637	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US 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	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Shendure, J. et al., Science, vol. 309, pp. 1728-1732, and supplementary material, pp. 1-41 (2005)
	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.

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.
13/750,768	01/25/2013	Doug Amorese	25115-755.201	1162

12754 7590 09/25/2014

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
KAUP, SAHANA S	
ART UNIT	PAPER NUMBER
1639	

NOTIFICATION DATE	DELIVERY MODE
09/25/2014	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 13/750,768	Applicant(s) AMORESE ET AL.	
	Examiner SAHANA KAUP	Art Unit 1639	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 *05 June 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) *1-3,6,9,11,14,40-42,45-49 and 51-56* 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) *1-3, 6, 9, 11, 14, 40-42, 45-49 and 51-56* 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 _____ 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 documents*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

The Applicants' response to the office action filed on 05 June 2014 has been considered and acknowledged.

Priority

This application filed on 25 January 2013 claims priority to US provisional 61/591,241 filed on 26 January 2012.

Status of the Application

Currently, claims 1-3, 6, 9-11, 14, 40-42, 45-49 and 51-54 are pending.

Claims 4-5, 7-8, 12-13, 15-39, 43-44, and 50 have been cancelled.

Claims 1 and 40 have been amended. Claims 55 and 56 have been added.

Claims 1-3, 6, 9-11, 14, 40-42, 45-49 and 51-56 are under examination.

Response to Arguments

With regards to the rejection of claims 1-3, 6, 9, 11, 14 and 51-53 under 35 USC 102(b) as being anticipated by Borodina et al.(Methods Enzymol. 2011 ;500:79-98, published January 1, 2011), Applicants' arguments and the amendment have been fully considered and deemed unpersuasive for the reasons that follow. Instant claim 1 has been amended. However, the scope of instant independent claim 1 has not changed. Therefore, the previous rejection has been maintained.

Instant claims 2, 3, 6, 9, 11, 14 and 51-53 depend from instant claim 1. Therefore, the previous rejections of instant claims 2, 3, 6, 9, 11, 14 and 51-53 are maintained.

With regards to the rejection of claims 40-42, 45-48 under 35 USC 102(b) as being anticipated by Borodina et al.(Methods Enzymol. 2011 ;500:79-98, published January 1, 2011), Applicants' arguments and the amendment have been fully considered and deemed persuasive for the reasons that follow. Applicants have introduced new limitations into instant claim 40 that were not previously examined. Therefore, the scope of the instant set of claims has changed and new rejections are applied. The previous rejection has been withdrawn. The action is made FINAL necessitated by the amendment.

With regards to the rejection of claims 10 and 54 under 35 USC 103(a) as being made obvious over Borodina et al. in view of Vaidyanathan et al.(US2011 01 04785), Applicants' arguments and the amendment have been fully considered and deemed unpersuasive for the reasons that follow. Instant claims 10 and 54 depend from instant claim 1. Instant claim 1 has been amended. However, the scope of instant independent claim 1 has not changed. Therefore, the previous rejection of instant claims 10 and 54 has been maintained.

With regards to the rejection of claim 49 under 35 USC 103(a) as being made obvious over Borodina et al. in view of Vaidyanathan et al.(US2011 01 04785), Applicants' arguments and the amendment have been fully considered and deemed persuasive for the reasons that follow. Instant claim 49 depends from instant claim 40.

Applicants have introduced new limitations into instant claim 40 that were not previously examined. Therefore, the scope of the instant set of claims has changed and new rejections are applied. The previous rejection of instant claim 49 has been withdrawn. The action is made FINAL necessitated by the amendment.

In addition, instant claims 55 and 56 have been added.

However, the cited art reads on instant claim 55. As instant claim 55 depends from instant claim 1 and the previous rejection of instant claim 1 has been maintained, the previous rejection of instant claim 1 has been reiterated to facilitate the rejection of instant claim 55 by the cited art.

As instant claim 56 depends from instant claim 40, it is included in the new rejections discussed below.

Claims 40-42, 45-49, 55 and 56 are pending and under examination.

Claim Rejections

Claim Rejections - 35 USC § 103

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been

> obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

Borodina et al. and Vaidyanathan et al.

Claim 55 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Borodina et al. in view of Vaidyanathan et al.(US20110104785).

As indicated in the previous Office action, Borodina et al. teach : A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a. fragmenting nucleic acids in a sample comprising the nucleic acids, thereby generating nucleic acid fragments;

b. appending a first adaptor sequence to the nucleic acid fragments, thereby generating nucleic acid fragments comprising the appended first adaptor sequence;

c. annealing one or more oligonucleotides to the nucleic acid fragments comprising the appended first adaptor sequence, wherein each of the one or more oligonucleotides comprises a 3' portion that is complementary to a nucleic acid sequence of interest present in one or more of the nucleic acid fragments, and a 5' portion comprising a second adaptor sequence;

d. extending the one or more oligonucleotides annealed to the nucleic acid fragments comprising the appended first adaptor sequence with a polymerase thereby generating one or more oligonucleotide extension products comprising sequence complementary to the appended first adaptor sequence at a first end, the nucleic acid sequence of interest, and the second adaptor sequence at a second end;

and e. amplifying the one or more oligonucleotide extension products using a first primer that is directed against the complement of the first adaptor sequence and a second primer that is directed against the complement of the second adaptor sequence to enrich for the one or more oligonucleotide extension products comprising the nucleic acid sequence of interest (e.g. as indicated in ssRNASEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 1.**

Borodina et al. teach adaptors for nucleic acid amplification and sequencing.

However, Borodina et al. do not teach the 3' portion of each of the one or more oligonucleotides comprises a random sequence.

Vaidyanathan et al. teach: The method of claim 1, wherein the 3' portion of each of the one or more oligonucleotides comprises a random sequence (e.g. barcode sequence as indicated in 0072, pg. 6-7) as set forth by instant **claim 55.**

It would have been prima facie obvious to a person of ordinary skill in the art at the time the invention was made to combine the above citations to develop the instant invention because both teach adaptors for RNA sequencing. In addition, Vaidyanathan et al. teach tagged adaptors which are random sequences(i.e. barcodes). Therefore, it would be obvious for a person of ordinary skill in the art to develop the instant invention of method for enriching for a nucleic acid sequence of interest in a sample, as taught by Borodina et al., in which the adaptors have a random sequence(i.e. barcodes). One of ordinary skill in the art would have been motivated to combine these citations and would have had a reasonable expectation of success in doing so because Vaidyanathan et al. Teach modified adaptors are suitable for RNA sequencing(e.g. as indicated para 0058, pg. 6).

Costa et al. and Kurn et al.

Claim 40-42, 45-49 and 56 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Costa et al.(US20090011959) in view of Kurn et al. (US20110224105).

Costa et al. teach: A method for enriching for a nucleic acid sequence of interest from a library comprising nucleic acid inserts with a first adaptor sequence on a first end and a second adaptor sequence on a second end, the method comprising:
a. denaturing the nucleic acid inserts with a first adaptor sequence on a first end and a second adaptor sequence on a second end, thereby generating a library of single stranded nucleic acid inserts with a first adaptor sequence on a first end and a second

adaptor sequence on a second end (e.g. as indicated in para 0011-0016, pg. 2; para 0034, pg. 3; para 0098, pg. 9) as set forth by instant **claim 40.**

Costa et al. teach: The method of claim 40, wherein the first adaptor and the second adaptor are common to each nucleic acid fragment (e.g. universal as indicated in para 0012, pg. 2) as set forth by instant **claim 45.**

Costa et al. teach: The method of claim 40, wherein the first adaptor and the second adaptor are distinct from each other as set forth by instant **claim 47 and** the method of claim 40, wherein the first adaptor and the second adaptor are the same as set forth by instant **claim 48 (**e.g. Following ligation with universal adaptors A and B (the two different adaptors are sometimes referred to as a "first" and "second" universal adaptor), double-stranded DNA will contain adaptors in four possible combinations: AA, BB, AB, and BA as indicated in para 0026, pg. 2).

Costa et al. teach: The method of claim 40, wherein the first adaptor, the second adaptor and/or the third adaptor sequence further comprise barcode sequence (e.g. unique discriminating key sequence as indicated in para 0036, pg. 3) as set forth by instant **claim 49.**

Costa et al. do not teach a third adaptor sequence.

Kurn et al. teach: b. annealing one or more oligonucleotides to the single stranded nucleic acid inserts with a first adaptor sequence on a first end and a second adaptor sequence on a second end, wherein each of the one or more oligonucleotides comprises a 3' portion that is complementary to a target nucleic acid sequence of

interest present in one or more of the nucleic acid inserts, and a 5' portion comprising a third adaptor sequence(e.g. as indicated in para 0070-0073, pg. 8-9);

c. extending the one or more oligonucleotides with a polymerase thereby generating one or more oligonucleotide extension products comprising sequence complementary to the first adaptor sequence at the first end, the nucleic acid sequence of interest, and the third adaptor sequence at the second end; and d. amplifying the one or more oligonucleotide extension products using a first primer that is directed against the complement of the first adaptor sequence and a second primer that is directed against the complement of the third adaptor sequence to enrich for the one or more oligonucleotide extension products comprising the nucleic acid sequence of interest (e.g. strand tracking as indicated in Fig.7; para 0128-0130, pg. 20) as set forth by instant **claim 40.**

Kurn et al. teach: The method of claim 40, further comprising an additional step of sequencing the amplified one or more oligonucleotide extension products from step d (e.g. as indicated in para 0042,pg. 4-5) as set forth by instant **claim 41.**

Kurn et al. teach: The method of claim 40, wherein the nucleic acid sequence of interest comprises genomic DNA, RNA or cDNA (e.g. as indicated in para 0039,pg. 4) as set forth by instant **claim 42.**

Kurn et al. teach: The method of claim 40, wherein the third adaptor sequence is common to the one or more oligonucleotides (e.g. as indicated in para 0070-0073, pg. 8-9)as set forth by instant **claim 46.**

Kurn et al. teach: The method of claim 40, wherein the 3' portion of each of the one or more oligonucleotides comprises a random sequence (e.g. as indicated in para 0070-0073, pg. 8-9) as set forth by instant **claim 56.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to combine the above citations because both teach generating of nucleic acid libraries. In addition, Kurn et al. teach a third random primer that can facilitate enrichment of first generation libraries. Therefore it would be obvious to develop the method as taught by Costa et al. in which in which the nucleic acid libraries are enriched by amplifying with a third random primer as taught by Kurn et al. One of ordinary skill in the art would have been motivated to combine the above citations and would have had a reasonable expectation of success in doing so because Kurn et al. teach the merits of random primers(e.g. as indicated in para 0043, pg. 5).

Conclusion

No claims are allowable.

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.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SAHANA KAUP whose telephone number is (571)272-6897. The examiner can normally be reached on Monday - Friday 7:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, HEATHER CALAMITA can be reached on (571) 272-2876. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/S. K./

Examiner, Art Unit 1639

/HEATHER CALAMITA/
Supervisory Patent Examiner, Art Unit 1639

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.
14/991,340	01/08/2016	Benjamin G. Schroeder	25115-760.301	4124

12754 7590 09/28/2018

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
KAUP, SAHANA S	

ART UNIT	PAPER NUMBER
1639	

NOTIFICATION DATE	DELIVERY MODE
09/28/2018	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	**Application No.** 14/991,340	**Applicant(s)** Schroeder et al.	
	Examiner SAHANA S KAUP	**Art Unit** 1639	**AIA 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 15 June 2018.

☐ 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) 1-13 and 49-51 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) 1-13 and 49-51 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 _____ 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 1 document.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date _____.

4) ☐ Other: _____.

Notice of Pre-AIA or AIA Status

The present application is being examined under the pre-AIA first to invent provisions.

Claim Rejections - 35 USC § 103

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

McKernan et al. and Dunn

Claims 1-6, 8, 11-13 and 49-51 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. (WO2010003153) in view of Dunn (US20100081174).

McKernan et al. teach methods of generating nucleic acid libraries comprising fragmenting nucleic acid (e.g. para 0057, pg. 15; para 0093, pg. 38), performing end repair on the fragmented target

DNA (e.g. Example 2, pg. 28; para 0093, pg. 38) and contacting the fragmented DNA with double stranded adapters and ligase to mediate ligation of the adapters (e.g. Example 2, pg. 28-29; para 0093-0098, pg. 38-39).

Furthermore, McKernan et al. teach an embodiment in which only the top strands of the adapters are ligated to the target nucleic acid (e.g. nick to prevent ligation event as in para 0061, pg. 17; During adapter ligation, only the top adapter strands were joined to the 5' - phosphorylated ends of the DNA fragments as in para 0098, pg. 39). The adapter ligated DNA is then subjected to extension using DNA polymerase during nick translation (e.g. para 0093-0098, Example 3, pg. 38-39).

McKernan et al. teach adapter ligated nucleic acid is denatured (e.g. para 0076, pg. 24-25) and subjected to bisulfite treatment and PCR amplification (e.g. para 0067, pg. 19; para 0071, pg. 21; para 0076, pg. 24-25; para 0098, pg. 39). McKernan et al. teach an embodiment of this method for analysis of human genomic DNA (e.g. para 0093-para 0101, pg. 38-42).

Therefore, McKernan et al. meet the requirements of the limitations: a method for generating a nucleic acid library, the method comprising: a) fragmenting double-stranded DNA, thereby generating a double-stranded DNA fragment (e.g. para 0057, pg. 15); b) performing end repair on the double-stranded DNA fragment, thereby generating an end repaired double-stranded DNA fragment (e.g. Example 2, pg. 28; para 0093, pg. 38); c) ligating a first strand of a first adapter duplex to a first 5' end of the end repaired double-stranded DNA fragment, wherein a second strand of the first adapter duplex is incapable of ligation to a 3' end of the end-repaired double-stranded DNA fragment, and ligating a first strand of a second adapter duplex to a second 5' end of the double-stranded DNA fragment, wherein the first adapter duplex and the second adapter duplex are the same, thereby generating an adapter-ligated double-stranded DNA fragment (e.g. para 0098, pg. 39); d)extending [[the]]3' ends of the adapter-ligated double-stranded DNA fragment with a DNA polymerase, thereby generating a double-

stranded DNA extension product (e.g. the bottom adapter sequence was filled-in by extension with DNA polymerase during nick-translation as in para 0098, pg. 39) as required by **claim 1.**

McKernan et al. also meet the limitations: e) denaturing the double-stranded DNA extension product, thereby creating a single-stranded DNA fragment comprising the first strand of the first adapter duplex ligated to a first 5' end(e.g. denaturing prior to bisulfite treatment as in para 0076, pg. 24-25) as required by **claim 1.**

McKernan et al. also teach sequencing the genomic DNA library (e.g. Sanger sequencing as in para 00112, p g. 44). Therefore, they meet the limitation: further comprising sequencing the amplified products as recited in **claim 2.**

McKernan et al. meet the limitation: wherein the double-stranded DNA comprises genomic DNA (e.g. gDNA as in para 0093, pg. 38) as recited in **claim 3.**

McKernan et al. meet the limitation: wherein a 5' and/or 3' end of the second strand of the first adapter duplex incapable of ligation is blocked to prevent adapter dimer formation (e.g. preventing ligation as in para 0061, pg. 17; The methyCAP adapter was missing a 5' phosphate from one of its oligonucleotides, which resulted in a nick on each strand as in para 00135, pg. 55) as recited in **claim 4.**

McKernan et al. meet the limitation: wherein a 3' end of the second strand of the first adapter duplex incapable of ligation is blocked with a terminal dideoxycytosine (e.g. Sanger sequencing as in para 00112, p g. 44) as recited in **claim 5.**

McKernan et al. meet the limitation: wherein a 5' end of the second strand of the first adapter duplex incapable of ligation contains a biotin moiety (e.g. internal adapter as in para 00137, pg. 55; dsMethylA adapter (SEQ ID. NO. 15), pg. 67) as recited in **claim 6.**

McKernan et al. meet the limitation: wherein the dCTP analog resistant to bisulfite treatment is 5-methyl dCTP (e.g. as in para 0052, pg. 14; para 0093, pg. 38) as recited in **claim 8.**

McKernan et al. meet the limitation: wherein step d) is performed in presence of dATP, dGTP, dTTP and a dCTP analog resistant to bisulfite treatment (e.g. as in para 0052, pg. 14; para 0093, pg. 38)as recited in **claim 49**.

McKernan et al. meet the limitation: wherein the first adapter duplex is a partial duplex (e.g. dsMethyIA adapter duplex as in para 00137, pg. 55; dsMethyIA adapter (SEQ ID. NO. 15), pg. 67) as recited in **claim 50**.

McKernan et al. meet the limitation: wherein a long strand of the partial duplex is the first strand of the first adapter duplex and a shorter strand of the partial duplex is the second strand of the first adapter duplex (e.g. as in para 0082, p g. 26; Fig 2) as recited in **claim 51**.

McKernan et al. teach a method of nucleic acid library preparation comprising subjecting adapter ligated nucleic acids to bisulfite treatment (e.g. para 0067, pg. 19; para 0071, pg. 21; para 0076, pg. 24-25; para 0098, pg. 39) and performing PCR with oligonucleotide primers corresponding to the priming sites on each strand of each adapter, thereby generating amplified products (e.g. library amplification as in para 00102-00107, pg. 42-43).

However, McKernan et al. teach adapters wherein the first strand comprises methylation resistant nucleotides and the second strand lacks methylation resistant nucleotides. They teach target nucleic acids that are ligated with these adapters are subjected to bisulfite treatment, resulting in converted target nucleic acids comprising adapters wherein the first strand is not converted by bisulfite and the second strand is converted by bisulfite. These converted target nucleic acids are amplified with primers that anneal to the converted second strands of the adapters ligated to the target nucleic acids (e.g. para 0013, pg. 4).

Therefore, McKernan et al. do not teach amplification with primers that anneal to a converted first strand of attached adapters. Therefore, they do not meet the requirements of the limitations: f) subjecting the single-stranded DNA fragment to bisulfite treatment, thereby converting cytosine

residues to uracils in the first strand of the first adapter duplex ligated to the first 5' end as required by **claim 1.**

Furthermore, they do not meet the limitation: g) performing polymerase chain reaction (PCR), with an oligonucleotide primer that anneals to a complement of the first strand of the first adaptor duplex ligated to the 5' end comprising uracils thereby generating a nucleic acid library comprising an amplified product as required by **claim 1.**

Like McKernan et al., Dunn teaches a method of treating adapter-ligated nucleic acids with bisulfite. Dunn also teaches unmethylated cytosine residues in the adapters (i.e. amplification cassettes) are converted by bisulfite treatment and the resulting converted fragments are amplified by PCR amplification using primers that are designed to the adapters which have cytosine residues converted to uracil (e.g. para 0063-0068, pg. 6).

Therefore, Dunn meets the limitation: f) subjecting the single-stranded DNA fragment to bisulfite treatment, thereby converting cytosine residues to uracils in the first strand of the first adapter duplex ligated to the first 5' end and g) performing polymerase chain reaction (PCR), with an oligonucleotide primer that anneals to a complement of the first strand of the first adaptor duplex ligated to the 5' end comprising uracils thereby generating a nucleic acid library comprising an amplified product as required by **claim 1.**

Furthermore, Dunn teach an embodiment wherein target DNA is fragmented and enriched for densely- methylated DNA by capture with methyl- binding ligands. The capture can occur before or after ligation of adapters (i.e. amplification cassettes) to target nucleic acids. Furthermore, Dunn teaches bisulfite treatment of target nucleic acids after adapter ligation and after enrichment by methyl capture (e.g. either prior to or after ligation of an amplification cassette to both ends of the double stranded fragments, the DNA is contacted with a Me-CpG binding ligand and the fragments which are rich in Me-CpG are separated from fragments having few Me-CpG sequences... After optional ligation of a double

stranded amplification cassette to the double stranded fragments and/or after separation of Me-CpG rich fragments, the DNA is treated under conditions suitable for conversion of all cytosine residues to uracil as in para 0033,pg. 3; para 0034-0041,pg. 3-4).

Dunn teach the methyl binding ligand is anti- methyl antibodies or methyl binding proteins such as MeCP2 and McrA (e.g. para 0044-0045, pg. 4).

Dunn teaches that McrA binds to 5-methylcytosine and 5-hydroxymethylcytosine (e.g. para 0101, pg. 8). Furthermore, they teach methylcytosine capture using recombinant McrA (e.g. para 0033-0041, pg. 3-4; para 0044-0045, pg. 4; DNA containing cMeCpGG sequences may be first separated from the bulk fragments by contact with a CMeCpGG binding ligand such rMcrA as in para 0073, pg. 6).

Therefore, Dunn meets the limitation: further comprising performing wherein 5-methylcytosine capture is performed prior to step f) (e.g. After optional ligation of a double stranded amplification cassette to the double stranded fragments and/or after separation of Me-CpG rich fragments, the DNA is treated under conditions suitable for conversion of all cytosine residues to uracil as in para 0033, pg. 3) as recited in **claim 11.**

As Dunn teaches methyl capture of fragmented DNA with a methyl binding ligand, McrA, that binds to 5-methylcytosine and 5-hydroxymethylcytosine (e.g. para 0033-0041, pg. 3-4; para 0044-0045, pg. 4; para 0073, pg. 6; para 0101, pg. 8), they also meet the limitation: wherein the dCTP cytosine analog resistant to bisulfite treatment is a cytosine analog other than 5-methyl dCTP 5 methylcytosine as recited in **claim 11.**

Furthermore, Dunn meets the limitation: wherein the 5-methylcytosine capture is performed using a methylcytosine binding protein (e.g. para 0033-0041, pg. 3-4; para 0044-0045, pg. 4) as recited in **claim 12.**

Furthermore, Dunn meets the limitation: wherein the 5-methylcytosine capture is performed using an anti-5-methylcytosine antibody (e.g. para 0044, pg. 4) as recited in **claim 13.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the method of McKernan to include providing adapters wherein both strands of the adapter are converted by bisulfite and amplified with primers that complement converted nucleic acid strands as taught by Dunn because this was a particular known technique recognized as part of the ordinary capabilities of one skilled in the art that was recognized for achieving the predictable result of a method of generating a nucleic acid library.

Furthermore, it would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the method of McKernan to include enriching target DNA by methyl capture after fragmentation and prior to bisulfite treatment as taught by Dunn because Dunn teaches enrichment by methyl capture facilitates subsequent processing of methylated target DNA (e.g. Ligation of an amplification cassette to the Me-CpG enriched fragment portion may facilitate ensuring that a molar excess of cassette to fragment ends had been used as in para 0033, pg. 3; Me-CpG DNA fragments are then further processed to enrich for me-CpG sequences of interest as in para 0038, pg. 4).

McKernan et al., Dunn and Margulies et al.

Claim 7 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. and Dunn as applied to claims 1-6, 8, 11-13 and 49-51 above, and further in view of Margulies et al.(Nature 437.7057 (2005): 376-380.).

The teachings of McKernan et al. and Dunn as applied above are incorporated in this rejection.

McKernan et al. and Dunn teach a method of nucleic acid library preparation comprising performing PCR with oligonucleotide primers corresponding to bisulfite- converted strands of the adapters, yielding amplified products. Furthermore, McKernan et al. teach clonal amplification by a method taught by Margulies et al. (e.g. clonal amplification including method of Margulies et al. as para 0044, pg. 11-12; para 0072, pg. 21-22).

However, McKernan et al. and Dunn do not expressly teach primers comprising barcodes as recited in **claim 7.**

Margulies et al. teach clonal amplification comprises using barcoded primers (e.g. annealing primers as in Emulsion-based sample preparation section. Pg. 376-377; Methods section, pg. 380; "key" sequence as indicated in 1st para, Base calling of individual reads section, pg. 378; 3rd , pg. 3 of 34, Supplementary Methods).

Therefore, Margulies et al. meet the limitation: wherein step g) comprises annealing to the single-stranded DNA fragment a sequence-specific oligonucleotide primer comprising a barcode sequence as recited in **claim 7.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan et al. and Dunn to feature annealing primers which have barcode sequences (i.e. key sequences) as taught by Margulies et al. as a skilled artisan would have recognized that the simple substitution the primers of McKernan and Dunn used for library amplification with the barcoded primers of Margulies would yield the predictable outcome of a method for generating bisulfite converted directional nucleic acid libraries with added benefit of easier identification of library members based on incorporated barcodes.

McKernan et al., Dunn and Huang et al.

Claim 9 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. and Dunn as applied to claims 1-6, 8, 11-13 and 49-51 above, and further in view of Huang et al. (PloS one 5.1 (2010): e8888; 9 pages).

The teachings of McKernan et al. and Dunn as applied above are incorporated in this rejection.

McKernan et al. and Dunn teach a method of nucleic acid library preparation comprising performing PCR with oligonucleotide primers corresponding to bisulfite- converted strands of the adapters, yielding amplified products. Furthermore, McKernan et al. teach bisulfite treatment of

genomic DNA containing 5-methylcytosine in their method of generating nucleic acid libraries. However, McKernan et al. and Dunn do not teach **claim 9.**

Huang et al. teach that 5-hydroxylmethyl cytosine, like 5-methylcytosine, is resistant to conversion by bisulfite (e.g. 2nd para, pg. 5 of 9 pages).

Therefore, Huang et al. meet the limitation: wherein the dCTP cytosine analog resistant to bisulfite treatment is 5-hydroxymethyl dCTP as recited in **claim 9.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan and Dunn to include 5-hydroxymethyl dCTP among bisulfite resistant methylcytosine residues in the target genomic DNA as taught by Huang et al. as a skilled artisan would have recognized that analysis of a population of methylated DNA comprising methylcytosine and hydroxymethylcytosine-containing DNA which would allow more comprehensive analysis of methylated DNA.

McKernan et al. and Granados et al.

Claim 10 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. and Dunn as applied to claims 1-6, 8, 11-13 and 49-51 above, and further in view of Granados et al. (US20110091882).

The teachings of McKernan et al. and Dunn as applied above are incorporated in this rejection.

McKernan et al. and Dunn teach a method of nucleic acid library preparation comprising performing PCR with oligonucleotide primers corresponding to bisulfite- converted strands of the adapters, yielding amplified products. Furthermore, McKernan et al. teach bisulfite treatment of genomic DNA containing 5-methylcytosine in their method of generating nucleic acid libraries. However, McKernan et al. and Dunn do not teach **claim 10.**

Granados et al. teach incorporation of propynyl analogs such as propynyl cytosine into the base composition of nucleic acid that is subjected to bisulfite treatment (e.g. para 0013, pg. 2; para 0022, pg.

3). Additionally, resistance to bisulfite conversion is considered an inherent feature of 5 propynyl cytosine residues.

Therefore, as Granados teaches bisulfite treatment of nucleic acid comprising 5 propynyl cytosine residues, their teaching meets the limitation: wherein the dCTP cytosine analog resistant to bisulfite treatment is 5-propynyl dCTP as recited in **claim 10.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan and Dunn to include 5-propynyl dCTP among bisulfite resistant methylcytosine residues in the target genomic DNA as taught by Granados et al. as a skilled artisan would have recognized that analysis of a population of methylated DNA comprising methylcytosine and propynylcytosine-containing DNA which would allow more comprehensive analysis of methylated DNA.

Response to the Arguments

Any rejection not reiterated or specifically addressed has been overcome by amendment. New rejections are set forth to address the amended claims.

However, previously cited references teach art relevant to the amended claims and therefore are included in the new rejections.

Applicants argue that the teaching of McKernan et al. does not anticipate amended claim 1 because "Example 1 and Example 2 of McKernan et al. do not disclose step c) of amended independent claim 1 ("ligating a first strand of a first adapter duplex to a first 5' end of the end repaired double-stranded DNA fragment"). This argument is persuasive.

It is noted that Example 1 of McKernan is not cited in the current rejections. Furthermore, Example 2 of McKernan is cited to show that end repair followed by ligase mediated reactions was known in the art at the time the claimed invention was made.

It is further noted that Example 3 of McKernan discloses an embodiment in which only the top strands of the adapters are ligated to the target nucleic acid (e.g. nick to prevent ligation event as in para 0061, pg. 17; During adapter ligation, only the top adapter strands were joined to the 5' -phosphorylated ends of the DNA fragments as in para 0098, pg. 39). The adapter ligated DNA is then subjected to extension using DNA polymerase during nick translation (e.g. para 0093-0098, Example 3, pg. 38-39). Therefore, this embodiment is cited to address the limitation: ligating a first strand of a first adapter duplex to a first 5' end of the end repaired double-stranded DNA fragment as recited in amended claim 1.

The teaching of Dunn is also applied to meet the requirements of the current amended claims.

Conclusion

No claims are allowable.

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.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SAHANA S KAUP whose telephone number is (571)272-6897. The examiner can normally be reached on M-F 7-10 EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, HEATHER CALAMITA can be reached on 571-272-2876. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/S.S.K/
Examiner, Art Unit 1639

/JULIET C SWITZER/
Primary Examiner, Art Unit 1634

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.
13/750,768	01/25/2013	Doug Amorese	25115-755.201	1162

12754 7590 10/30/2015

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
BERTAGNA, ANGELA MARIE	

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
10/30/2015	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 13/750,768	Applicant(s) AMORESE ET AL.	
	Examiner Angela M. Bertagna	Art Unit 1637	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 *10/15/15*.

☐ 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) *1-3,6,9,10,14,40-42,46-49,51,52,54-58 and 60-93* 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) *1-3,6,9,10,14,40-42,46-49,51,52,54-58 and 60-93* is/are rejected.

8)☒ Claim(s) *1,40 and 93* 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 _____ 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

Notice of Pre-AIA or AIA Status

1. The present application is being examined under the pre-AIA first to invent provisions.

Status of the Application

2. Applicant's response filed on October 15, 2015 has been entered. Claims 1-3, 6, 9, 10, 14, 40-42, 46-49, 51, 52, 54-58, and 60-93 are pending and under examination.

Response to Arguments

3. Applicant's arguments filed on October 15, 2015 have been fully considered.

Objection to claim 55

Applicant argues that the objection should be withdrawn in view of the amendment to claim 55 (page 11). This argument was persuasive. The objection has been withdrawn.

Duplicate Claims

As noted by Applicant, see page 11, claim 45 has been canceled. The issue of whether claims 45 and 48 are duplicates is therefore moot.

Rejection of claims 53 and 59 under pre-AIA 35 U.S.C. 112, fourth paragraph

Applicant argues that the rejection should be withdrawn since claims 53 and 59 have been canceled (page 11). This argument was persuasive. The rejection has been withdrawn.

Rejection of claims 1-3, 6, 9, 10, 14, 40-42, 45-49, 51-54, and 57-62 under pre-AIA 35 U.S.C. 102(b) as being anticipated by Jones et al. (US 2003/0232348 A1, "Jones 1")

Applicant argues that the rejection should be withdrawn in view of the amendments to independent claims 1 and 40, which require the first adaptor sequence or its complement, the second adaptor sequence or its complement, the first primer or its complement, or the second primer or its complement to comprise sequence to initiate sequencing on a massively parallel sequencing platform (pages 12 and 13).

This argument was persuasive. The rejection has been withdrawn.

Rejection of claims 1-3, 6, 9, 10, 14, 40-42, 45-49, and 51-62 under pre-AIA 35 U.S.C. 102(b) as being anticipated by Jones et al. (US 2005/0142577 A1, "Jones 2")

Applicant argues that the rejection should be withdrawn in view of the amendments to independent claims 1 and 40, which require the first adaptor sequence or its complement, the third adaptor sequence or its complement, the first primer or its complement, or the second primer or its complement to comprise sequence to initiate sequencing on a massively parallel sequencing platform (pages 12 and 13).

This argument was persuasive. The rejection has been withdrawn.

Rejection of claims 55 and 56 under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Jones 1 in view of Jones 2

Applicant argues that the rejection should be withdrawn because neither reference teaches all of the elements of amended independent claims 1 and 40 (page 14).

This argument was persuasive. The rejection has been withdrawn.

Claim Objections

4. Claims 1 and 40 are objected to because the word “the” should be inserted before

"second primer" in the third-to-last line of each claim.

Claim 93 is objected to because "comprise" should be replaced with "comprises".

Claim Rejections - 35 USC § 112 (New Matter)

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.

Claims 1-3, 6, 9, 10, 14, 40-42, 46-49, 51, 52, 54-58, and 60-93 are rejected under 35 U.S.C. 112(a) or 35 U.S.C. 112 (pre-AIA), first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor or a joint inventor, or for pre-AIA the inventor(s), at the time the application was filed, had possession of the claimed invention. **This is a new matter rejection.**

Independent claim 1 has been amended to state that "the first adaptor sequence or its complement, the second adaptor sequence or its complement, the first primer or its complement, or the second primer or its complement comprise sequence to initiate sequencing on a massively parallel sequencing platform". Independent claim 40 has been amended to state that "the first adaptor sequence or its complement, the third adaptor sequence or its complement, the first primer or its complement, or the second primer or its complement comprise sequence to initiate

sequencing on a massively parallel sequencing platform". Claims 68-77 and 83-92 further limit this clause in claims 1 and 40.

Applicant states that the amendments to claims 1 and 40 find support throughout the original disclosure and specifically points to Figure 1, Figure 4, and paragraphs 32, 37, 41, 52, 54, 60, 64, and 71 (page 10). Applicant states that claims 68-77 and 83-92 are supported throughout the original disclosure and specifically points to Figure 4 and paragraphs 32, 41, 60, 74, and 95 (page 10).

The original disclosure, including the portions cited by Applicant, has been reviewed, but support was not found for all of the options recited in independent claims 1 and 40. Paragraph 32 provides support a method in which the first adaptor or the second adaptor comprises sequence to initiate sequencing on a massively parallel sequencing platform. Paragraph 95 provides support for a method in which the first and second primers comprise the sequence to initiate sequencing on the massively parallel sequencing platform. Figure 4 and paragraphs 32 and 37 provide support for a method in which the third adaptor comprises sequence for initiating sequencing on a massively parallel sequencing platform. However, neither these passages nor the remainder of the original disclosure provides support for methods in which the complements of the first, second, or third adaptor or the complements of the first or second primer comprise the sequence to initiate sequencing on the massively parallel sequencing platform. As a result, claims 1 and 40, and those claims depending therefrom, contain new matter.

The following is suggested to overcome the rejection and also to facilitate prosecution in general: Replacing the concluding clause in claims 1 and 40 with a positively recited massively parallel sequencing step and canceling claims 68-77 and 83-92.

Claim Rejections - 35 USC § 112

6. 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.

Claims 1-3, 6-9, 10, 14, 40-42, 46-49, 51, 52, 54-58, 60-67, 70-82, and 85-93 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.

Claims 1 and 40 are indefinite because the requirements of "comprise sequence to initiate sequencing on a massively parallel sequencing platform", which is recited in the concluding clause of each claim, are unclear. More specifically, it is entirely unclear as to what is meant by "to initiate sequencing". The word "initiate" suggests that an active step occurs at the sequence, such as polymerase-mediated extension, ligation, or exonuclease digestion, but this does not make sense for all of the options in the claims because the adaptor sequences and their complements are covalently attached to the nucleic acid of interest and cannot be extended in a polymerase-mediated extension reaction. The language also suggests that particular structural features are required, but neither the claims nor the specification discusses what these features might be. Paragraphs 70-74 of the specification describe sequencing methods, but they do not use the language "comprises a sequence to initiate sequencing" and, therefore, do not clarify the language, particularly since paragraph 73 seems to indicate that no special structural features are required to initiate sequencing. In view of the above, the requirements of claims 1 and 40 are entirely unclear.

Claims 2, 3, 6, 9, 10, 14, 41, 42, 46-49, 51, 52, 54-58, 60-67, 70-82, and 85-93 are also indefinite since they depend from claim 1 or claim 40.

Claims 70 and 85 are also vague and indefinite because the requirements of the claims are unclear. The claims recite that "the sequence to initiate sequencing on the massively parallel sequencing platform is a flow cell sequence". It is not clear whether this language requires a sequence that is immobilized in a flow cell or if it only requires a sequence capable of being used in a flow cell.

Claims 75-77 and 90-92 are further indefinite because it is unclear what is meant by the "allows annealing" language in each of claims 75 and 90. It is not clear whether the language requires annealing of the enriched nucleic acid of interest to a solid support via the sequence to initiate sequencing or if another, broader meaning is intended (e.g., that the sequence to initiate sequencing does not prevent annealing).

Claim Rejections - 35 USC § 112

7. The following is a quotation of 35 U.S.C. 112(d):

> (d) REFERENCE IN DEPENDENT FORMS.—Subject to subsection (e), a claim in dependent form shall contain a reference to a claim previously set forth and then specify a further limitation of the subject matter claimed. A claim in dependent form shall be construed to incorporate by reference all the limitations of the claim to which it refers.

The following is a quotation of 35 U.S.C. 112 (pre-AIA), fourth paragraph:

> Subject to the [fifth paragraph of 35 U.S.C. 112 (pre-AIA)], a claim in dependent form shall contain a reference to a claim previously set forth and then specify a further limitation of the subject matter claimed. A claim in dependent form shall be construed to incorporate by reference all the limitations of the claim to which it refers.

Claims 70 and 85 are rejected under 35 U.S.C. 112(d) or pre-AIA 35 U.S.C. 112, 4th paragraph, as being of improper dependent form for failing to further limit the subject matter of

the claim upon which it depends, or for failing to include all the limitations of the claim upon which it depends.

Claims 70 and 85 depend from claims 1 and 40, respectively, and recite that "the sequence to initiate sequencing on the massively parallel sequencing platform is a flow cell sequence". As discussed above, these claims are indefinite because it is not clear whether any structural difference relative to claims 1 and 40 is required. If no structural difference is required (i.e., if the language is intended to simply require the sequence to be capable of being used in a flow cell), the claims are not further limiting because the sequences in claims 1 and 40 are necessarily capable of being used in a flow cell. In other words, if no structural difference is required, claims 70 and 85 are not further limiting because they merely recite an inherent property of the sequences in claims 1 and 40.

Applicant may cancel the claim(s), amend the claim(s) to place the claim(s) in proper dependent form, rewrite the claim(s) in independent form, or present a sufficient showing that the dependent claim(s) complies with the statutory requirements.

Claim Rejections - 35 USC § 103

8. The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the

various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

9. Claims 1-3, 6, 9, 10, 14, 40-42, 46-49, 51, 52, 54-58, 60-66, 68-81, and 83-93 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Jones et al. (US 2005/0142577 A1) in view of Tucker et al. (The American Journal of Human Genetics (2009) 85: 142-154).

Regarding claims 1, 2, and 14, Jones teaches a method for enriching a nucleic acid sequence of interest in a sample comprising nucleic acids that comprises the following steps (see, for example, Figures 3-4 and paragraphs 86-90, 96, 99, and 100): (a) fragmenting nucleic acids in the sample to generate fragments comprising the sequence of interest, (b) ligating a first adaptor to a 5' end of the nucleic acid fragments, (c) annealing one or more oligonucleotides in solution to the nucleic acid sequence of interest in the fragments, wherein said oligonucleotides comprise a 3' portion with at least ten bases that are complementary to the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence, (d) extending said annealed oligonucleotides using a DNA polymerase in a reaction mixture that does not contain a primer that anneals to sequence complementary to the first adaptor sequence, thereby obtaining extension products comprising sequence complementary to the ligated first adaptor, sequence complementary to the nucleic acid of interest, and the second adaptor, (e) amplifying the extension products using a primer that anneals to the complement of the first adaptor sequence

and a primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest, and (f) sequencing the enriched nucleic acid of interest.

Regarding claim 3, Jones teaches that the nucleic acid sequence of interest comprises genomic DNA (paragraph 100). The reference also teaches that the nucleic acid sequence of interest may comprise RNA (paragraphs 43-44).

Regarding claims 6 and 52, Jones teaches that the nucleic acid fragments are double-stranded and that method comprises denaturing the nucleic acid fragments before the step of annealing the oligonucleotides (step (c) above), thereby generating single-stranded nucleic acid fragments comprising the first adaptor sequence and the nucleic acid sequence of interest (Figure 3 and paragraphs 87 and 100).

Regarding claim 9, Jones teaches that the first and second adaptor sequences are distinct from each other (Figure 3, paragraph 87, and paragraph 100).

Regarding claims 10 and 54, Jones teaches that the second adaptor sequence may comprise a barcode sequence (paragraphs 58 and 61) and also that the first adaptor may include a restriction site (paragraph 49).

Regarding claim 51, in the method of Jones, the oligonucleotides comprise a 3' portion that is complementary to a different nucleic acid sequence of interest in the nucleic acid fragments and a 5' portion comprising the second adapter sequence (paragraphs 87, 99, and 100; see also Figure 3).

Regarding claim 55, Jones teaches that the 3' portion of the oligonucleotides comprises a random sequence (see Figure 4, paragraph 100 and paragraphs 88-90).

Regarding claims 57, 58, 60, and 61, Jones teaches that the first adaptor sequence is in a long strand of a partial duplex adaptor comprising a long strand and a short strand, the long strand is ligated to a 5' end of the nucleic acid fragments, and the short strand is ligated to a 3' end of the nucleic acid fragments (see Figure 3 and paragraph 87).

Regarding claim 62, Jones teaches that the first adaptor sequence is in a ligation strand of a first adaptor, wherein the first adaptor comprises the ligation strand and a non-ligation strand, wherein the ligation strand ligates to a 5' end of the nucleic acid fragments (paragraph 50).

Regarding claim 40, Jones teaches a method for enriching a nucleic acid sequence of interest from a library (see Figure 3, paragraphs 86-90, 99, and 100), wherein the library comprises nucleic acid inserts comprising the nucleic acid sequence of interest (Figure 3 and paragraphs 86 and 100), wherein the nucleic acid inserts have a first adaptor sequence attached on a first end and a second adaptor sequence attached on a second end (Figure 3 and paragraphs 86 and 100), the method comprising:

(a) denaturing the nucleic acid inserts, thereby generating a library of single-stranded nucleic acids with the first and second adaptor sequences flanking the nucleic acid sequence of interest (Figure 3 and paragraphs 86 and 100);

(b) annealing one or more oligonucleotides in solution to the nucleic acid sequence of interest in the single-stranded nucleic acids, wherein the oligonucleotides comprise a 3' portion with at least ten bases that are complementary to the nucleic acid sequence of interest and a 5' portion comprising a third adaptor sequence that is distinct from the first and second adaptor sequences (Figure 3, paragraph 86, and paragraph 100, where the oligonucleotides containing an A1 or degenerate sequence are annealed);

(c) extending said annealed oligonucleotides using a polymerase in a reaction mixture that does not include a primer that anneals to the complement of the first or second adaptor sequence to obtain extension products comprising sequence complementary to the first adaptor, sequence complementary to the nucleic acid of interest, and the third adaptor (Figure 3, paragraph 86, and paragraph 100); and

(d) amplifying the extension products using a primer that anneals to the complement of the first adaptor sequence and a primer that anneals to the complement of the third adaptor sequence to enrich for the nucleic acid sequence of interest (Figure 3 and paragraphs 86 and 100).

Regarding claim 41, Jones teaches that the method further comprises sequencing the enriched nucleic acid sequence of interest (paragraph 100).

Regarding claim 42, Jones teaches that the nucleic acid sequence of interest comprises genomic DNA (paragraph 100). The reference also teaches that the nucleic acid sequence of interest may comprise RNA (paragraphs 43-44).

Regarding claim 46, Jones teaches that the third adaptor sequence is the same among the one or more oligonucleotides (Figure 3 and paragraphs 86 and 100).

Regarding claim 47, Jones teaches that the first and second adaptor sequence may be distinct from each other (Figure 3 and also paragraph 53).

Regarding claim 48, Jones teaches that the first and second adaptor sequences are the same among each of the inserts in the library (Figure 3 and paragraphs 86 and 100).

Regarding claim 49, Jones teaches that the third adaptor sequence may comprise a barcode sequence (paragraph 58).

Regarding claim 56, Jones teaches that the 3' portion of the oligonucleotides comprises a random sequence (paragraph 100 and paragraphs 88-90).

Regarding claim 93, Jones teaches that the neither the first adaptor sequence nor the second adaptor sequence comprises a 3' block (see, for example, paragraph 87, which indicates that the 3' blocking group is optional).

In the method of Jones, the amplification products contain adaptor sequences at each end (see Figure 3, for example; see also paragraphs 87 and 100). It is not clear that these adaptor sequences, their complements, the amplification primers described in paragraph 100, or their complements comprise sequence to initiate sequencing on a massively parallel sequencing platform as required by independent claims 1 and 40. Also, Jones teaches that the nucleic acids amplified by the disclosed method may be analyzed by array hybridization or any other method known in the art to obtain sequence information about a region of interest (paragraph 79-80). The reference does not disclose massively parallel sequencing, however.

Tucker describes methods of obtaining sequence information about a region of interest in a nucleic acid, and specifically describes a number of massively parallel sequencing methods (abstract, pages 142-146, and Figures 2-5). The reference teaches that "Massively parallel sequencing will allow simultaneous screening for mutations in hundreds of loci in genetically heterogeneous disorders, whole-genome screening for novel mutations, and sequence-based detection of novel pathogens that cause human disease. In addition, massively parallel sequencing will permit clinical application of our expanding knowledge of pharmacogenetics, cancer genetics, epigenetics, and complex traits" (page 142, column 2). The reference further teaches that massively parallel sequencing "has reduced the cost and increased the throughput of

genomic sequencing by more than three orders of magnitude...and may also replace arrays and Sanger sequencing in clinical applications where they are currently being used" (abstract).

Further regarding claims 63 and 78, Tucker describes massively parallel methods developed by Illumina, Applied Biosystems, Roche, and Helicos (pages 142-144 and Figs. 2-5).

Regarding claims 64, 65, 79, and 80, the Illumina method comprises bridge amplification and the use of four labeled reversible terminators (Figure 2).

Regarding claims 68-72, 75, 77, 83-87, 90, and 92, the Illumina method comprises preparing a nucleic acid fragment of interest with an adaptor at each end (Figure 2). The adaptors hybridize to flow cell sequences (Figure 2). Each of the two adaptors is also complementary to a sequencing primer (Figure 2). In addition, the complement of each of the two adaptors contains a sequencing primer sequence as well as a flow cell sequence (Figure 2).

Further regarding claims 66 and 81, Tucker also describes nanopore sequencing (page 143, column 1).

Further regarding claims 76 and 91, the 454 method uses beads as the support for the sequencing method (Figure 4). As in the Illumina method, the adaptors contain a sequence that allows annealing to the solid support (Figure 4).

It would have been *prima facie* obvious for one of ordinary skill in the art at the time of the invention to analyze the amplification products generated by the method of Jones using one of the massively parallel sequencing methods described in Tucker. The ordinary artisan would have been motivated to do so since Tucker taught that such methods offered very high-throughput sequencing (abstract and page 142) and as well as the ability to accurately detect minor alleles (page 144, column 2). Additional motivation is provided in the abstract of Tucker,

which teaches that massively parallel sequencing methods may replace array methods. The ordinary artisan would have had a reasonable expectation of success in view of the guidance provided by Tucker concerning the use of massively parallel sequencing methods and also since Jones taught that the amplification products could be analyzed using any desired method. In using the Illumina sequencing method, for example, the ordinary artisan would have recognized that the adaptor sequences or their complements should include sequence to initiate sequencing on a parallel sequencing platform and would have been motivated to use such adaptors with a reasonable expectation of success since the methods of Illumina and 454 have been publicly disclosed and commercialized. Further regarding claims 73, 74, 88, and 89, since the disclosure of Tucker indicates that the adaptor sequences or their complements contain sequence to initiate sequencing on the massively parallel sequencing platform (Figure 2), the ordinary artisan also would have been motivated to perform the amplification step in the method of Jones using first and second primers that either contain the sequence to initiate sequencing on the massively parallel sequencing platform or contain a sequence complementary thereto. Thus, the methods of claims 1-3, 6, 9, 10, 14, 40-42, 46-49, 51, 52, 54-58, 60-66, 68-81, and 83-93 are *prima facie* obvious.

10. Claims 67 and 82 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Jones et al. (US 2005/0142577 A1) in view of Tucker et al. (The American Journal of Human Genetics (2009) 85: 142-154) and further in view of Rothberg et al. (US 2009/0127589 A1).

As discussed above, the teachings of Jones in view of Tucker render obvious the methods

of claims 1-3, 6, 9, 10, 14, 40-42, 46-49, 51, 52, 54-58, 60-66, 68-81, and 83-93.

Jones does not teach semiconductor sequencing as required by claims 67 and 82. Tucker does not remedy this deficiency.

However, Rothberg describes a massively parallel semiconductor sequencing method (abstract, paragraphs 28-29, 33, and 51-58; see also paragraphs 315-343).

It would have been *prima facie* obvious for one of ordinary skill in the art at the time of the invention to substitute semiconductor sequencing as taught by Rothberg for one of the massively parallel sequencing methods disclosed in Tucker. The ordinary artisan would have recognized from the teachings of Rothberg that the disclosed semiconductor sequencing method was a massively parallel sequencing method and, accordingly, would have considered it an art-recognized equivalent of the methods disclosed in Tucker. As discussed in MPEP 2144.06, it is *prima facie* obvious to substitute art-recognized equivalents in the absence of unexpected results. Attention is also directed to MPEP 2144.07, which states that it is *prima facie* obvious to select a known method based on its suitability for the intended purpose. In this case, the teachings of Rothberg indicate that the disclosed method is suitable for massively parallel sequencing. Accordingly, the ordinary artisan would have been motivated to select this method with a reasonable expectation of success. It is further noted that no evidence of unexpected results with respect to the sequencing method have been presented. Thus, the methods of claims 67 and 82 are *prima facie* obvious.

Conclusion

11. No claims are currently allowable.

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.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela M. Bertagna whose telephone number is (571)272-8291. The examiner can normally be reached on Monday-Friday, 9-5.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on (571)272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/Angela M. Bertagna/
Primary Examiner, Art Unit 1637

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.
13/750,768	01/25/2013	Doug Amorese	25115-755.201	1162

12754 7590 09/25/2014

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
KAUP, SAHANA S	
ART UNIT	PAPER NUMBER
1639	

NOTIFICATION DATE	DELIVERY MODE
09/25/2014	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 13/750,768	Applicant(s) AMORESE ET AL.	
	Examiner SAHANA KAUP	Art Unit 1639	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 *05 June 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) *1-3,6,9,11,14,40-42,45-49 and 51-56* 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) *1-3, 6, 9, 11, 14, 40-42, 45-49 and 51-56* 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 _____ 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 documents*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

The Applicants' response to the office action filed on 05 June 2014 has been considered and acknowledged.

Priority

This application filed on 25 January 2013 claims priority to US provisional 61/591,241 filed on 26 January 2012.

Status of the Application

Currently, claims 1-3, 6, 9-11, 14, 40-42, 45-49 and 51-54 are pending.

Claims 4-5, 7-8, 12-13, 15-39, 43-44, and 50 have been cancelled.

Claims 1 and 40 have been amended. Claims 55 and 56 have been added.

Claims 1-3, 6, 9-11, 14, 40-42, 45-49 and 51-56 are under examination.

Response to Arguments

With regards to the rejection of claims 1-3, 6, 9, 11, 14 and 51-53 under 35 USC 102(b) as being anticipated by Borodina et al.(Methods Enzymol. 2011 ;500:79-98, published January 1, 2011), Applicants' arguments and the amendment have been fully considered and deemed unpersuasive for the reasons that follow. Instant claim 1 has been amended. However, the scope of instant independent claim 1 has not changed. Therefore, the previous rejection has been maintained.

Instant claims 2, 3, 6, 9, 11, 14 and 51-53 depend from instant claim 1. Therefore, the previous rejections of instant claims 2, 3, 6, 9, 11, 14 and 51-53 are maintained.

With regards to the rejection of claims 40-42, 45-48 under 35 USC 102(b) as being anticipated by Borodina et al.(Methods Enzymol. 2011 ;500:79-98, published January 1, 2011), Applicants' arguments and the amendment have been fully considered and deemed persuasive for the reasons that follow. Applicants have introduced new limitations into instant claim 40 that were not previously examined. Therefore, the scope of the instant set of claims has changed and new rejections are applied. The previous rejection has been withdrawn. The action is made FINAL necessitated by the amendment.

With regards to the rejection of claims 10 and 54 under 35 USC 103(a) as being made obvious over Borodina et al. in view of Vaidyanathan et al.(US2011 01 04785), Applicants' arguments and the amendment have been fully considered and deemed unpersuasive for the reasons that follow. Instant claims 10 and 54 depend from instant claim 1. Instant claim 1 has been amended. However, the scope of instant independent claim 1 has not changed. Therefore, the previous rejection of instant claims 10 and 54 has been maintained.

With regards to the rejection of claim 49 under 35 USC 103(a) as being made obvious over Borodina et al. in view of Vaidyanathan et al.(US2011 01 04785), Applicants' arguments and the amendment have been fully considered and deemed persuasive for the reasons that follow. Instant claim 49 depends from instant claim 40.

Applicants have introduced new limitations into instant claim 40 that were not previously examined. Therefore, the scope of the instant set of claims has changed and new rejections are applied. The previous rejection of instant claim 49 has been withdrawn. The action is made FINAL necessitated by the amendment.

In addition, instant claims 55 and 56 have been added.

However, the cited art reads on instant claim 55. As instant claim 55 depends from instant claim 1 and the previous rejection of instant claim 1 has been maintained, the previous rejection of instant claim 1 has been reiterated to facilitate the rejection of instant claim 55 by the cited art.

As instant claim 56 depends from instant claim 40, it is included in the new rejections discussed below.

Claims 40-42, 45-49, 55 and 56 are pending and under examination.

Claim Rejections

Claim Rejections - 35 USC § 103

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been

> obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

Borodina et al. and Vaidyanathan et al.

Claim 55 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Borodina et al. in view of Vaidyanathan et al.(US20110104785).

As indicated in the previous Office action, Borodina et al. teach : A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a. fragmenting nucleic acids in a sample comprising the nucleic acids, thereby generating nucleic acid fragments;

b. appending a first adaptor sequence to the nucleic acid fragments, thereby generating nucleic acid fragments comprising the appended first adaptor sequence;

c. annealing one or more oligonucleotides to the nucleic acid fragments comprising the appended first adaptor sequence, wherein each of the one or more oligonucleotides comprises a 3' portion that is complementary to a nucleic acid sequence of interest present in one or more of the nucleic acid fragments, and a 5' portion comprising a second adaptor sequence;

d. extending the one or more oligonucleotides annealed to the nucleic acid fragments comprising the appended first adaptor sequence with a polymerase thereby generating one or more oligonucleotide extension products comprising sequence complementary to the appended first adaptor sequence at a first end, the nucleic acid sequence of interest, and the second adaptor sequence at a second end;

and e. amplifying the one or more oligonucleotide extension products using a first primer that is directed against the complement of the first adaptor sequence and a second primer that is directed against the complement of the second adaptor sequence to enrich for the one or more oligonucleotide extension products comprising the nucleic acid sequence of interest (e.g. as indicated in ssRNASEQ protocol section, pg. 83-86 and Figure 5.1, pg. 84) as set forth by instant **claim 1.**

Borodina et al. teach adaptors for nucleic acid amplification and sequencing.

However, Borodina et al. do not teach the 3' portion of each of the one or more oligonucleotides comprises a random sequence.

Vaidyanathan et al. teach: The method of claim 1, wherein the 3' portion of each of the one or more oligonucleotides comprises a random sequence (e.g. barcode sequence as indicated in 0072, pg. 6-7) as set forth by instant **claim 55.**

It would have been prima facie obvious to a person of ordinary skill in the art at the time the invention was made to combine the above citations to develop the instant invention because both teach adaptors for RNA sequencing. In addition, Vaidyanathan et al. teach tagged adaptors which are random sequences(i.e. barcodes). Therefore, it would be obvious for a person of ordinary skill in the art to develop the instant invention of method for enriching for a nucleic acid sequence of interest in a sample, as taught by Borodina et al., in which the adaptors have a random sequence(i.e. barcodes). One of ordinary skill in the art would have been motivated to combine these citations and would have had a reasonable expectation of success in doing so because Vaidyanathan et al. Teach modified adaptors are suitable for RNA sequencing(e.g. as indicated para 0058, pg. 6).

Costa et al. and Kurn et al.

Claim 40-42, 45-49 and 56 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Costa et al.(US20090011959) in view of Kurn et al. (US20110224105).

Costa et al. teach: A method for enriching for a nucleic acid sequence of interest from a library comprising nucleic acid inserts with a first adaptor sequence on a first end and a second adaptor sequence on a second end, the method comprising:
a. denaturing the nucleic acid inserts with a first adaptor sequence on a first end and a second adaptor sequence on a second end, thereby generating a library of single stranded nucleic acid inserts with a first adaptor sequence on a first end and a second

adaptor sequence on a second end (e.g. as indicated in para 0011-0016, pg. 2; para 0034, pg. 3; para 0098, pg. 9) as set forth by instant **claim 40.**

Costa et al. teach: The method of claim 40, wherein the first adaptor and the second adaptor are common to each nucleic acid fragment (e.g. universal as indicated in para 0012, pg. 2) as set forth by instant **claim 45.**

Costa et al. teach: The method of claim 40, wherein the first adaptor and the second adaptor are distinct from each other as set forth by instant **claim 47 and** the method of claim 40, wherein the first adaptor and the second adaptor are the same as set forth by instant **claim 48 (**e.g. Following ligation with universal adaptors A and B (the two different adaptors are sometimes referred to as a "first" and "second" universal adaptor), double-stranded DNA will contain adaptors in four possible combinations: AA, BB, AB, and BA as indicated in para 0026, pg. 2).

Costa et al. teach: The method of claim 40, wherein the first adaptor, the second adaptor and/or the third adaptor sequence further comprise barcode sequence (e.g. unique discriminating key sequence as indicated in para 0036, pg. 3) as set forth by instant **claim 49.**

Costa et al. do not teach a third adaptor sequence.

Kurn et al. teach: b. annealing one or more oligonucleotides to the single stranded nucleic acid inserts with a first adaptor sequence on a first end and a second adaptor sequence on a second end, wherein each of the one or more oligonucleotides comprises a 3' portion that is complementary to a target nucleic acid sequence of

interest present in one or more of the nucleic acid inserts, and a 5' portion comprising a third adaptor sequence(e.g. as indicated in para 0070-0073, pg. 8-9);

c. extending the one or more oligonucleotides with a polymerase thereby generating one or more oligonucleotide extension products comprising sequence complementary to the first adaptor sequence at the first end, the nucleic acid sequence of interest, and the third adaptor sequence at the second end; and d. amplifying the one or more oligonucleotide extension products using a first primer that is directed against the complement of the first adaptor sequence and a second primer that is directed against the complement of the third adaptor sequence to enrich for the one or more oligonucleotide extension products comprising the nucleic acid sequence of interest (e.g. strand tracking as indicated in Fig.7; para 0128-0130, pg. 20) as set forth by instant **claim 40.**

Kurn et al. teach: The method of claim 40, further comprising an additional step of sequencing the amplified one or more oligonucleotide extension products from step d (e.g. as indicated in para 0042,pg. 4-5) as set forth by instant **claim 41.**

Kurn et al. teach: The method of claim 40, wherein the nucleic acid sequence of interest comprises genomic DNA, RNA or cDNA (e.g. as indicated in para 0039,pg. 4) as set forth by instant **claim 42.**

Kurn et al. teach: The method of claim 40, wherein the third adaptor sequence is common to the one or more oligonucleotides (e.g. as indicated in para 0070-0073, pg. 8-9)as set forth by instant **claim 46.**

Kurn et al. teach: The method of claim 40, wherein the 3' portion of each of the one or more oligonucleotides comprises a random sequence (e.g. as indicated in para 0070-0073, pg. 8-9) as set forth by instant **claim 56.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to combine the above citations because both teach generating of nucleic acid libraries. In addition, Kurn et al. teach a third random primer that can facilitate enrichment of first generation libraries. Therefore it would be obvious to develop the method as taught by Costa et al. in which in which the nucleic acid libraries are enriched by amplifying with a third random primer as taught by Kurn et al. One of ordinary skill in the art would have been motivated to combine the above citations and would have had a reasonable expectation of success in doing so because Kurn et al. teach the merits of random primers(e.g. as indicated in para 0043, pg. 5).

Conclusion

No claims are allowable.

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.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SAHANA KAUP whose telephone number is (571)272-6897. The examiner can normally be reached on Monday - Friday 7:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, HEATHER CALAMITA can be reached on (571) 272-2876. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/S. K./

Examiner, Art Unit 1639

/HEATHER CALAMITA/
Supervisory Patent Examiner, Art Unit 1639

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.
14/991,340	01/08/2016	Benjamin G. Schroeder	25115-760.301	4124

12754 7590 12/15/2017

Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
KAUP, SAHANA S	

ART UNIT	PAPER NUMBER
1639	

NOTIFICATION DATE	DELIVERY MODE
12/15/2017	ELECTRONIC

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.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@wsgr.com

Office Action Summary	Application No. 14/991,340	Applicant(s) Schroeder et al.	
	Examiner SAHANA S KAUP	Art Unit 1639	AIA 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 03 June 2016 .

☐ 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) 1-13 and 49-52 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) 1-13 and 49-52 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 _____ 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 10 documents.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date _____.

4) ☐ Other: _____.

The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

Claim Rejections

Claim Rejections - 35 USC § 102

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

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.

Claims 1-6, 8 and 49-52 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by McKernan et al. (WO2010003153).

McKernan et al. teach methods of generating directional nucleic acid libraries comprising fragmenting nucleic acid(e.g. para 0057, pg. 15);contacting the fragmented DNA with adapters and ligating the adapters to the fragmented DNA(e.g. para 0059, pg. 16). McKernan also teaches an embodiment in which the ligation events do not occur (e.g. nick to prevent ligation event as in para 0061, pg. 17). The adapter ligated DNA is then subjected to extension using DNA polymerase during nick translation(e.g. para 0093-0098, Example 3, pg. 38-39). The nucleic acid is treated with bisulfite and subjected to PCR amplification (e.g. para 0067,pg. 19; para 0071, pg. 21). McKernan et al. teach an embodiment of this method for analysis of human genomic DNA(e.g. para 0093-para 0101, pg. 38-42).

Therefore, McKernan et al. teach the following limitations: McKernan et al. teach: fragmenting double-stranded DNA, thereby generating double-stranded DNA fragments(e.g. para 0093, pg. 38) ; b. performing end repair on the double-stranded DNA fragments(e.g. End-It DNA end-repair kit (Epicentre Biotechnologies, Madison, WI, USA) was used as in para 0093, pg. 38); c. ligating a single oligonucleotide adapter duplex, wherein one strand of the single oligonucleotide adapter duplex is capable of ligation to a 5' end of the double-stranded DNA fragments and an opposing strand of the single oligonucleotide adapter duplex is incapable of ligation at a 3' end of the double-stranded DNA fragments, to both ends of each of the double-stranded DNA fragments(e.g. nick to prevent ligation event as in para 0061, pg. 17; During adapter ligation, only the top adapter strands were joined to the 5' - phosphorylated ends of the DNA fragments as in para 0098, pg. 39) ; d. extending the 3' end of the double-stranded DNA fragments with a DNA polymerase(e.g. bottom adapter sequence was filled-in by extension with DNA polymerase as in para 0098, pg. 39); e. denaturing the double stranded DNA fragments, thereby creating single-stranded DNA fragments(e.g. as in para 0099, pg. 39-40); f. subjecting the single-stranded DNA fragments to bisulfite treatment, thereby converting cytosine residues to uracils and creating PCR priming sites at [[the]] 5' and 3' ends of the single-stranded DNA fragments(e.g. as in para 00100-00101, pg. 40-42); and g. performing PCR with oligonucleotide primers corresponding to the priming sites, thereby generating amplified products (e.g. library amplification as in para 00102-00107, pg. 42-43) as recited in **claim 1.**

McKernan et al. also teach sequencing genomic DNA library(e.g. Sanger sequencing as in para 00112, p g. 44). Therefore, they teach the limitation: further comprising an additional step of sequencing the amplified products as recited in **claim 2.**

McKernan et al. teach the limitation: wherein the double-stranded DNA comprises genomic DNA (e.g. gDNA as in para 0093, pg. 38) as recited in **claim 3.**

McKernan et al. teach the limitation: wherein a 5' and/or 3' end of the opposing strand of the single oligonucleotide adapter duplex incapable of ligation [[are]] is blocked and enzymatically unreactive to prevent adapter dimer formation(e.g. preventing ligation as in para 0061, pg. 17; The methyCAP adapter was missing a 5' phosphate from one of its oligonucleotides, which resulted in a nick on each strand as in para 00135, pg. 55) as recited in **claim 4.**

McKernan et al. teach the limitation: wherein a 3' end of the opposing strand of the single oligonucleotide adapter duplex incapable of ligation is blocked with a terminal dideoxycytosine (e.g. Sanger sequencing as in para 00112, p g. 44) as recited in **claim 5.**

McKernan et al. teach the limitation: wherein a 5' end of the opposing strand of the single oligonucleotide adapter duplex incapable of ligation contains a biotin moiety (e.g. internal adapter as in para 00137, pg. 55; dsMethylA adapter (SEQ ID. NO. 15), pg. 67) as recited in **claim 6.**

McKernan et al. teach the limitation: wherein the dCTP analog resistant to bisulfite treatment is 5-methyl dCTP (e.g. as in para 0052, pg. 14; para 0093, pg. 38) as recited in **claim 8.**

McKernan et al. teach the limitation: wherein the extension reaction is performed in presence of dATP, dGTP, dTTP and a dCTP analog resistant to bisulfite treatment (e.g. as in para 0052, pg. 14;para 0093, pg. 38)as recited in **claim 49**.

McKernan et al. teach the limitation: wherein the single oligonucleotide adapter duplex is a partial duplex (e.g. dsMethylA adapter duplex as in para 00137, pg. 55; dsMethylA adapter (SEQ ID. NO. 15), pg. 67) as recited in **claim 50**.

McKernan et al. teach the limitation: wherein a long strand of the partial duplex is the one strand of the single oligonucleotide adapter duplex capable of ligation to the 5' end of the double-stranded DNA fragments and a shorter strand of the partial duplex is the opposing strand of the single oligonucleotide adapter duplex incapable of ligation at the 3' end of the double-stranded DNA fragments (e.g. as in para 0082, p g. 26; Fig 2) as recited in **claim 51**.

McKernan et al. teach the limitation: a. fragmenting double-stranded DNA, thereby generating double-stranded DNA fragments(e.g. para 0093, pg. 38); b. performing end repair on the double stranded DNA fragments(e.g. para 0093, pg. 38); c. ligating distinct oligonucleotide adapters of similar construct to both ends of each DNA fragment, wherein at least one cytosine residue in a ligation compatible arm of the adapters has been replaced with a cytosine analog resistant to bisulfite treatment(e.g. para 0093-0097, pg. 38); d. extending a 3' end of the double-stranded DNA fragments with a DNA polymerase(e.g. para 0098, pg. 39); e. denaturing the double stranded DNA fragments, thereby creating single-stranded DNA fragments(e.g. para 0099, pg. 39-40); f. subjecting the single-stranded DNA fragments to bisulfite treatment, thereby converting cytosine residues to uracils and creating PCR priming sites at 5' and 3' ends of the single-stranded DNA fragments (e.g. as in para 00100-00101, pg. 40-42)and g. performing PCR with oligonucleotide primers corresponding to the priming sites, thereby generating amplified products (e.g. library amplification as in para 00102-00107, pg. 42-43) as recited in **claim 52**.

Claim Rejections - 35 USC § 103

In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under pre-AIA 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.

2. Ascertaining the differences between the prior art and the claims at issue.

3. Resolving the level of ordinary skill in the pertinent art.

4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

This application currently names joint inventors. In considering patentability of the claims under pre-AIA 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of pre-AIA 35 U.S.C. 103(c) and potential pre-AIA 35 U.S.C. 102(e), (f) or (g) prior art under pre-AIA 35 U.S.C. 103(a).

McKernan et al. and Margulies et al.

Claim 7 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. in view of Margulies et al.(Nature 437.7057 (2005): 376-380.).

The teachings of McKernan et al. as applied in the previous rejection above are incorporated in this rejection.

McKernan et al. teach clonal amplification by a method taught by Margulies et al. (e.g. clonal amplification including method of Margulies et al. as para 0044, pg. 11-12; para 0072, pg. 21-22). Furthermore, McKernan et al. teach performing PCR with oligonucleotide primers corresponding to the

priming sites, thereby generating amplified products (e.g. library amplification as in para 00102-00107, pg. 42-43). However, they do not expressly teach primers comprising barcodes as recited in **claim 7.**

Margulies et al. teach: wherein step g) comprises annealing to the single-stranded DNA fragments a sequence-specific oligonucleotide primer, or multiple sequence-specific oligonucleotide primers, wherein the sequence-specific oligonucleotide primer or multiple sequence-specific oligonucleotide primers comprise a barcode sequence (e.g. annealing primers as in Emulsion-based sample preparation section. Pg. 376-377; Methods section, pg. 380; "key" sequence as indicated in 1st para, Base calling of individual reads section, pg. 378; 3rd , pg. 3 of 34, Supplementary Methods) as recited in **claim 7.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan et al. to feature annealing primers which have barcode sequences (i.e. key sequences) as taught by Margulies et al. as a skilled artisan would have recognized that the simple substitution the primers of McKernan used for library amplification with the barcoded primers of Margulies would yield the predictable outcome of a method for generating bisulfite converted directional nucleic acid libraries with added benefit of easier identification of library members based on incorporated barcodes.

McKernan et al. and Huang et al.

Claim 9 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. in view of Huang et al. (PloS one 5.1 (2010): e8888; 9 pages).

The teachings of McKernan et al. as applied in the previous rejection above are incorporated in this rejection.

McKernan et al. teach bisulfite treatment of genomic DNA containing 5-methylcytosine in their method of generating nucleic acid libraries. However, they do not teach **claim 9.**

Huang et al. teach that 5-hydroxylmethyl cytosine, like 5-methylcytosine, is resistant to conversion by bisulfite (e.g. 2nd para, pg. 5 of 9 pages).

Therefore, Huang et al. teach the limitation: wherein the dCTP cytosine analog resistant to bisulfite treatment is 5-hydroxymethyl dCTP as recited in **claim 9.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan to include 5-hydroxymethyl dCTP among bisulfite resistant methylcytosine residues in the target genomic DNA as taught by Huang et al. as a skilled artisan would have recognized that analysis of a population of methylated DNA comprising methylcytosine and hydroxymethylcytosine-containing DNA which would allow more comprehensive analysis of methylated DNA .

McKernan et al. and Granados et al.

Claim 10 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. in view of Granados et al. (US20110091882).

The teachings of McKernan et al. as applied in the previous rejection above are incorporated in this rejection.

McKernan et al. teach bisulfite treatment of genomic DNA containing 5-methylcytosine in their method of generating nucleic acid libraries. However, they do not teach **claim 10.**

Granados et al. teach incorporation of propynyl analogs such as propynyl cytosine into the base composition of nucleic acid that is subjected to bisulfite treatment (e.g. para 0013, pg. 2; para 0022, pg. 3). Additionally, resistance to bisulfite conversion is considered an inherent feature of 5 propynyl cytosine residues.

Therefore, as Granados teaches bisulfite treatment of nucleic acid comprising 5 propynyl cytosine residues, their teaching meets the limitation: wherein the dCTP cytosine analog resistant to bisulfite treatment is 5-propynyl dCTP as recited in **claim 10.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan to include 5-propynyl dCTP among bisulfite resistant methylcytosine residues in the target genomic DNA as taught by Granados et al. as a skilled artisan would have recognized that analysis of a population of methylated DNA comprising methylcytosine and propynylcytosine-containing DNA which would allow more comprehensive analysis of methylated DNA .

McKernan et al. , Ehrich et al. and Mulligan et al.

Claims 11 and 12 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al. in view of Ehrich et al.(US20100105049) and Mulligan et al.(Nucleic acids research 38.6 (2009): 1997-2005.).

The teachings of McKernan et al. as applied in the previous rejection above are incorporated in this rejection.

McKernan et al. teach bisulfite treatment of genomic DNA in their method of generating nucleic acid libraries. However, they do not teach **claims 11 and 12.**

At the time the invention was made, Ehrich et al. teach detection of methylation by treating genomic DNA with a methyl cytosine binding protein (i.e. MBD-Fc protein) to enrich for methylated DNA prior to bisulfite treatment(e.g. para 0179-0190, Example 1, pg. 22-24).

Therefore, as Ehrich et al. teach methylated DNA enrichment by binding of a methyl cytosine binding protein prior to bisulfite treatment(e.g. para 0179-0190, Example 1, pg. 22-24),they teach the limitation: wherein 5-methylcytosine capture is performed prior to step f) as recited in **claim 11.**

Furthermore, as Ehrich teaches methylated DNA capture using a methyl cytosine binding protein (i.e. MBD-Fc protein), they teach the limitation: wherein 5-methylcytosine capture is performed using a methylcytosine binding protein as recited in **claim 12.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan to include methylcytosine capture as taught by Ehrich as a skilled artisan would have appreciate that methylcytosine capture results in enrichment for methylated DNA which would allow for more efficient bisulfite treatment of target DNA.

McKernan and Ehrich teach treatment of methylated DNA. However, they do not expressly teach the limitation: wherein the dCTP cytosine analog resistant to bisulfite treatment is a cytosine analog other than 5-methyl dCTP 5 methylcytosine.

At the time the invention was made, Mulligan et al. teach another methyl cytosine binding protein(i.e. McrA protein) that binds to 5-methylcytosine and 5-hydroxymethylcytosine(e.g. 2nd para, Introduction, pg. 1997). Furthermore, they teach methylcytosine capture using recombinant McrA(e.g. rMcrA-S affinity enrichment of DNA fragments containing m5CpG section, pg. 1999).

Therefore, the teaching of Mulligan meets the limitation: wherein the dCTP cytosine analog resistant to bisulfite treatment is a cytosine analog other than 5-methyl dCTP 5 methylcytosine as recited in **claim 11.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan and Ehrich to include methylcytosine capture using the methylcytosine binding protein McrA as taught by Mulligan et al. as a skilled artisan would have recognized that , as McrA recognizes methylcytosine and hydroxymethylcytosine-containing DNA, that the simple substitution of McrA for another would result in enrichment for larger population of methylated DNA comprising methylcytosine and hydroxymethylcytosine-containing DNA which would allow more comprehensive analysis of methylated DNA .

McKernan et al., Ehrich et al., Mulligan et al. and Weber et al.

Claim 13 is rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over McKernan et al., Ehrich et al. and Mulligan et al. as applied to claim 11 and 12 above, and further in view of Weber et al.(Nature genetics 37.8 (2005): 853-862.).

The teachings of McKernan et al., Ehrich et al. and Mulligan et al. as applied in the previous rejection above are incorporated in this rejection.

McKernan et al., Ehrich et al. and Mulligan et al. teach a method of generating nucleic acid libraries comprising enriching methylated DNA by 5-methylcytosine capture. However, they do not teach **claim 13.**

At the time the invention was made, Weber et al. teach a capture with methyl cytosine binding antibody is known in the art(e.g. Unbiased detection of methylated DNA by immunoprecipitation section, pg. 853-854; Fig. 1, pg. 854; MeDIP assay, pg. 860).

Therefore, Weber et al. teach the limitation: wherein 5-methylcytosine capture is performed using an anti-5-methylcytosine antibody as recited in **claim 13.**

It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of McKernan, Ehrich and Mulligan to include methylcytosine capture using the methylcytosine binding antibody as taught by Weber et al. as a skilled artisan would have recognized that the simple substitution of one methyl binding agent for another would result in enrichment of methylated DNA.

Conclusion

No claims are allowable.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SAHANA S KAUP whose telephone number is (571)272-6897. The examiner can normally be reached on M-F 7-10 EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, HEATHER CALAMITA can be reached on 571-272-2876. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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.

/S.S.K/
Examiner, Art Unit 1639

/JULIET C SWITZER/
Primary Examiner, Art Unit 1634

Notice of References Cited	Application/Control No. 14/991,340	Applicant(s)/Patent Under Reexamination Schroeder et al.	
	Examiner SAHANA S KAUP	Art Unit 1639	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-20110091882-A1	04-2011	Granados; Edward N.	C12Q1/6858	435/6.13
*	B	US-20100105049-A1	04-2010	Ehrich; Mathias	C12Q1/6804	435/6.12
	C					
	D					
	E					
	F					
	G					
	H					
	I					
	J					
	K					
	L					
	M					

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	CPC Classification
	N	WO-2010003153-A2	01-2010	WO	BOYD VICTORIA L	C12Q1/6827
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Margulies et al.(Nature 437.7057 (2005): 376-380.). (Year: 2005)
	V	Huang et al. (PloS one 5.1 (2010): e8888; 9 pages). (Year: 2010)
	W	Mulligan et al.(Nucleic acids research 38.6 (2009): 1997-2005.). (Year: 2009)
	X	Weber et al.(Nature genetics 37.8 (2005): 853-862.). (Year: 2005)

*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.

METHODS

Human Mutation

OFFICIAL JOURNAL HGVS HUMAN GENOME VARIATION SOCIETY www.hgvs.org

Deep Sequencing to Reveal New Variants in Pooled DNA Samples

Astrid A. Out,[1] Ivonne J.H.M. van Minderhout,[2] Jelle J. Goeman,[3] Yavuz Ariyurek,[2,4] Stephan Ossowski,[6] Korbinian Schneeberger,[6] Detlef Weigel,[6] Michiel van Galen,[2,4] Peter E.M. Taschner,[2,4] Carli M.J. Tops,[1] Martijn H. Breuning,[1] Gert-Jan B. van Ommen,[2,4] Johan T. den Dunnen,[1,2,4] Peter Devilee,[2,5] and Frederik J. Hes[1]

[1]Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands; [2]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands; [3]Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands; [4]Leiden Genome Technology Center, Leiden University Medical Center, Leiden, The Netherlands; [5]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands; [6]Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany

Communicated by Graham R. Taylor

Received 20 April 2009; accepted revised manuscript 18 August 2009.

Published online 4 September 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/humu.21122

ABSTRACT: We evaluated massive parallel sequencing and long-range PCR (LRP) for rare variant detection and allele frequency estimation in pooled DNA samples. Exons 2 to 16 of the MUTYH gene were analyzed in breast cancer patients with Illumina's (Solexa) technology. From a pool of 287 genomic DNA samples we generated a single LRP product, while the same LRP was performed on 88 individual samples and the resulting products then pooled. Concentrations of constituent samples were measured with fluorimetry for genomic DNA and high-resolution melting curve analysis (HR-MCA) for LRP products. Illumina sequencing results were compared to Sanger sequencing data of individual samples. Correlation between allele frequencies detected by both methods was poor in the first pool, presumably because the genomic samples amplified unequally in the LRP, due to DNA quality variability. In contrast, allele frequencies correlated well in the second pool, in which all expected alleles at a frequency of 1% and higher were reliably detected, plus the majority of singletons (0.6% allele frequency). We describe custom bioinformatics and statistics to optimize detection of rare variants and to estimate required sequencing depth. Our results provide directions for designing high-throughput analyses of candidate genes.

Hum Mutat 30:1703–1712, 2009. © 2009 Wiley-Liss, Inc.

KEY WORDS: Illumina sequencing; pooling; association study; MUTYH; PCR

Introduction

Massive parallel sequencing (MPS) technology enables the molecular genetic analysis of diseases at an unprecedented scale, both in terms of the number of genes and the number of samples to be analyzed [Pettersson et al., 2009; Ropers, 2007]. Even though the 1000 Genomes project will give insight into human variation below a 1% allele frequency [Kaiser, 2008] (www.1000genomes.org), it will remain necessary to investigate rare variants in disease genes in selected patient cohorts. Targeted MPS is an attractive cost-efficient method to resolve rare and unique variants. MPS can be employed as a follow-up of a genome-wide association (GWA) study to extensively sequence regions surrounding associated single nucleotide polymorphisms (SNPs) or copy number variations (CNVs) [Goossens et al., 2009; Varley and Mitra, 2008; Yeager et al., 2008]. In this manner, the full extent of allelic variation might be recovered, e.g., in an 8- to 10-kb region of linkage disequilibrium (LD) or a large number of candidate genes [Bodmer and Bonilla, 2008; Houlston and Peto, 2004].

MPS capacity ensures a high sequence depth at specific loci, which can be exploited in different pooling strategies. Coverage depth of a DNA sample at a locus is primarily determined by the size of the resequenced region and capacity of the sequencing equipment. For example, with a total sequence yield of 1 Gb, each position of a 10-kb region can be read approximately 100,000 times. Multiple genomic regions, and/or DNA from multiple individuals, might be pooled without losing statistical power to detect unique variants. For detection of rare alleles, e.g., in DNA from viruses, tumors or pooled DNA, 0.02 to 1% has been reported as minimal detectable allele frequency [Campbell et al., 2008; Druley et al., 2009; Rohlin et al., 2009; Simen et al., 2009; Thomas et al., 2006; Wang et al., 2007]. However, no clear directions are available for the number of required reads on a position in order to reliably detect a rare allele at frequencies of 0.1 to 3% [Bodmer and Bonilla, 2008]. A limiting factor for the minimal detectable allele frequency in MPS data is the method's error rate, from which genuine variants should be significantly distinguished [Dohm et al., 2008; Wang et al., 2007]. The allele frequency of one heterozygous variant in a pool of samples, i.e., a singleton, is dependent on the pool size, which in turn needs to be adapted to the error rate.

Here we have evaluated the suitability of MPS by Illumina's Solexa sequencing technology (Illumina Inc., San Diego, CA) to detect rare variants and to accurately estimate allele frequencies in pooled samples. mutY (*E. coli*) human homolog (*MUTYH*; MIM# 604933) was analyzed as a candidate gene in DNA samples from breast cancer patients, since an association between *MUTYH* and breast cancer has been suggested [Nielsen et al., 2005]. Moreover, *MUTYH* is a relatively compact gene making it attractive for resequencing. We demonstrate that reliable detection of an allele frequency of 1% (2/176 alleles, doubletons) or higher is possible in a pool of 6-kb long-range PCR (LRP) products derived from 88

Additional Supporting Information may be found in the online version of this article.

Correspondence to: Astrid A. Out, Department of Clinical Genetics, Leiden University Medical Center, Post zone S6-P, P.O. Box 9600, 2300 RC, Leiden, The Netherlands.
E-mail: a.out@lumc.nl

© 2009 WILEY-LISS, INC.

samples. As expected, the detection of singleton variants (1/176, 0.6%) was limited by the error rate and dependent on the local coverage depth. Our data in combination with statistical power analyses for detection of unique and rare variants as a function of coverage depth provide directions for optimal design of high-throughput analysis of candidate gene regions.

Materials and Methods

DNA Samples, Pooling, and PCR

Patients were selected from a cohort of 1,359 breast cancer patients from whom DNA was collected between October 1996 and December 2005 at the Leiden University Medical Center and the Erasmus MC-Daniel Den Hoed Cancer Center in The Netherlands. Informed consent was obtained for DNA testing according to protocols approved by the local ethics review board. Genomic DNA was isolated from peripheral white blood cells by a salting out method [Huijts et al., 2007; Miller et al., 1988]. A 5,763-bp LRP fragment located at Chr1 (NCBI 36): g.45573093 to g.45567331 was designed, including 323 bp before exon 2 to 170 bp after the 3′UTR of the *MUTYH* gene (www.ncbi.nlm.nih.gov/GenBank). Two approaches for equimolar pooling were tested. First, a single LRP was performed on pooled genomic DNA of 287 patients. Automated fluorimetric quantification of DNA (PicoGreen® dsDNA Quantitation Kit; Molecular Probes, Eugene, OR), followed by dilution to 5 ng/μl, was performed using a Tecan Robot. From each sample, 2 μl (10 ng) were pooled. LRPs were performed with a TaKaRa LA PCR Kit (Version.2.1; TaKaRa Bio, Inc., Otsu, Shiga, Japan) according to manufacturer's instructions with minor adaptations. To the LRP, 75 ng pooled DNA were added; i.e., 0.26 ng per sample. LRP yield was measured by a Nanodrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). To generate 1 μg of LRP product, as required for the MPS procedure, products of five LRPs were pooled, followed by purification with a QIAquick® PCR purification kit (Qiagen GmbH, Hilden, Germany). A second pool was made from products of individual LRPs, performed on 125 of the 287 samples, on 30 ng of genomic DNA. The concentration of each LRP product was measured by high-resolution melting curve analysis (HR-MCA). LRPs that failed or generated products showing aberrant melting curves were excluded, leaving 88 LRP products of sufficient DNA yield and quality. These LRP products were pooled equimolarly. To minimize contaminating template DNA and primers, gel electrophoresis of the pooled LRP product was performed, after which the 6-kb band was cut out on a Dark Reader® Transilluminator (Clare Chemical Research, Dolores, CO) and purification was performed with a QIAquick® gel extraction kit (Qiagen).

Sanger Sequencing

Of the 287 samples present in the first pool (of genomic DNA), 217 had been sequenced by the Sanger method, including all exons, exon–intron boundaries, and UTRs of the *MUTYH* gene [Nielsen et al., 2005]. Given that in the MPS data from the pool of 287 samples the majority of singletons and doubletons were not detectable, and the frequencies of common variants deviated greatly from frequencies expected from Sanger data (Supp. Table S1), no further effort was put into Sanger sequencing of the remaining 70 samples. Hence, the common variant frequencies were derived from 217 samples. It was considered highly unlikely that this reduced sample size would dramatically affect the allele frequency estimates for the total pool, because all samples were randomly selected from the same population and the allele frequencies were comparable to those in the Database of Single Nucleotide Polymorphisms (dbSNP; [www.ncbi.nlm.nih.gov/projects/SNP]). All 88 samples, which were individually amplified in the second experiment, had been analyzed by Sanger sequencing. Additionally, three amplicons, one in intron 2 and two in intron 14, were designed for Sanger sequencing to confirm novel variants detected by MPS. Sanger sequenced areas (68% of the LRP product) are shown in Figure 1A.

Concentration Measurement by HR-MCA

To facilitate equimolar pooling, LRP products were analyzed with HR-MCA with the LCGreen®Plus+Melting Dye (Idaho Technology, Inc., Salt Lake City, UT). Fluorescence intensity was taken as a measure for DNA quantity and the melting curve shape for quality. To test if concentration measurement was accurate, the dilution series of one LRP product were analyzed (Fig. 2; Supp. Fig. S1). From each LRP product, 3 μl were added to a 10-μl solution containing a 1 × dye concentration and a 1 × concentration of PCR buffer (FastStart Taq DNA Polymerase PCR buffer; Roche Applied Science, Indianapolis, IN). Background fluorescence was measured in a 1 × dye solution and in blank LRPs. Melting and analysis were performed on an Idaho Lightscanner® instrument (Idaho Technology, Inc.), with temperature increasing from 35°C to 98°C, using Lightscanner® software with Call-IT™ 1.5 (Idaho Technology, Inc.). Fluorescence intensities of the LRP products were measured at the temperature at which the fluorescence of blank LRPs reached a level comparable to that of the dye only, at 60°C (Fig. 2A).

MPS and Data Analysis

Purified LRP product (1 μg) was fragmented by nebulization, from which a library was prepared according to the Illumina protocol. Fragments of approximately 250 bp were selected for MPS. A total of 4 pmol of the library were added to the flowcell lane (one lane per pool). Single-end sequencing of 36 cycles was performed on an Illumina Genome Analyzer I (GA1) instrument. GA1 images were analyzed using the Bustard and Firecrest software of the GAPipeline (Illumina). Filtering, mapping, and consensus analysis were performed with the SHORE software package (http://1001genomes.org/downloads/shore.html; and Supp. Methods) [Ossowski et al., 2008]. ELAND (Illumina) was used for ungapped and GenomeMapper (part of SHORE) for gapped alignments against the expected amplicon sequence selected from the human genome reference (NC_000001.9, g.45573093 to g.45567331). This sequence does not contain a 327-bp *Alu*-insertion-polymorphism (dbSNP: rs10527342), which we added to obtain a reference sequence of 6,090 bp. ELAND was run to align the first 32 bp of each 36-bp read, allowing for up to two mismatches, and the last 4 bp irrespective of mismatches. Subsequently, three sets of read alignments were selected for comparison featuring up to two, three, and four mismatches in total. GenomeMapper was run multiple times allowing for two to four mismatches and gaps. All best alignments were recorded and the resulting set of all alignments was transformed into a consensus summary. This summary contained the counts of all 4 nucleotides per position, separately for forward and reverse reads to detect strand-specific abnormalities. In addition, base counts were recorded after excluding the three nucleotides at the beginning and the end of each aligned read in order to minimize alignment ambiguities as well as to reduce the effects of sequencing errors, which are highest in the last positions

QIAGEN Exhibit 1002
Page 725

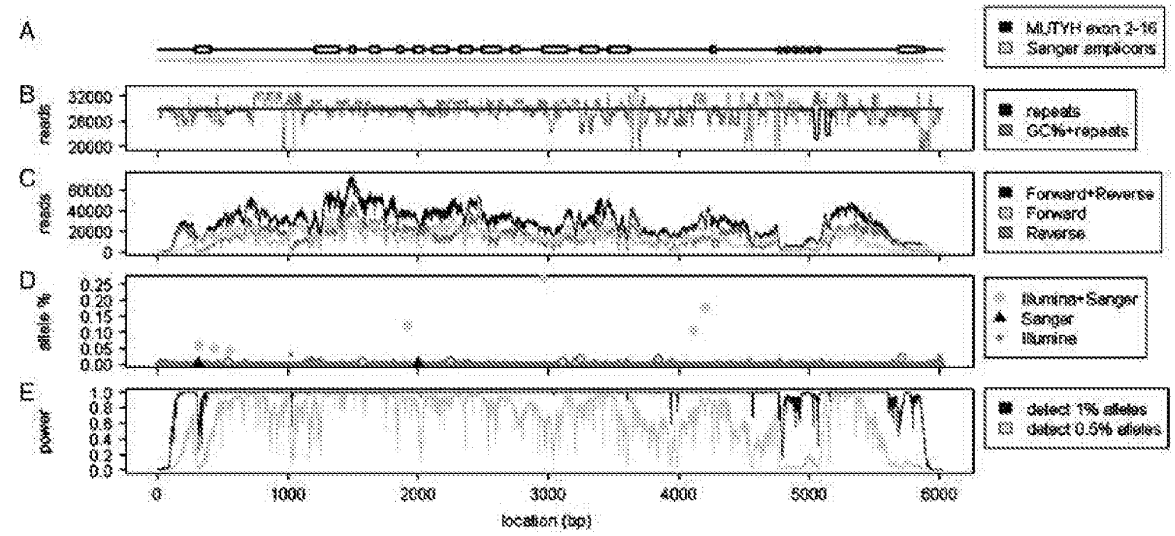


Figure 1. Coverage depth, allele frequencies, and statistical power along the 6-kb LRP fragment in the data from the 88 LRP products. **A:** *MUTYH* gene with exons 2–16 (boxes), 3′UTR (box with half height), and the *Alu*-insertion (box with serrated sides). The light gray/purple line below the gene picture is thick for areas covered and thin for areas not covered by Sanger sequencing. **B:** Expected coverage depth variation, on basis of an ELAND alignment of all 32-mers in the reference sequence, showing dips for nonunique regions (black), and on the basis of a combination of nonunique regions and GC% (green), calculated as described [Ossowski et al., 2008]. **C:** Obtained coverage depth for all, forward, and reverse reads. Depth is decreased at both fragment ends along 400 bp and at the location of the insertion. **D:** Detected variants and allele frequencies (mean of forward and reverse), by both Sanger and Illumina sequencing (light gray/purple), only by Sanger (black), and only by Illumina (dark gray/green). Two singletons were not detected by Illumina. **E:** Power for detection of allele frequencies 0.5% (singletons) and 1% (doubletons) above an error proportion of 0.31%. [Color figure can be viewed in the online issue, which is available at http://www.interscience.wiley.com.]

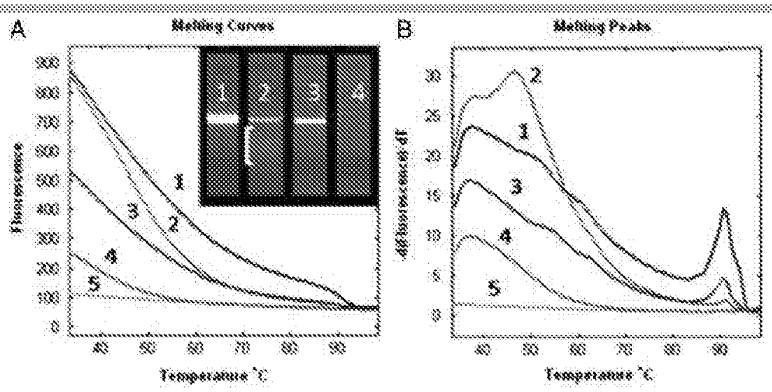


Figure 2. A representative selection of HR-MCA results from LRP products and control samples. The fluorescence intensity measurements appear proportionally comparable to intensities of bands on agarose gel and LRP products with aberrant bands on gel also show aberrant melting curves. **A:** Raw melting curves, and the corresponding gel pictures (insert). **B:** Derivative curves. (curves 1–3) LRP products: 2 shows aberrant curves, in (A) and (B), and also a faint smear on gel as marked by "[," probably consisting of nonspecific product. (curve 4) Blank LRP, which shows approximately the same fluorescence level and melting speed as the curve of the dye only, around 60°C. (curve 5) Only the dye LCGreen®Plus+. [Color figure can be viewed in the online issue, which is available at http://www.interscience.wiley.com.]

of the read. For each base in a read, two values for the quality of the base call were calculated by the GAPipeline image analysis. The quality value Q (–5 to 40), corresponds to the probability P (25–99.99%) that a base call is correct, where $Q = 10\log_{10}(P/(1-P))$ [Dohm et al., 2008]. The "chastity" value (50–100%) is used to filter out mixed clusters that have grown from two or more molecules that are too close together to be measured separately. It is calculated by dividing highest intensity of a called base through the sum of the highest and second highest intensity. Five Q-value cutoffs (5, 7, 11, 20, and 27) together with a chastity value cutoff of 67% were tested for single-base masking (Supp. Methods; Supp. Fig. S2), to evaluate their impact on error rate and rare allele calling.

Statistics

After excluding the primer sequences (both 34 bp), the remaining 6,022 positions were analyzed for minor allele frequencies. To

QIAGEN Exhibit 1002
Page 726

model sensitivity we adopted a Poisson-model approach similar to Wang et al. [2007]. Errors were assumed to occur at a constant proportion q, and a frequency described by a Poisson-distribution with rate $\lambda = Nq$, where N is the local coverage depth. The P-value for the null hypothesis that the number of occurrences of a minor allele r is due to sequencing errors is given by

$$P = 1 - \sum_{k=0}^{r-1} \frac{e^{-\lambda}\lambda^k}{k!}. \quad (1)$$

For the error proportion q_{null} under the null hypothesis we took the conservative value of 3.1×10^{-3}, which is at the 97.5th percentile of the observed distribution of minor allele frequencies (Supp. Fig. S3a). To adjust for multiple testing of $K = 3 \times 6{,}022$ possible minor alleles, we called a sequence variant only when $P \leq \sqrt{1-(1-\alpha)^{1/K}}$ for both the forward and reverse strands, with $\alpha = 0.05$. Under the assumption that sequencing errors in different locations and strands are independent, this P-value threshold makes sure that the family-wise error rate (FWER), the probability of a single, simultaneous, false detection on both strands, is bounded by α. Correcting for multiple testing based on the FWER is the most rigorous approach to multiple testing, and is in line with current practice in genetic association studies [Hoggart et al., 2008]. The threshold for a significant number of occurrences plotted against coverage depth is shown in Supp. Figure S3b.

To calculate the power to detect sequence variants we used the same Poisson-model, assuming that the rate q under the alternative hypothesis was $q = r/R + b$, where r is the number of alleles out of $R = 2 \times 88$ in which the variant occurs, and b is the error rate, which we estimated at 3×10^{-4}, the median of the observed minor allele frequencies (Supp. Fig. S3a). The addition of b to the expected rate r/R reflects the fact that, even for a true sequence variant, some of the observed reads of the sequence variant can be erroneous reads from the major allele. The different estimate of q_{null} and b is meant to accommodate, to some extent, the heterogeneity of the true error rate: a high estimate of q_{null} protects against false-positive detection of sequence variants in locations with a high error proportion, whereas the more moderate estimate of b based on the median reflects the fact that the presence of sequence variants is not related to the error proportion, and prevents inflated power estimates. Power as a function of coverage for different error proportions and allele frequencies is given in Figure 3. Statistical calculations and graphs were made with the Statistical Package for Social Studies version 16 (SPSS; SPSS Inc., Chicago, IL) and R version 2.8.1 [R Development Core Team; www.r-project.org]. R-scripts are available on request.

Results

Illumina Sequencing Results of the Pool of Genomic DNA

Illumina sequencing of a single LRP product generated from an equimolar pool of 287 genomic DNA samples delivered 3.9 million reads. Of these reads, 1.7 million (43%) mapped to the reference sequence, resulting in a mean coverage of 8,800 per site (30 per site per sample). A poor match was observed between allele frequencies determined by Sanger and Illumina sequencing (Supp. Table S1). The most extreme difference was a minor allele frequency of 29% in the Illumina data for one variant, which was 5% according to the Sanger data. This discrepancy suggested that not all individuals contributed equally to the LRP product. In addition, the error proportion was too high to detect rare variants (i.e., 1–4 alleles in 287 samples; 0.2–0.9%).

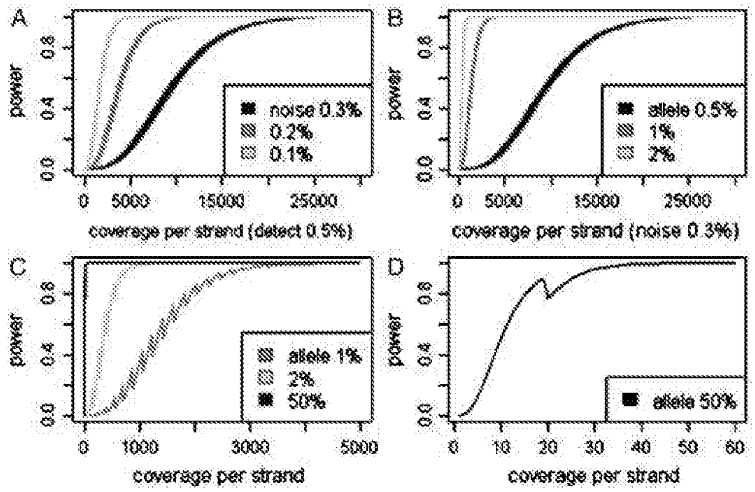


Figure 3. Power for detection of different allele frequencies above error proportion in relation to coverage depth on one strand (see Materials and Methods). **A:** Depth per strand required for detection of an 0.5% allele frequency above different error proportions (0.1%, 0.2%, and 0.3%). **B:** Depth required for detection of different allele frequencies above an error proportion of 0.3%. **C:** Like (B), zoomed in on depth below 5,000. **D:** Like (B), zoomed in depth below 60. An optimal difference between the error proportion and the singleton frequency can be obtained by decreasing the pool size; e.g., for reliable detection of 2% (1 allele in 25 samples) a depth of 1,000 (40 per sample) per strand is needed, whereas for 0.5% (1 allele in 100 samples) the required depth is 20,000-fold (200 per sample). Predicted depth required for reliable detection of a 50% allele frequency (D) was 30-fold per strand. The dents in this power function arise as a consequence of the discreteness of the Poisson distribution, which makes the significance level strictly smaller than α for many coverage values. This results in a small loss of power for these coverage values. [Color figure can be viewed in the online issue, which is available at http://www.interscience.wiley.com.]

QIAGEN Exhibit 1002
Page 727

Illumina Sequencing Results of the Pool of LRP Products

To eliminate the possible influence of DNA concentration and PCR efficiency of individual samples on the equimolar constitution of the pool, we performed Illumina sequencing of a pool of 88 LRP products. This produced 11.1 million reads, of which 7.2 million (65%) mapped to the reference sequence of *MUTYH* with three mismatches or less, resulting in a mean coverage depth of 34,983 reads per bp (400 per bp per sample). After per-base quality filtering (thresholds of 67% for chastity and 7 for *Q*-value), 28,746 per bp depth remained (82%). In this analysis, 15 of 17 variants detected by Sanger sequencing were called in the Illumina data (labeled as true positives; see Tables 1 and 2). Four of these 15 were variants newly detected with Illumina sequencing and confirmed by Sanger sequencing, in the middle introns. Three of these were present in dbSNP, whereas one, a singleton, was new (NM_001128425.1: c.157+136G>A). In silico analyses revealed no clues for pathogenicity (Supp. Tables S2 and S3). Furthermore, two A>G changes were detected that were not seen in the Sanger data (labeled as false positives) and three T>C changes were detected outside the regions covered using Sanger sequencing (to be confirmed).

There was a good concordance between allele frequency estimates by Illumina and Sanger sequencing (Pearson correlation coefficient 0.99). However, frequencies in the Illumina data were generally overestimated. We reasoned this could have been caused by a known *Alu*-insertion-polymorphism (see Materials and Methods). Gel electrophoresis of individual LRP products revealed two fragments, of 5.8 and 6.1 kb. Additionally, the presence of the insertion sequence was confirmed in a few homozygous samples by Sanger sequencing. Of the 88 samples, 48 were heterozygous and 22 were homozygous for the insertion, yielding an allele frequency of 52%. In heterozygotes, the intensity of the longer LRP product was lower compared to that of the shorter one (Supp. Fig. S4). Variants in LD with the wild-type or the insertion allele were inferred from the Sanger sequencing data and the bands on gel. In heterozygotes we applied a correction factor, assuming that the smallest allele was on average amplified twice as much as compared to the largest allele (Supp. Table S4). This correction improved the correspondence between variant frequencies as determined by Illumina and Sanger (Tables 1 and 2; Fig. 4).

Optimization of Allele Frequency Assessment

To assess the critical parameters for alignment and allele frequency prediction we compared the results obtained with different settings of SHORE. We found that using a quality threshold for base masking (*Q*5–*Q*27) and additional base masking with a chastity threshold substantially influenced the error rate. Furthermore, these quality thresholds influenced the allele call rate (i.e., the false positives and false negatives, as compared to the Sanger sequencing data; Table 1; Supp. Table S5). In addition, the error rate and the number of positive calls, but not the true positives, could be reduced by discarding bases at the beginning and the end of each aligned read (equal to the number of allowed mismatches). This truncation reduces alignment errors when allowing for gaps but also reduces the effective read depth per site. Altering the number of allowed mismatches (two, three, and four) did not lead to large differences in allele calling. The optimal ratio of true allele calls to false positives was obtained with the following settings: three or fewer mismatches (in 36-bp reads), truncated ends of aligned reads (3 bp discarded at each end), a *Q*-value threshold of 7, and a chastity threshold of 67%.

SHORE predictions with the optimal parameter settings showed a Pearson correlation coefficient of 0.99 to the Sanger results independent of the alignment tool. This correlation was marginally higher (at the third decimal place) in the data not filtered for quality and chastity (Table 1). This effect was mainly caused by a change in frequencies of two common variants, which were affected by a filtering bias against G-bases (Fig. 4; Table 2; and Supp. Table S6). By filtering, correct bases were also discarded (Supp. Fig. S2). In this experiment the chastity filtering showed the most prominent effect, which was probably due to intensity differences between the different bases (run-dependent).

Gapped alignments revealed two potential deletions, both in homopolymer runs. A deletion of a T (8% of alleles) in a stretch of nine Ts and a deletion of one A in seven As (1%) were detected. However, both deletions were not observed in the Sanger data. No significant differences in calling of false negatives and false positives could be found between the ELAND and GenomeMapper alignment.

Detection of Rare Variants and Sequence Depth

All variants with an expected frequency of a doubleton or higher (1.1% and 3.4–21.6%), were significantly detected in the Illumina data (0.7–1.8% and 3.9–26.3%), with only a few false-positive predictions. Of the five singleton variants in the Sanger data, three were readily detected by Illumina sequencing (0.5–0.6%). The other two (0.3%) were close to the error proportion q_{null} (0.31%) (Table 2). Taking a lower percentile as error proportion (e.g., the 92.5th percentile of 0.21%) led to calling of many discrepancies with the Sanger data (false positives, mostly A>G and T>C changes), without additional true positives (Supp. Table S5).

The statistical power in relation to sequence depth was calculated for detection of different allele frequencies in our experimental setup. For detection of singletons, read depths of 25,000 and 15,000 per site per strand were predicted to be required for reaching statistical power $(1-\alpha)$ of 1.0 and 0.9, respectively. The obtained sequence depth from the 88 LRP products approximately followed a normal distribution with a mean of 14,000 reads per site per strand, and 99% of the positions were covered over 2,200-fold on each strand, with exclusion of the insertion region and 400 bp at each fragment end. Distributions and percentiles of coverage depth are shown in Supp. Table S7 and Supp. Figure S5. The depth was partly sufficient for detection of singletons with high power along the majority of the 6-kb fragment. The observed sequence depth appeared more than sufficient for detection of doubleton variants with statistical power of 1.0 along almost the entire fragment (Figs. 1 and 3).

To assess the influence of repeat and GC-content of the sequence on the coverage depth in our experiment, two predictions of depth were made. ELAND was used to uniquely align reads to the reference sequence. Repetitive regions in the reference sequence, expected to show decreased depth of aligned reads, were detected by alignment of all 32-mers present in the reference sequence uniquely back to the reference sequence. In this alignment, sequence depth was significantly lower in three small repetitive regions. A second prediction of the coverage depth was made by taking into account both GC-content and repetitive regions around each site. Depth was empirically determined in relation to GC-content in a 100-bp window and repeats aligning with up to three mismatches (Fig. 1B), as performed in Ossowski et al. [2008]. Coverage depth obtained by Illumina sequencing of

QIAGEN Exhibit 1002
Page 728

Table 1. Analysis Settings vs. Correlation and Discordances Between Illumina and Sanger Sequencing Data

Settings						Descriptive statistics						Variant detection		
Data	Aligner	Q	C (%)	MM	Truncation	Mean depth	Mean MAF	Median MAF	Correlation before/after correction[a]	PCTL	PCTL MAF = q	I−; S+ 0.6% (5), 1.1% (4), 1.7% (1)[b]	I+; S− (A>G or T>C)	I+; S NA (A>G or T>C)
All	Eland	**0**	**50**	3	Yes	34983	0.0046	0.0022	0.994/0.995	97.5	0.022	5, 4, 1	5 (0)	0
All	Eland	**5**	**50**	3	Yes	32968	0.0016	0.00096	0.994/0.996	97.5	0.0063	5, 1	1 (1)	0
All	Eland	7	**50**	3	Yes	31976	0.0013	0.00076	0.994/0.996	97.5	0.0050	5, 1	1 (1)	1 (1)
All	Eland	**5**	67	3	Yes	28977	0.00073	0.00033	0.986/0.988	97.5	0.0036	3	0	3 (3)
All	**Eland**	**7**	**67**	**3**	**Yes**	**28746**	**0.00071**	**0.00032**	**0.986/0.988**	**97.5**	**0.0031**	**2**	**2 (2)**	**3 (3)**
										92.5	0.0021	2	26 (26)	11 (11)
All	Eland	**11**	67	3	Yes	27585	0.00068	0.00029	0.986/0.988	97.5	0.0032	4	2 (2)	3 (3)
All	Eland	**20**	67	3	Yes	22697	0.00065	0.00023	0.981/0.982	97.5	0.0035	5, 2, 1	2 (2)	3 (3)
All	Eland	**27**	67	3	Yes	18873	0.00066	0.00019	0.976/0.977	97.5	0.0039	5, 2, 1	1 (1)	2 (2)
All	Eland	7	67	3	**No**	34097	0.00079	0.00037	0.987/0.989	97.5	0.0033	2	2 (1)	7 (5)
All	Eland	7	67	**2**	Yes	29516	0.00065	0.00029	0.985/0.987	97.5	0.0030	2	2 (2)	4 (3)
All	Eland	7	67	**4**	Yes	27200	0.00072	0.00032	0.985/0.988	97.5	0.0032	3	1 (1)	3 (3)
All	**Genome-Mapper**	7	67	3	Yes	29459	0.00083	0.00039	0.989/0.990	97.5	0.0035	3	0, 2[c]	5 (3)
50%	Eland	7	67	3	Yes	14377	0.00071	0.00030	0.986/0.988	97.5	0.0033	3	0	2 (2)
25%	Eland	7	67	3	Yes	7185	0.00071	0.00026	0.988/0.990	97.5	0.0034	5, 1, 1	0	0
										92.5	0.0021	3	0	2 (2)

The analysis that was chosen for variant detection (fifth from above) and the altered parameters in the other analyses are in bold.

[a]Pearson correlation coefficient before and after correction for amplification bias caused by the presence of the *Alu*-insertion-polymorphism (Supp. Table S4).

[b]In total, 17 variants were found by Sanger sequencing. Seven had a frequency of 3.4% or higher and were detected with all settings. Five, four, and one had frequencies of 0.6%, 1.1%, and 1.7% in the Sanger sequencing data, respectively. Of these 10 low frequency variants zero to eight were detected with the different settings.

[c]Single-nucleotide changes and 1-bp deletions.

Q, Illumina quality value; C, chastity; MM, mismatches; MAF, minor allele frequency; PCTL, percentile; q, error proportion used in the statistical formula to distinguish true variants from errors; S, Sanger; I, Illumina; NA, not analyzed; "–," not detected single-nucleotide changes; "+," detected single-nucleotide changes.

Table 2. Detected Variants (24) by Illumina Sequencing (22), Sanger Sequencing (17), or Both (15) in the 88 Pooled LRP-Products

Position in fragment	Variant cDNA	Illumina allele counts (F+R)					Illumina minor allele %				Sanger		Detection with	
		A	C	G	T	–	F	R	F+R	Corrected	Allele %	Count	Illumina	Sanger
343	c.56G>A[a,b,c]	**46**	1	13828	3		0.3	0.2	0.3	0.5	0.6	1	–	+
351	c.64G>A[a]	**808**	1	13411	1		5.6	6.3	5.7	5.3	5.7	10	+	+
474	c.157+30A>G	19633	7	**970**	1		4.6	4.8	4.7	3.9	3.4	6	+	+
580	c.157+136G>A[d]	**199**	3	31065	7		0.7	0.6	0.6	0.6	0.6	1	+	+
595	c.157+151G>A[e]	**1440**	5	35883	6		3.8	3.9	3.9	3.9	4.0	7	+	+
1228	c.158−46A>G[f]	19944	48	**159**	15		0.5	1.6	0.8	?	0	0	+	–
1282	c.166A>G[f]	12078	25	**108**	8		0.5	1.5	0.9	?	0	0	+	–
1606	c.388+56G>A	**317**	5	51166	16		0.5	0.7	0.6	0.6	0.6	1	+	+
1952	c.504+35G>A[h]	**4721**	0	34574	16		12.4	11.8	12.0	10.3	10.8	19	+	+
2032	c.536A>G[c]	25881	16	**82**	5		0.3	0.4	0.3	0.5	0.6	1	–	+
2282	c.690+21C>A	**386**	47402	1	26		0.7	1.1	0.8	1.0	1.1	2	+	+
3002	c.1014G>C[i]	14	**5684**	15937	3		28.7	22.4	26.3	23.5	21.6	38	+	+
3141	c.1153G>A	**130**	3	26678	14		0.5	0.5	0.5	?	0.6	1	+	+
3252	c.1187−27C>T	12	27888	2	**191**		0.7	0.7	0.7	?	1.1	2	+	+
3279	c.1187G>A	**458**	5	29929	2		1.2	1.9	1.5	1.3	1.1	2	+	+
3742	c.1476+89delT[g,j]	5	59	11	23520	**2084**	9.1	7.3	8.1	?	0	0	+	–
3881	c.1476+228C>T[e]	6	17891	2	**250**		1.8	0.9	1.4	1.4	1.7	3	+	+
4153	c.1477−125C>A[e]	2775	24257	12	47		8.7	12.8	10.2	9.5	7.4	13	+	+
4238	c.1477−40G>C[h,i]	9	**3982**	18984	12		24.5	10.0	17.3	14.8	10.8	19	+	+
4494	c.1518+175delA[g,j]	24041	5	65	2	**279**	1.1	1.1	1.1	?	0	0	+	–
4584	c.1518+265T>C[f,g]	0	**105**	6	12741		0.9	0.7	0.8	?	NA	NA	+	NA
4809	c.1518+490T>C[f,g]	0	**62**	0	6699		0.8	1.1	0.9	?	NA	NA	+	NA
4812	c.1518+493T>C[f,g]	0	**50**	0	4048		1.2	1.2	1.2	?	NA	NA	+	NA
5750	c.1544C>T	1	8160	0	**149**		1.7	1.8	1.8	1.5	1.1	2	+	+

Illumina data were obtained from the 5th analysis in Table 1. Variants were considered as detected when their frequency was significantly above 3.1×10^{-3} in the described statistical model. cDNA nomenclature is according to HGVS guidelines (www.hgvs.org), NCBI sequence NM_001128425.1 (see also Supp. Table S2). For the Illumina data, allele counts obtained from all reads (both strands), allele frequencies for reads in forward (F), reverse (R), and both directions and frequencies corrected (both strands) for amplification bias caused by the presence of the *Alu*-insertion-polymorphism (Supp. Table S4) are given. For the Sanger data allele counts and allele frequencies are shown. Detection by Illumina and Sanger sequencing is indicated by "+" if detected, "–" if not (significantly) detected, and "NA" if not analyzed.

[a]Found in the same sample with Sanger sequencing, but in separate reads with Illumina sequencing.

[b]Low reverse coverage.

[c]Amplified less, because most probably in LD with the insertion allele (based on data from this experiment and unpublished data from our laboratory, see Supp. Table S4).

[d]Newly detected by Illumina sequencing, unknown in dbSNP.

[e]Newly detected by Illumina sequencing, known in dbSNP.

[f]Using a lower error proportion q_{null} in the formula described in methods led to calling of many false-positive A>G and T>C changes (Table 1; Supp. Table S5).

[g]Located at the end of a homopolymer (T or A).

[h]In 100% LD with each other in the 88 samples.

[i]The difference between allele frequency in forward and reverse reads increased after filtering (see Fig. 4 and Supp. Table S6).

[j]Detected with the 9th analysis in Table 1.

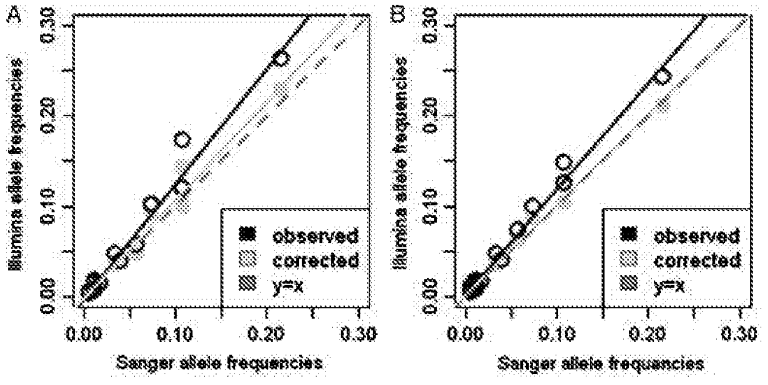


Figure 4. Allele frequencies observed in the Illumina sequencing data from the 88 pooled LRP products compared to frequencies observed in the Sanger sequencing data, of the 17 variants detected by Sanger sequencing. Of these variants, 15 were significantly detected in the Illumina data, whereas two were observed, although not significantly above the error proportion. Observed frequencies are shown in black and frequencies corrected for preferential amplification due to the *Alu*-insertion in gray/purple. Linear fit lines and an y = x line (dashed) are shown. Correction moved the frequencies toward the y = x line. **A:** Filtered data (*Q*-value cutoff 7, chastity cutoff 0.67). **B:** Unfiltered data. Frequencies of two common alleles (10% and 25%) were affected by filtering (see also Table 2 and Supp. Table S6). [Color figure can be viewed in the online issue, which is available at http://www.interscience.wiley.com.]

QIAGEN Exhibit 1002
Page 730

the 88 LRP products showed a variable pattern, which corresponded well, but not perfectly, to the predicted coverage depths (Fig. 1C). Coverage depth decreased markedly at both ends of the LRP fragment (400 bp), more pronounced at the beginning for forward and at the end for reverse reads. Depth was also reduced in the region of the *Alu*-insertion-polymorphism, to approximately one-third of the average depth, reflecting its underrepresentation in the pool.

The obtained coverage depth appeared more than sufficient for detection of doubletons, as revealed by analyzing random subsets of 50% and 25% of our Illumina sequencing output. In the 25% subset, above the error proportion q_{null} of 0.34%, five singletons, one doubleton, and one variant with a frequency of 3 in 176 were not detected. However, when taking the 92.5th percentile of minor allele frequencies as q_{null} (0.21%) in this dataset, the number of false positives remained low and all variants except three singletons were detected (Table 1; Supp. Table S5; and Supp. Fig. S6).

Discussion

We demonstrate that use of Illumina technology and the SHORE pipeline is highly appropriate for rare variant detection in a candidate gene analyzed in pooled samples. Analysis of a pool of 88 normalized 6-kb LRP products in one flow cell lane results in reliable detection of alleles with an occurrence of at least 2 in 176 (1.1%). Importantly, allele frequencies detected by Sanger and Illumina sequencing highly correlate. Furthermore, three out of five singleton variants are detected above error rate in the Illumina data, depending on the depth and the observed frequency. Thus our data suggest that sequencing with the Illumina GA1 can reliably detect variants down to a 0.5% minor allele frequency, corresponding to a singleton variant in a pool of 100 diploid samples. This efficient and broadly applicable approach proves the suitability of Illumina technology for candidate gene analysis.

Characteristics of DNA samples may require different pooling strategies. Pooling of 287 genomic DNA samples, followed by one LRP, provided inaccurate allele frequency estimates and rare variants were undetectable. The latter is attributable to insufficient coverage depth in relation to the pool size. Equimolar amplification in the LRP may be hampered by factors like unequal quality of the DNA samples and, in this case, preferential allelic amplification due to a common large insertion-polymorphism. A plausible and relevant issue in a clinical setting is that collection of DNA samples over many years may involve variability in quality due to gradual changes in the DNA isolation protocol, variable storage time, and usage. Genome centers typically analyze large sets of DNA samples extracted in a high-throughput automated fashion, most likely yielding homogeneous DNA quality. These centers tend to apply the pooling step first, followed by LRP [Unnur Thorsteinsdottir, Decode Genetics; David Altschuler, Broad Institute, MIT; personal communication]. Reliable results with Roche 454 sequencing have been demonstrated in a pool of 96 clinical genomic DNA samples, after either starting with pooling or with LRP, for validated allele frequencies down to 3.7% [Ingman and Gyllensten, 2009]. However, it can be difficult to measure genomic DNA quality and to control uneven sample representation in the LRP. First applying LRP, followed by equimolar pooling of measured products, provides more certainty about the contents of the pool. Measurements of both quantity and quality with HR-MCA have enabled us to exclude LRP products with a high one-point fluorescence level, but aberrant quality. Our accurate data obtained from the pool of 88 LRP products, support the validity of this easy and high-throughput method.

On the basis of our data, we have computed the correlation between sequence depth per nucleotide and statistical power to detect minor allele frequencies. The correlations depicted in Figure 3 provide insight into the variables that determine the potential possibilities of MPS. In our dataset obtained from 88 samples the power is much higher for detection of doubletons as compared to singletons, which can be explained by a singleton allele frequency close to the error rate. Sequence depth per basepair is approximately the ratio between the instrument's sequencing capacity and the size of the target region, while the pool size determines the allele frequency of a singleton variant. Thus, to reliably detect a singleton (i.e., power > 0.95) one can simply reduce the pool size. Importantly, the allele frequencies detected by Illumina show variation around the expected frequency (0.2–0.7% for singletons), probably arising from errors in the pooling and the sequencing procedure. This variation needs to be taken into account when calculating the aimed depth for an experiment.

Although sequence depth is an important indicator of sensitivity, the lower detection limit is determined by the error rate. The error rate per site remaining after quality filtering (97.5th percentile, 0.31%), probably mainly consists of polymerase errors in the different amplification steps [Campbell et al., 2008; Dohm et al., 2008]. The error rate and data quality are experiment-dependent. Determining and applying optimal quality filter settings, with the SHORE pipeline, proves crucial for obtaining an optimal balance between detection of true and false positives.

At the positions of the detected singleton variants, sequence depth was markedly higher as compared to positions of undetected singletons. Sequence depth variation is dependent on, among others, stochastic effects and sequence characteristics [Bentley et al., 2008; Dohm et al., 2008; Ossowski et al., 2008]. A computationally predicted depth distribution along the 6-kb fragment only partly correlated with the obtained depth. A model has been proposed for calculation of the haploid redundancy (i.e., the total number of sequenced bases divided by the size of the sequenced region) required for an optimal fraction of the region to be covered with sufficient reads [Wendl and Wilson, 2008]. According to this model, a haploid redundancy of 21-fold is required for covering an optimal fraction (0.99996) of both alleles in a human genome with at least one read. For a short fragment and many alleles a different model may be required. We have obtained a haploid redundancy of 28,000-fold, covering 99% of the 6-kb fragment at least 2,200 deep per strand (sufficient for detection of 1% alleles with power 0.8), with exclusion of the insertion-polymorphism and 400 bp at each end of the fragment. Depth was reduced at the fragment ends. This can be explained by edge-bias, which depends on the size of the sequenced region, the selected fragmentation size, and the read length [Wendl and Barbazuk, 2005]. Increased fragment size (600 bp instead of 200 bp) and blocked primers have been shown to improve sequence coverage uniformity [Harismendy and Frazer, 2009]. Bentley et al. [2008] describe that 30-fold haploid redundancy allowed for detection of an optimal proportion of heterozygous variants in the sequence of an individual genome. Our model predicts that read depth of 30–40 per strand is required for detection of a 50% allele frequency with power 0.95–0.99. Separation of reads per strand for detection of such a high frequency may not be necessary, suggesting a too conservative estimate. Overall, these numbers indicate that haploid redundancy in an experiment needs to be set in great excess (at least 10–20×) of the minimal required depth per position.

QIAGEN Exhibit 1002
Page 731

MPS provides the opportunity to analyze a single LRP product in multiple tagged subpools [Meyer et al., 2008; Parameswaran et al., 2007]. Alternatively, subpools can involve a combination of different parts of the genome [Yeager et al., 2008]. The tangible optimal pool size and number of subpools needs to be determined by experiments and cost-utility calculations. Furthermore, LRP is an attractive method for selecting a region of interest in case of many samples, a relatively small number of target regions, and a limited target size. With fewer samples, but larger or many regions, use of hybridization capturing may be more attractive [Albert et al., 2007; Bau et al., 2009].

Comparable studies with Roche 454 sequencing show detectable allele frequencies of 0.2 to 1% [Simen et al., 2009; Thomas et al., 2006; Wang et al., 2007]. One study distinguishes subclones from each other down to 1 in 5,000 (0.02%) in mononuclear cell samples from leukemia patients, by sequencing a hypervariable region of 250 bp, approximately the read length, and by haplotype clustering [Campbell et al., 2008]. A recent study with Illumina sequencing validated detection of multiple variants down to 0.5% and showed detection of one singleton in 2,222 pooled alleles (0.05%), with SNPSeeker [Druley et al., 2009]. This algorithm allows calling variants below the raw error rate, in bases 3–12 from each read (low error rate). Furthermore, it builds a per run error model based on a control sequence without SNPs, taking into account positions and sequence context of errors in reads [Druley et al., 2009]. However, they did not make use of Illumina quality values. A combination of an advanced error model and quality threshold might facilitate detection of variants in the entire reads, giving more depth. Further development of algorithms is needed to optimize rare variant detection.

The present study provides insight in the potential of pooling strategies to be applied in large association studies, especially aiming at variant discovery, for small to moderately large genomic regions of, e.g., a few to a few hundred kilobases. We determined important parameters in the analysis and in the design of a pooling experiment. We conclude that in a clinical setting, sample variability due to collection over time requires adaptation of protocols for optimal results. MPS provides excellent possibilities for exploring candidate genes for all variants in large groups. Substantial increase in throughput (as lately announced by Illumina) and new developments like Pacific Biosciences Single Molecule Real Time (SMRT)^{TM} technology [Eid et al., 2009], will further lower sequencing costs and should further increase our ability to effectively conduct large-scale association studies.

Acknowledgments

We thank Petra Huijts, Karin Kroeze, and Maaike Vreeswijk for their management of and laboratory work on the DNA samples used; Matthew Hestand and Jaap van der Heijden for skillful assistance in bioinformatics; and Rolf Vossen for sharing his expertise with HR-MCA. Finally, we acknowledge all anonymous breast cancer patients, and the clinicians and researchers (Leiden University Medical Center [LUMC], Leiden, The Netherlands, and Daniel Den Hoed Kliniek [DDHK], Rotterdam, The Netherlands) involved in supplying DNA and clinical data. Grant sponsors: Dutch Cancer Society (KWF-UL-2006-3601) and the Max Planck Society.

References

Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, Richmond TA, Middle CM, Rodesch MJ, Packard CJ, Weinstock GM, Gibbs RA. 2007. Direct selection of human genomic loci by microarray hybridization. Nat Methods 4:903–905.

Aretz S, Uhlhaas S, Goergens H, Siberg K, Vogel M, Pagenstecher C, Mangold E, Caspari R, Propping P, Friedl W. 2006. MUTYH-associated polyposis: 70 of 71 patients with biallelic mutations present with an attenuated or atypical phenotype. Int J Cancer 119:807–814.

Bau S, Schracke N, Kränzle M, Wu H, Stähler PF, Hoheisel JD, Beier M, Summerer D. 2009. Targeted next-generation sequencing by specific capture of multiple genomic loci using low-volume microfluidic DNA arrays. Anal Bioanal Chem 393:171–175.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Cheetham RK, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Catenazzi MCE, Chang S, Cooley RN, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Furey WS, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Jones TAH, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling NB, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris PD, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Sohna JE, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59.

Bodmer W, Bonilla C. 2008. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 40:695–701.

Campbell PJ, Pleasance ED, Stephens PJ, Dicks E, Rance R, Goodhead I, Follows GA, Green AR, Futreal PA, Stratton MR. 2008. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. Proc Natl Acad Sci USA 105: 13081–13086.

Cheadle JP, Sampson JR. 2007. MUTYH-associated polyposis—from defect in base excision repair to clinical genetic testing. DNA Repair (Amst) 6:274–279.

Cleary SP, Cotterchio M, Jenkins MA, Kim H, Bristow R, Green R, Haile R, Hopper JL, LeMarchand L, Lindor N, Parfrey P, Potter J, Younghusband B, Gallinger S. 2009. Germline MutY human homologue mutations and colorectal cancer: a multisite case-control study. Gastroenterology 136:1251–1260.

Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda, MD: National Center for Biotechnology Information, National Library of Medicine (dbSNP Build ID: 129). Available at: www.ncbi.nlm.nih.gov/projects/SNP. Accessed 27 August 2009.

den Dunnen JT, Antonarakis SE. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat 15:7–12.

den Dunnen JT, Paalman MH. 2003. Standardizing mutation nomenclature: why bother? Hum Mutat 22:181–182.

Dohm JC, Lottaz C, Borodina T, Himmelbauer H. 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res 36:e105.

Druley TE, Vallania FL, Wegner DJ, Varley KE, Knowles OL, Bonds JA, Robison SW, Doniger SW, Hamvas A, Cole FS, Fay JC, Mitra RD. 2009. Quantification of rare allelic variants from pooled genomic DNA. Nat Methods 6:263–265.

Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-time DNA sequencing from single polymerase molecules. Science 323: 133–138.

QIAGEN Exhibit 1002
Page 732

Goossens D, Moens LN, Nelis E, Lenaerts AS, Glassee W, Kalbe A, Frey B, Kopal G, De Jonghe P, De RP, Del-Favero J. 2009. Simultaneous mutation and copy number variation (CNV) detection by multiplex PCR-based GS-FLX sequencing. Hum Mutat 30:472–476.

Grantham R. 1974. Amino acid difference formula to help explain protein evolution. Science 185:862–864.

Harismendy O, Frazer K. 2009. Method for improving sequence coverage uniformity of targeted genomic intervals amplified by LR-PCR using Illumina GA sequencing-by-synthesis technology. Biotechniques 46:229–231.

Hoggart CJ, Clark TG, De Iorio M, Whittaker JC, Balding DJ. 2008. Genome-wide significance for dense SNP and resequencing data. Genet Epidemiol 32:179–185.

Houlston RS, Peto J. 2004. The search for low-penetrance cancer susceptibility alleles. Oncogene 23:6471–6476.

Huijts PE, Vreeswijk MP, Kroeze-Jansema KH, Jacobi CE, Seynaeve C, Krol-Warmerdam EM, Wijers-Koster PM, Blom JC, Pooley KA, Klijn JG, Tollenaar RA, Devilee P, van Asperen CJ. 2007. Clinical correlates of low-risk variants in FGFR2, TNRC9, MAP3K1, LSP1 and 8q24 in a Dutch cohort of incident breast cancer cases. Breast Cancer Res 9:R78.

Ingman M, Gyllensten U. 2009. SNP frequency estimation using massively parallel sequencing of pooled DNA. Eur J Hum Genet 17:383–386.

Kaiser J. 2008. DNA sequencing. A plan to capture human diversity in 1000 genomes. Science 319:395.

Küry S, Buecher B, Robiou-du-Pont S, Scoul C, Colman H, Lelièvre B, Olschwang S, Le Houérou C, Le Neel N, Faroux R, Ollivry J, Lafraise B, Chupin LD, Bèzieau S. 2007. The thorough screening of the MUTYH gene in a large French cohort of sporadic colorectal cancers. Genet Test 11:373–379.

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

Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian SV. 2006. Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. Nucleic Acids Res 34:1317–1325.

Meyer M, Stenzel U, Hofreiter M. 2008. Parallel tagged sequencing on the 454 platform. Nat Protoc 3:267–278.

Miller SA, Dykes DD, Polesky HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215.

Nielsen M, Franken PF, Reinards TH, Weiss MM, Wagner A, van der Klift H, Kloosterman S, Houwing-Duistermaat JJ, Aalfs CM, Ausems MG, Brocker-Vriends AH, Gomez Garcia EB, Hoogerbrugge N, Menko FH, Sijmons RH, Verhoef S, Kuipers EJ, Morreau H, Breuning MH, Tops CM, Wijnen JT, Vasen HF, Fodde R, Hes FJ. 2005. Multiplicity in polyp count and extracolonic manifestations in 40 Dutch patients with MYH associated polyposis coli (MAP). J Med Genet 42:e54.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205–217.

Ossowski S, Schneeberger K, Clark RM, Lanz C, Warthmann N, Weigel D. 2008. Sequencing of natural strains of *Arabidopsis thaliana* with short reads. Genome Res 18:2024–2033.

Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ. 2007. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Res 35: e130.

Pettersson E, Lundeberg J, Ahmadian A. 2009. Generations of sequencing technologies. Genomics 93:105–111.

R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: www.r-project.org. Accessed 27 August 2009. [ISBN 3-900051-07-0.]

Rohlin A, Wernersson J, Engwall Y, Wiklund L, Björk J, Nordling M. 2009. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. Hum Mutat 30:1012–1020.

Ropers HH. 2007. New perspectives for the elucidation of genetic disorders. Am J Hum Genet 81:199–207.

Simen BB, Simons JF, Huppler Hullsiek K, Novak RM, Macarthur RD, Baxter JD, Huang C, Lubeski C, Turenchalk GS, Braverman MS, Desany B, Rothberg JM, Egholm M, Kozal MJ. 2009. Low-abundance drug-resistant viral variants in chronically HIV-infected, antiretroviral treatment-naive patients significantly impact treatment outcomes. J Infect Dis 199:693–701.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989.

Stephens M, Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73: 1162–1169.

Tavtigian SV, Byrnes GB, Goldgar DE, Thomas A. 2008. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. Hum Mutat 29:1342–1354.

Thomas RK, Nickerson E, Simons JF, Jänne PA, Tengs T, Yuza Y, Garraway LA, LaFramboise T, Lee JC, Shah K, O'Neill K, Sasaki H, Lindeman N, Wong KK, Borras AM, Gutmann EJ, Dragnev KH, DeBiasi R, Chen TH, Glatt KA, Greulich H, Desany B, Lubeski CK, Brockman W, Alvarez P, Hutchison SK, Leamon JH, Ronan MT, Turenchalk GS, Egholm M, Sellers WR, Rothberg JM, Meyerson M. 2006. Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing. Nat Med 12:852–855.

Varley KE, Mitra RD. 2008. Nested Patch PCR enables highly multiplexed mutation discovery in candidate genes. Genome Res 18:1844–1850.

Wang C, Mitsuya Y, Gharizadeh B, Ronaghi M, Shafer RW. 2007. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. Genome Res 17:1195–1201.

Wendl MC, Barbazuk WB. 2005. Extension of Lander-Waterman theory for sequencing filtered DNA libraries. BMC Bioinformatics 6:245.

Wendl MC, Wilson RK. 2008. Aspects of coverage in medical DNA sequencing. BMC Bioinformatics 9:239.

Yeager M, Xiao N, Hayes RB, Bouffard P, Desany B, Burdett L, Orr N, Matthews C, Qi L, Crenshaw A, Markovic Z, Fredrikson KM, Jacobs KB, Amundadottir L, Jarvie TP, Hunter DJ, Hoover R, Thomas G, Harkins TT, Chanock SJ. 2008. Comprehensive resequence analysis of a 136 kb region of human chromosome 8q24 associated with prostate and colon cancers. Hum Genet 124: 161–170.

QIAGEN Exhibit 1002
Page 733

USER GUIDE

Ovation® Target Enrichment System

PART NOS. 9079, 9087, 9099, 9111, 9114, 9115, 9116 AND CUSTOMS

NuGEN
imagine more from less®

Patents, Licensing and Trademarks

© 2014–2016 NuGEN Technologies, Inc. All rights reserved. The Encore®, Ovation® and Applause® families of products and methods of their use are covered by several issued U.S. and International patents and pending applications (www.nugen.com). NuGEN, Ovation, SPIA, Ribo-SPIA, Applause, Encore, Prelude, Mondrian and Imagine More From Less are trademarks or registered trademarks of NuGEN Technologies, Inc. Other marks appearing in these materials are marks of their respective owners.

The purchase of this product conveys to the buyer the limited, non-exclusive, non-transferable right (without the right to modify, reverse engineer, resell, repackage or further sublicense) under these patent applications and any patents issuing from these patent applications to use this product and methods, accompanying this user guide, for research and development purposes solely in accordance with the intended use described and the written instructions provided in this user guide. No license to make or sell products by use of this product is granted to the buyer whether expressly, by implication, by estoppels or otherwise. In particular, the purchase of this product does not include or carry any right or license to use, develop or otherwise exploit this product commercially and no rights are conveyed to the buyer to use the product or components of the product for purposes including commercial services or clinical diagnostics.

For information on purchasing a license to the NuGEN patents for uses other than in conjunction with this product or to use this product for purposes other than research, please contact NuGEN Technologies, Inc., 201 Industrial Road, Suite 310, San Carlos, CA 94070. Phone 888-654-6544 or 650-590-3600; FAX 888-296-6544 or 650-590-3630.

Certain genomic variants, genes, nucleic acid sequences and/or genomic regions and their use in specific applications may be protected by patents. Customers are advised to determine whether licensing is required from the party that owns or controls such patents in order to use the product in Customer's specific application. Unless expressly stated otherwise in writing by NuGEN, Customer, and not NuGEN, is solely responsible for obtaining such licenses.

Warranty

NuGEN warrants that this product meets the performance standards described in the Company's product and technical literature for a period of six months from the date of purchase, provided that the product is handled and stored according to published instructions, and that the product is not altered or misused. If the product fails to meet these performance standards, NuGEN will replace the product free of charge or issue a credit for the purchase price. NuGEN's liability under this warranty shall not exceed the purchase price of the product. NuGEN shall assume no liability for direct, indirect, consequential or incidental damages arising from the use, results of use or inability to use its products. NuGEN reserves the right to change, alter or modify any product to enhance its performance and design.

NuGEN's products are developed, designed and sold FOR RESEARCH USE ONLY. This product is not to be used for diagnostic or therapeutic purposes, nor is it to be administered to humans or animals.

Except as expressly set forth herein, no right to modify, reverse engineer, distribute, offer to sell or sell NuGEN's product is conveyed or implied by buyer's purchase of this NuGEN product. The buyer agrees to use NuGEN products accompanying the product insert in accordance with the intended use and the written instructions provided.

Contents

I. Introduction 1
- A. Background 1
- B. Workflow 1
- C. Performance Specifications 3
- D. Quality Control 3
- E. Storage and Stability 3
- F. Material Safety Data Sheet (MSDS) 3

II. Components 4
- A. Reagents Provided 4
- B. Additional Equipment, Reagents and Labware 8

III. Planning the Experiment 9
- A. Input DNA Requirements 9
- B. Working with the Target Enrichment 32-Plex and 96-Plex Adaptor Plates 9
- C. Library Quantification and Loading Recommendations 9
- D. Sequencing Recommendations and Guidelines 10
- E. Index Structure and Index Read Recommendations 11
- F. Ovation Target Enrichment Data Pre-Processing using BaseSpace 12
- G. Enriched, Amplified Library Storage 14

IV. Overview 15
- A. Overview 15
- B. Protocol Notes 15
- C. Agencourt® RNAClean® XP Purification Beads 16
- D. Programming the Thermal Cycler 17

V. Protocol 19
- A. Preparation of gDNA 19
- B. End Repair 20
- C. FFPE DNA Repair (Optional) 20
- D. Adaptor Ligation 21
- E. Ligation Purification 22
- F. Target Enrichment 23
- G. Enriched Library Purification 24
- H. qPCR (Optional) 26
- I. Library Amplification 26
- J. Amplified Library Purification 27
- K. Quantitative and Qualitative Assessment of the Library 29

VI. Technical Support 30

VII. Appendix 31
A. Normalization of Libraries 31
B. Guidelines for Optional qPCR 32
C. Barcode Sequences and Guidelines for Multiplex Experiments 33
D. Copy Number Alteration/Variation Analysis 35
E. Frequently Asked Questions (FAQs) 39
F. Update History 41

I. Introduction

A. Background

The Ovation Target Enrichment Systems are available as pre-defined panels or complete custom target enrichment solutions for singleplex or multiplex library preparation for targeted resequencing. Powered by Single Primer Enrichment Technology (SPET), the fully customizable probe design enables the interrogation of up to 10 Mb of custom content defined by the user. The system works with as little as 10 ng of fresh/frozen tissue gDNA or higher inputs of DNA isolated from FFPE tissue. The simple, single-tube workflow can be completed in 24 hours with an optional fast protocol that can be completed in as little as 8 hours from sample to target enriched library. A novel barcoding design allows both a high degree of sample multiplexing and the unambiguous identification of true PCR duplicates.

Multiplex kits include reagents for multiplexed target enrichment and Illumina sequencing library creation for 8, 32 or 96 samples. Singleplex sample enrichment versions of the 8-, 32- and 96-sample enrichment kits are also available. A singleplex sample enrichment workflow may be beneficial when working with samples of variable quality such as FFPE blocks or small sample batch sizes.

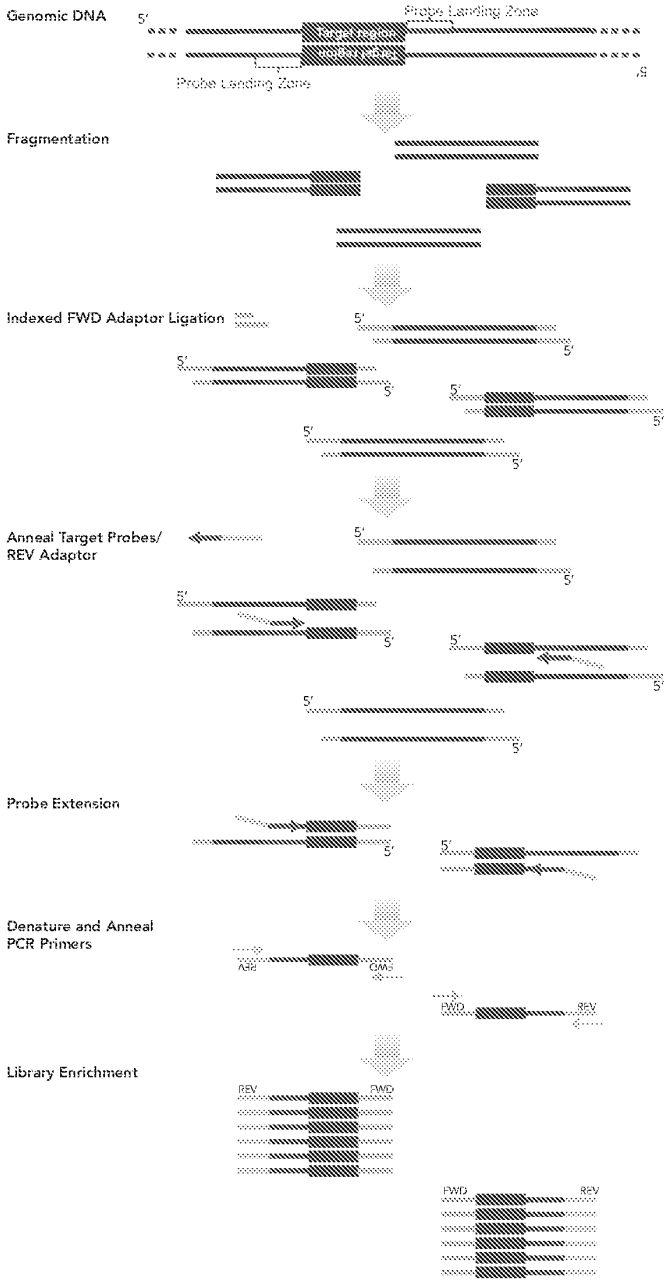
B. Workflow

In the simple, add and incubate workflow shown in Figure 1, fragmented gDNA samples are end-repaired and individually ligated with barcoded adaptors. When starting with variable quality or quantity of gDNA, an optional normalization step can be used to balance the input into the enrichment step. Following a bead purification step, samples are combined for multiplex target enrichment during which time the targeting probes are annealed and extended, creating amplifiable library molecules. Following a second bead purification, an optional qPCR step determines the optimal number of PCR cycles to use in the final library amplification step. After library amplification, a final bead purification prepares the target-enriched library for sequencing. Using this simple workflow, a single user can easily process 32 samples or more at a time.

QIAGEN Exhibit 1002
Page 738

I. Introduction

Figure 1. The Ovation Target Enrichment System DNA workflow.



I. Introduction

C. Performance Specifications

The Ovation Target Enrichment System is designed to produce panel or custom target-enriched, singleplex or multiplex libraries suitable for sequencing on Illumina NGS platforms. It is a fast, robust and simple system capable of producing target-enriched libraries from 10–500 ng of genomic DNA from fresh or FFPE tissue samples. When starting with FFPE material, the maximum input available should be used to produce best results. Sequencing metrics such as on-target rates and uniformity of coverage will likely vary from one custom probe design to another.

D. Quality Control

Each lot of library reagents in the Ovation Target Enrichment System undergoes functional testing to meet strict performance specifications.

E. Storage and Stability

The Ovation Target Enrichment System is shipped on dry ice and should be unpacked immediately upon receipt.

Note: This product contains components with multiple storage temperatures.

- Vials labeled Agencourt® RNAClean® XP Beads (clear cap) should be removed from the top of the shipping carton upon delivery and stored at 4°C.
- All other components should be stored at –20°C on internal shelves of a freezer without a defrost cycle.

The kit has been tested to perform to specifications after as many as six freeze/thaw cycles. Kits handles and stored according to the above guidelines will perform to specifications for at least six months.

F. Material Safety Data Sheet (MSDS)

If applicable, an MSDS for this product is available on the NuGEN website at www.nugen.com/products/ngs/ovation-target-enrichment-system.

QIAGEN Exhibit 1002
Page 740

II. Components

A. Reagents Provided

Table 1. Ovation Target Enrichment System Target Enrichment System Reagents (Multiplex Kits) (Part Nos. 0400-08, 0400-32, 0400-96)

COMPONENTS	0400-08 PART NUMBER	0400-32 PART NUMBER	0400-96 PART NUMBER	VIAL CAP	VIAL NUMBER
End Repair Buffer Mix	S01860	S01899	S01907	Blue	ER1 VER 9
End Repair Enzyme Mix	S01510	S01533	S01909	Blue	ER2 VER 4
End Repair Enhancer	S01862	S01709	S01908	Blue	ER3 VER 2
End Repair FFPE Enhancer	S01938	S01935	S01939	Blue	ER5
Ligation Buffer Mix	S01936	S01534	S01910	Yellow	L1 VER 4
Ligation Enzyme Mix	S01467	S01535	S01911	Yellow	L3 VER 4
DNA Resuspension Buffer Mix	S01901	S01901 (2)	S01901 (3)	Clear	DR1
Target Extension Buffer Mix	S01866	S01900	S01912	Clear	TX1
Target Extension Enzyme Mix	S01870	S01902	S01913	Clear	TX2
Amplification Buffer Mix	S01871	S01903	S01914	Red	P1 VER 4
Amplification Primer Mix	S01934	S01906	S01917	Red	P2 VER 14
Amplification Enzyme Mix	S01873	S01764	S01895	Red	P3 VER 2
Amplification Enhancer Mix	S01872	S01904	S01915	Red	P6
Nuclease Free Water	S01001	S01001, S01113	S01001 (2), S01113 (3)	Green	D1
RNACleanXP Beads	S01502	S01502	S01502 (3)	Clear	----

QIAGEN Exhibit 1002
Page 741

II. Components

Ovation Target Enrichment System Target Enrichment System Reagents (Multiplex Kits) (Part Nos. 0400-08, 0400-32, and 0400-96), *continued*					
COMPONENTS	0400-08 PART NUMBER	0400-32 PART NUMBER	0400-96 PART NUMBER	VIAL CAP	VIAL NUMBER
Ligation Adaptor Mix	S01941 S01942 S01943 S01944 S01945 S01946 S01947 S01948	N/A	N/A	Yellow	L2V15-BC1 L2V15-BC2 L2V15-BC3 L2V15-BC4 L2V15-BC5 L2V15-BC6 L2V15-BC7 L2V15-BC8
Target Enrichment 32-Plex Adaptor Plate	N/A	S01937	N/A	N/A	N/A
Target Enrichment 96-Plex Adaptor Plate	N/A	N/A	S01940	N/A	N/A

Table 2. Ovation Target Enrichment Reagents for Singleplex Enrichment (Part Nos. 0404-08, 0404-32 and 0404-96)

COMPONENTS	0404-08 PART NUMBER	0404-32 PART NUMBER	0404-96 PART NUMBER	VIAL CAP	VIAL NUMBER
End Repair Buffer Mix	S01860	S01899	S01899 (3)	Blue	ER1 ver 9
End Repair Enzyme Mix	S01510	S01533	S01533 (3)	Blue	ER2 ver 4
End Repair Enhancer	S01862	S01709	S01709 (3)	Blue	ER3 ver 2
End Repair FFPE Enhancer	S01938	S01935	S01935 (3)	Blue	ER5
Ligation Buffer Mix	S01936	S01534	S01534 (3)	Yellow	L1 ver 4
Ligation Enzyme Mix	S01467	S01535	S01535 (3)	Yellow	L3 ver 4
DNA Resuspension Buffer Mix	S01901 (2)	S01901 (2)	S01901 (6)	Clear	DR1
Target Extension Buffer Mix	S01866	S01900	S01900 (3)	Clear	TX1

QIAGEN Exhibit 1002
Page 742

II. Components

Ovation Target Enrichment Reagents for Singleplex Enrichment (Part Nos. 0404-08, 0404-32, and 0404-96), *continued*					
COMPONENTS	**0404-08 PART NUMBER**	**0404-32 PART NUMBER**	**0404-96 PART NUMBER**	**VIAL CAP**	**VIAL NUMBER**
Target Extension Enzyme Mix	S01870	S01902	S01902 (3)	Clear	TX2
Amplification Buffer Mix	S01903	S01914	S01914 (3)	Red	P1 ver 4
Amplification Primer Mix	S01906	S01917	S01917 (3)	Red	P2 ver 14
Amplification Enzyme Mix	S01764	S01895	S01895 (3)	Red	P3 ver 2
Amplification Enhancer Mix	S01904	S01915	S01915 (3)	Red	P6
Nuclease Free Water	N/A	S01001	N/A	Green	D1
Nuclease Free Water	S01113	S01113	S01113 (5)	Green	D1
Ligation Adaptor Mixes	S01941 S01942 S01943 S01944 S01945 S01946 S01947 S01948	N/A	N/A	Yellow	L2V15-BC1 L2V15-BC2 L2V15-BC3 L2V15-BC4 L2V15-BC5 L2V15-BC6 L2V15-BC7 L2V15-BC8
Target Enrichment 32-Plex Adaptor Plate	N/A	S01937	N/A	N/A	N/A
Target Enrichment 96-Plex Adaptor Plate	N/A	N/A	S01940	N/A	N/A

Note: Each Ovation Target Enrichment System kit also includes a solution of target probes specific to the kit ordered. For details, please refer to Table 3.

QIAGEN Exhibit 1002
Page 743

II. Components

Table 3. Target Enrichment Probes (Part Nos. 9079-xx, 9087-xx, 9099-xx, 9111-xx, 9114-xx, 9115-xx, 9116-xx)

COMPONENT	PART NO. 9079-08	PART NO. 9079-32	PART NO. 9079-96
Ovation Cancer Panel Target Enrichment System Target Enrichment Probe	S02007	S02008	S02008 (3)
COMPONENT	**PART NO. 9087-08**	**PART NO. 9087-32**	**PART NO. 9087-96**
Ovation Human Mitochondrion Target Enrichment Probe	S02031	S02032	S02032 (3)
COMPONENT	**PART NO. 9099-08**	**PART NO. 9099-32**	**PART NO. 9099-96**
Ovation Cancer Panel FFPE Target Enrichment Probe	S02185	S02175	S02175 (3)
COMPONENT	**PART NO. 9111-08**	**PART NO. 9111-32**	**PART NO. 9111-96**
Ovation Cancer Panel FFPE Target Enrichment Probe Singleplex	S02176	S02162	S02162 (3)
COMPONENT	—	**PART NO. 9114-32**	**PART NO. 9114-96**
Ovation Cancer Panel Target Enrichment Probe Singleplex	—	S02171	S02171 (3)
COMPONENT	**PART NO. 9115-08**	**PART NO. 9115-32**	**PART NO. 9115-96**
Ovation Cancer Panel 2.0 Target Enrichment Probe	S02173	S02174	S02174 (3)
COMPONENT	**PART NO. 9116-08**	**PART NO. 9116-32**	**PART NO. 9116-96**
Ovation Cancer Panel 2.0 Target Enrichment Probe Singleplex	S02185	S02175	S02175 (3)

QIAGEN Exhibit 1002
Page 744

II. Components

B. Additional Equipment, Reagents and Labware

Required Materials

- **Equipment**
  - Covaris Sonication System
  - Agilent 2100 Bioanalyzer or materials and equipment for electrophoretic analysis of nucleic acids
  - Real-time PCR system capable of SYBR Green detection
  - Microcentrifuge for individual 1.5 mL and 0.5 mL tubes
  - Microcentrifuge for individual or strip 0.2 mL tubes
  - Vortexer
  - Thermal cycler with 0.2 mL tube heat block, heated lid, and 100 µL reaction capacity
  - Appropriate spectrophotometer and cuvettes, or Nanodrop® UV-Vis Spectrophotometer for quantitation of fragmented DNA
- **Reagents**
  - Ethanol (Sigma-Aldrich, Cat. #E7023), for purification steps
  - EvaGreen Dye (Biotium, Cat. #31000) or SYBR Green, for optional qPCR and/or normalization protocols
  - MyTaq HS DNA Polymerase (Bioline, Cat. #BIO-21111), for optional normalization protocol
- **Supplies and Labware**
  - Amicon Ultra 30K MWCO (Millipore, Cat. #UFC503024) or Vacuum Concentrator
  - Nuclease-free pipette tips
  - 1.5 mL and 0.5 mL RNase-free microcentrifuge tubes
  - 0.2 mL individual thin-wall PCR tubes or 8 X 0.2 mL strip PCR tubes or 0.2 mL thin-wall PCR plates
  - Magnetic bead separation device
  - Disposable gloves
  - Kimwipes
  - Ice bucket
  - Cleaning solutions such as DNA-OFF™ (MP Biomedicals, Cat. #QD0500)

To Order

- Agilent, Inc., www.agilent.com
- Biotium, www.biotium.com
- Covaris, www.covarisinc.com
- Millipore, www.millipore.com
- MP Biomedicals, www.mpbio.com
- Sigma-Aldrich, Inc., www.sigmaaldrich.com

QIAGEN Exhibit 1002
Page 745

III. Planning the Experiment

A. Input DNA Requirements

The Ovation Target Enrichment System is designed to work with inputs of 10 ng to 500 ng of fragmented genomic DNA. DNA samples must be free of contaminating proteins, RNA, organic solvents (including phenol and ethanol) and salts. Use of a commercially available system for genomic DNA isolation is recommended. The A260:A280 ratio for DNA samples should be between 1.8–2.0. Use of DNA samples with lower ratios may result in low amplification yield. Genomic DNA derived from FFPE samples has also been demonstrated to work well in the Ovation Target Enrichment System but may require higher inputs.

B. Working with the Target Enrichment 32-Plex and 96-Plex Adaptor Plates

The Target Enrichment Adaptor Plates contain ligation adaptor mixes, each with a unique eight-base barcode. Each well (first 32 wells (A01–H04) or all 96 wells, respectively) contains sufficient volume for preparation of a single library. The Target Enrichment Adaptor Plates are sealed with a foil seal designed to provide for airtight storage.

Prior to thawing the adaptor plate, spin it for 5 to 10 minutes at room temperature at ~1000 X g in a centrifuge with a rotor appropriate for microwell plates. This will allow the plate to thaw while spinning and help to minimize cross-contamination between wells. To use only a portion of the plate, remove the seal covering only the portion of the plate you wish to use. The remaining wells of the plate should remain sealed for use at a later date. Used wells can be covered with a new foil seal (e.g., AlumaSeal II) to prevent any remaining adaptor-containing liquid from contaminating future reactions.

For details regarding barcode balancing for multiplex sequencing, please see Appendix C on page 33.

C. Library Quantification and Loading Recommendations

We recommend using a qPCR quantification method for the most accurate quantification of libraries produced with the Ovation Target Enrichment System. These libraries may also produce a broad profile when analyzed using the Agilent Bioanalyzer depending on the input DNA quality and size. In order to avoid underestimating the library concentration in this case, we recommend using a library fragment size approximately 50–100 bp smaller than the actual peak determined by the Agilent Bioanalyzer to calculate the size adjusted concentration. Additionally, we recommend loading approximately 60% of the Illumina recommend maximum input onto the flowcell during initial sequencing. The amount of library loaded can be adjusted for subsequent sequencing runs in order to optimize the cluster density.

QIAGEN Exhibit 1002
Page 746

D. Sequencing Recommendations and Guidelines

We recommend Illumina single end sequencing with a 14-base Indexing read (see Section E. below) for most gDNA applications using the Ovation Target Enrichment System. Ensure that the SBS reagent kit chosen contains enough reagent to also allow for a 14-base index read.

Since the randomly fragmented end of the insert is directly ligated to the forward adaptor, high quality targeted data can be obtained from as little as 50-base single end reads. Longer reads will generally improve target coverage, and 100–150 base pair forward read lengths are good options for these libraries. It may occasionally be beneficial or necessary to perform paired end sequencing on these libraries. For example, to detect chromosomal translocation or fusion events it is useful to perform paired end sequencing in order to ensure that the translocation break point can be identified.

Sequencing recommendations:

For SNP studies:

Index read: 14 bp (8 nt barcode + 6 nt random sequence).

Forward read length for high quality DNA: 100–150 bp.

Forward read length for FFPE DNA: based on the final library size.

Reverse read: Not required. Minimum of 100 bp if performed.

For genomic rearrangement studies:

Index read: 14 bp (8 nt barcode + 6 nt random sequence) or 8 bp if duplicate determination not required.

Forward read length for high quality DNA: 100–150 bp.

Forward read length for FFPE DNA: based on the final library size.

Reverse read: 75–100 bp.

When performing paired end sequencing, the initial 68 bases of every reverse read is derived from the probe and an 18 bp linker sequence. A reverse read of at least 75–100 bp is required to sequence through the linker and probe sequence and allow unambiguous identification of the probe hybridization site. While the probe sequence must be trimmed for variant analysis, it can provide useful information for chromosomal translocation and fusion events. The extent of trimming from the reverse read will depend on the experimental application.

QIAGEN Exhibit 1002
Page 747

E. Index Structure and Index Read Recommendations

The Ovation Target Enrichment System uses 8-base barcodes for sample multiplexing. In addition to the 8-base barcode, the index contains 6 random bases immediately following the 8-base barcode, for a total of 14 bases. The additional 6 bases are used for duplicate read determination using the NuGEN Duplicate Marking Tool, NuDup (http://nugentechnologies.github.io/nudup/). The libraries should be sequenced using 14 cycles for the index read. The 8-base barcode sequences can be found in this user guide. The 8-base barcodes followed by NNNNNN should be entered into the sample sheet to enable proper multiplex library parsing.

Note: If the data will be pre-processed using the NuGEN Ovation Target Enrichment App for BaseSpace (see Section III.F for more details), demultiplexing the samples is not recommended. To avoid automatic demultiplexing, substitute the barcode information in the sample sheet with 14 N's. It is also necessary to generate and upload the 14-bp Index fastq file (see below) and the target manifest file (from NuGEN Technical Support) to BaseSpace as input for the app.

The MiSeq and HiSeq instruments do not provide a simple way to obtain the sequence information contained in the 14-base pair index read including the 6 random bases that are necessary for duplicate read determination. Several recommended methods to generate the necessary index fastq file are provided below.

• MiSeq Instrument

Parsing multiplex runs using the MiSeq built-in Illumina software replaces the barcode sequence from each library with a numerical substitute. This removes the duplicate information provided by the N6 sequence present after the barcode. To retrieve this information using the MiSeq instrument, we recommend modification of the MiSeq config file to allow generation of an index fastq file during data analysis. This will generate a 14-base index file that is compatible with the NuGEN Target Enrichment BaseSpace App or use of the NuDup script.

If you are unfamiliar with editing the config file, we recommend you request assistance from Illumina Technical Support to make this modification. The steps are as follows:

1. Stop the MiSeq Reporter process.
2. Locate the "MiSeq Reporter.exe.config" file located in C:/Illumina/MiSeq Reporter
3. Open config file and search for a line that reads:
   "<add key="CreateFastqForIndexReads" value="0"/>".
   - If this line is present, change the value from "0" to "1".
   - If this line is not present, add the line to the config file using the add keys function under the appSettings tab.
4. Restart the MiSeq reporter process.
5. Requeue the run for data analysis if required.

QIAGEN Exhibit 1002
Page 748

• Other Illumina Sequencers

Use the method described below to generate the read and index fastq files for use with the NuGEN Ovation Target Enrichment App for BaseSpace. This workflow uses bcl2fastq2 version 2.17.0 (formerly CASAVA).

1. Maneuver to the location of the sequence run folder (referred to as RunFolder in this document) and rename the sample sheet (e.g., SampleSheet.csv.bak).

   Note: In order to parse the data during the fastq file generation, modify the sample sheet to remove the six N's located at the end of the barcodes. Run bcl2fastq using the command:

   "/usr/local/bin/bcl2fastq --runfolder-dir . --output-dir ./Data/Intensities/BaseCalls/ --sample-sheet SampleSheet.csv --use-bases-mask y*,i8y* --minimum-trimmed-read-length 0" (for paired-end reads use "--use-bases-mask y*,i8y*, y*).

   This command will produce an R1 fastq file with the forward read, an R2 fastq files containing just the N6 information and, if present, an R3 fastq file containing the reverse read.

2. To generate the read and index fastq files, use the following command:

   "usr/local/bin/bcl2fastq --runfolder-dir . --output-dir ./Data/Intensities/BaseCalls/ --use-bases-mask y*,y*--minimum-trimmed-read-length 0 (for paired-end reads use --use-bases-mask y*,y*,y*.

3. The fastq files will be located in /RunFolder/Data/Intensities/BaseCalls/ unless otherwise specified. The generated fastq files can then be uploaded to BaseSpace for input into the NuGEN Ovation Target Enrichment App.

F. Ovation Target Enrichment Data Pre-Processing using BaseSpace

The NuGEN Ovation Target Enrichment BaseSpace Application (App) is designed for initial data processing of sequencing data produced from libraries prepared with NuGEN's Ovation DNA Target Enrichment System family of products. The App is suitable for use with either single or paired end read data from any Illumina sequencing platform. Data input requirements for this App are human DNA sequence data in FASTQ format (including the index read) as well as the NuGEN provided manifest file containing probe and target information. The App initially demultiplexes the data, using the 96 barcodes provided in the NuGEN Ovation Target Enrichment kits. An optional quality and/or probe trimming can be performed on the parsed data to remove data derived from bases with low sequencing quality scores or from probe sequences. The reads are then aligned to the UCSC human reference genome hg19 using Bowtie2. Subsequently, NuGEN's custom data processing tool, NuDup is used to identify PCR duplicate reads based on the use of the random N6 sequence positioned adjacent to the barcode index in the forward adaptor. Multiple output files are available with NuDup-identified duplicates either removed or alternately, retained and marked. By marking duplicates, downstream analyses, such as those for variant (SNP) detection,

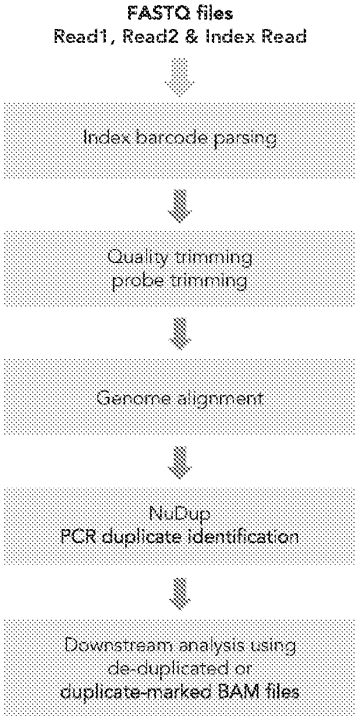
QIAGEN Exhibit 1002
Page 749

can take advantage of duplicate read resequencing. Processed BAM files with PCR duplicates either removed or marked, processed FASTQ files with duplicates removed, and UCSC bigWig tracks with duplicates removed for visualization using the UCSC human genome browser are provided by the App. For each library data set processed, the App provides a summary table of Picard generated sequencing metrics detailing the total number of reads, percent aligned to the genome, percent on-target, percent of targets with zero coverage, mean coverage, and percent duplicates.

The NuGEN Ovation Target Enrichment Application (App) is available on the Illumina BaseSpace website (https://basespace.illumina.com). When planning to use this App for initial data pre-processing, ensure that the sequencing center is able to provide the un-demultiplexed read and index files for upload to BaseSpace. Additional information on file requirements, file uploads, and App usage can be found in the NuGEN Ovation Target Enrichment BaseSpace App user guide (www.nugen.com).

Alternatively, the standalone NuDup script for duplicate removal or marking can be downloaded from GitHub (http://nugentechnologies.github.io/nudup/) for use in a local data analysis pipeline. A general data pre-processing workflow is provided below.

Figure 2. General data processing workflow for the Ovation Target Enrichment System DNA.



Contact NuGEN Technical Support for additional guidance or questions regarding processing of Ovation Target Enrichment generated data.

G. Enriched, Amplified Library Storage

Amplified libraries may be stored at –20°C.

IV. Overview

A. Overview

The library preparation process used in the Ovation Target Enrichment System is performed in four stages:

1. DNA end repair	0.75 hours
2. Adaptor ligation and purification	1.75 hours
3. Probe hybridization and extension	3–12+ hours*
4. Amplification and purification	2.25 hours
Total time to prepare enriched library	**8–20 hours**

*Short protocol ~3 hours

Components in the Ovation Target Enrichment System are color coded, with each color linked to a specific stage of the process. Performing each stage requires making a master mix then adding it to the reaction, followed by incubation. Master mixes are prepared by mixing components provided for that stage.

B. Protocol Notes

- We recommend the routine use of a positive control DNA. Especially the first time you set up a reaction, the use of a positive control DNA will allow the establishment of a baseline of performance and provide the opportunity to become familiar with the protocol steps.
- Use the water provided with the kit (green: D1) or an alternate source of nuclease-free water. We do not recommend the use of DEPC-treated water with this protocol.
- Thaw components used in each step and immediately place them on ice. Do not thaw all reagents at once.
- Always keep thawed reagents and reaction tubes on ice unless otherwise instructed.
- After thawing and mixing buffer mixes, if any precipitate is observed, re-dissolve it completely prior to use. You may gently warm the buffer mix for 2 minutes at room temperature followed by brief vortexing. Do not warm any enzyme mixes.
- When placing small amounts of reagents into the reaction mix, pipet up and down several times to ensure complete transfer.
- When instructed to pipet mix, gently aspirate and dispense a volume that is at least half of the total volume of the reaction mix.
- Always allow the thermal cycler to reach the initial incubation temperature prior to placing the tubes or plates in the block.
- When preparing master mixes, use the minimal amount of extra material to ensure that you are able to run the maximal number of reactions using the components provided in the kit.
- Components and reagents from other NuGEN kits should not be used with the Ovation Target Enrichment Systems.

QIAGEN Exhibit 1002
Page 752

- Make the ethanol mixes fresh, carefully measuring both the ethanol and water with pipettes.
- Use only fresh ethanol stocks to make 85% ethanol used in the purification steps.

C. Agencourt® RNAClean® XP Purification Beads

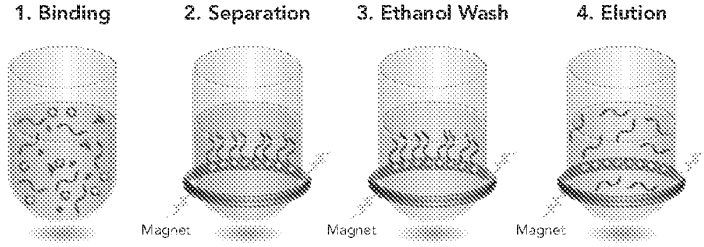
Tips and Notes

There are significant modifications to the Agencourt RNAClean XP beads standard procedure; therefore, it is important to follow the protocols outlined in this user guide for the use of these beads. However, the Beckman Coulter user guide may be reviewed to become familiar with the manufacturer's recommendations.

The bead purification processes used in this kit consist of the following steps:

1. Binding of DNA to RNAClean XP beads
2. Magnetic separation of beads from supernatant
3. Ethanol wash of bound beads to remove contaminants
4. Elution of bound DNA from beads

Figure 3. Agencourt® RNAClean® XP Bead purification process overview.



Reproduced from original picture from Agencourt/Beckman Coulter Genomics

Additional Tips and Notes

- Remove beads from 4°C and leave at room temperature for at least 30 minutes before use, and ensure that they have completely reached room temperature. Cold beads reduce recovery.
- Fully resuspend beads by inverting and tapping before adding to sample.
- It is critical to let the beads separate on the magnet for a full 5 minutes. Removing binding buffer before the beads have completely separated will impact DNA yields.
- After completing the binding step, it is important to minimize bead loss when removing the binding buffer.

QIAGEN Exhibit 1002
Page 753

- Any significant loss of beads during the ethanol washes will impact DNA yields, so make certain to minimize bead loss throughout the procedure.
- Ensure that the ethanol wash is freshly prepared from fresh ethanol stocks at the indicated concentration. Lower percent ethanol mixes will reduce recovery.
- During the ethanol washes, keep the samples on the magnet. The beads should not be allowed to disperse; the magnet will keep the beads on the walls of sample wells or tubes in a small ring. It is critical that all residual ethanol be removed prior to continuing with the next step. Therefore, when removing the final ethanol wash, first remove most of the ethanol, then allow the excess to collect at the bottom of the tube before removing the remaining ethanol. This reduces the required bead air drying time.
- After drying the beads for the time specified in the protocol, inspect each tube carefully and make certain that all the ethanol has evaporated before proceeding.
- It is strongly recommended that strip tubes or partial plates are firmly placed when used with the magnetic plate. We do not advise the use of individual tubes as they are difficult to position stably on the magnetic plates.

D. Programming the Thermal Cycler

Use a thermal cycler with a heat block designed for PCR plates, equipped with a heated lid, and with a capacity of 100 µL reaction volume. Prepare the programs shown in Table 4 following the operating instructions provided by the manufacturer. For thermal cyclers with an adjustable heated lid, set the lid temperature to 100°C only when sample temperature reaches above 30°C. For thermal cyclers with a fixed temperature heated lid (e.g., ABI GeneAmp® PCR 9600 and 9700 models), use the default settings (typically 100 to 105°C).

QIAGEN Exhibit 1002
Page 754

Table 4. Thermal Cycler Programming

END REPAIR	
Program 1 End Repair/Ligation	25°C – 30 min, 70°C – 10 min, hold at 4°C
DNA REPAIR (OPTIONAL)	
Program 2 FFPE DNA Repair (optional)	16°C – 10 min, 70°C – 10 min, hold at 4°C
HYBRIDIZATION	
Program 3 Probe Hybridization & Extension	95°C – 5 min, 200 cycles (80°C – 10 sec, decrease 0.1°C each cycle), hold at 60°C, 72°C – 10 min, hold at 4°C
AMPLIFICATION	
Program 4 Library Amplification	37°C – 10 min, 95°C – 3 min, 15–30 cycles* (95°C – 30 sec, 62°C – 15 sec, 72°C – 20 sec), 72°C – 2 min, hold at 4°C

*The number of cycles may be adjusted and optimized for desired yield based on qPCR assessment of the unamplified library.

QIAGEN Exhibit 1002
Page 755

V. Protocol

Important Note: The 8-sample kits are designed for multiplex enrichment of 4 samples at a time. Processing fewer than 4 samples will result in insufficient Targeting Probe Pool and PCR reagents to complete all reactions.

The 32-sample and 96-sample enrichment kits are designed for multiplex enrichment of 8 samples at a time. Processing fewer than 8 samples will result in insufficient Targeting Probe Pool and PCR reagents to complete all reactions.

Singleplex sample enrichment versions of the 8-, 32- and 96-sample enrichment kits are also available. A singleplex sample enrichment workflow may be beneficial when working with samples of variable quality such as from FFPE blocks or small sample batch sizes.

Use a heated lid in all steps to prevent evaporation. Use PCR tubes or plates that seal tightly and will not deform when the thermal cycler lid is closed. Insufficient tube sealing will result in sample loss due to evaporation during extended incubation steps.

Use a heated lid in all steps to prevent evaporation. Use PCR tubes or plates that seal tightly and will not deform when the thermal cycler lid ***is closed. Insufficient tube*** *sealing will result in sample loss due to evaporation during extended incubation steps.*

A. Preparation of gDNA

Genomic DNA used in the target enrichment protocol should be fragmented to an average of 500 bp in size. We recommend using Covaris acoustic shearing following the manufacturer's protocols. The gDNA may need to be concentrated after fragmentation as the protocol requires the gDNA to be in a volume of 10 µL. You may reduce the volume of the gDNA by using a vacuum concentrator (Speedvac) according to the manufacturer's instructions.

Alternatively, you may use Agencourt AMPure XP or RNAClean XP beads to purify and concentrate the fragmented DNA. However, since some loss of material will occur during bead purification, you should repeat the quantitation of the samples prior to End Repair.

Important Note: Samples intended to be multiplexed together at Target Enrichment (step F) must be of equivalent mass, quality and average fragment length. Alternatively, when using variable gDNA samples, use the optional Normalization of Libraries protocol (Appendix A) to determine the quantity of library ready material to balancing inputs into the Target Enrichment reaction. Failure to normalize samples that are processed together will result in unequal sample representation in the final target-enriched library.

QIAGEN Exhibit 1002
Page 756

V. Protocol

B. End Repair

1. Obtain the End Repair Buffer Mix (blue: ER1), End Repair Enzyme Mix (blue: ER2) and End Repair Enhancer (blue: ER3) from –20°C storage.
2. Thaw ER1 at room temperature. Mix by vortexing, spin and place on ice.
3. Spin down contents of ER2 and ER3 and place on ice.
4. Prepare a master mix by combining ER1, ER2 and ER3 in a 0.5 mL capped tube, according to the volumes shown in Table 5.

Table 5. End Repair Master Mix (volumes listed are for a single reaction)

END REPAIR BUFFER MIX (BLUE: ER1 VER 9)	END REPAIR ENZYME MIX (BLUE: ER2 VER 4)	END REPAIR ENHANCER (BLUE: ER3 VER 2)
4 µL	0.5 µL	0.5 µL

*Mix by pipetting and spin down the master mix **briefly. Place on ice. Use** immediately.*

5. Add 5 µL of the End Repair Master Mix to 10 µL of fragmented DNA (10 ng to 500 ng).
6. Mix by pipetting, cap and spin tubes and place on ice.
7. Place the tubes in a pre-warmed thermal cycler programmed to run Program 1 (End Repair/Ligation; see Table 4):

   25°C – 30 min, 70°C – 10 min, hold at 4°C
8. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice.
9. Proceed to FFPE DNA Repair (optional, Section D) or continue immediately with the Adaptor Ligation protocol (Section E).

C. FFPE DNA Repair (Optional)

This is an optional repair step for use with damaged DNA, e.g., FFPE DNA.

1. Add 1 µL of End Repair FFPE Enhancer (blue: ER5).
2. Mix well by pipetting.
3. Place the tubes in a pre-warmed thermal cycler programmed to run Program 2 (FFPE DNA Repair; see Table 4):

   16°C – 10 min, 70°C – 10 min, hold at 4°C
4. Remove the tubes from the thermal cycler, spin briefly and place on ice.

QIAGEN Exhibit 1002
Page 757

D. Adaptor Ligation

1. Remove the Ligation Buffer Mix (yellow: L1), Ligation Enzyme Mix (yellow: L3) and Nuclease-free Water (green: D1) from –20°C storage.

   For 8-reaction kits: Remove Ligation Adaptor Mixes (yellow: L2V15-BC1 to BC8) from –20°C storage.

   For 32- and 96-reaction kits: Remove Target Enrichment Adaptor Plate (32-plex or 96-plex) from –20°C storage.

2. Thaw L1, Nuclease-free Water and Ligation Adaptor Mixes at room temperature. Mix L1 by vortexing, spin and place on ice.

3. **For 8-reaction kits:** Spin down Ligation Adaptor Mixes and place on ice.

   For 32- and 96-reaction kits: Spin down Target Enrichment Adaptor Plate and place on ice.

4. Spin down L3 and place on ice.

5. **Add 3 µL of Ligation Adaptor Mix directly to each end-repaired sample.** Each tube of Ligation Adaptor Mix and each well of Target Enrichment Adaptor Plates contains adaptors with a unique barcode.

6. Make a master mix by combining the Nuclease-free Water (green: D1), L1 and L3 in a 0.5 mL capped tube, according to the volumes shown in Table 6. Prepare the master mix by combining water and Ligation Buffer Mix, vortex to mix, spin down, then add Ligation Enzyme Mix. Mix thoroughly by repeatedly pipetting up **and down using a large volume (100–200 µL) pipettor — the larger bore tips used with such pipettors are necessary to sufficiently mix the Ligation Master Mix and** sample.

   Note: The L1 Ligation Buffer Mix is very viscous. Please be sure to pipet this reagent slowly.

Table 6. Ligation Master Mix (volumes listed are for a single reaction)

WATER (GREEN: D1)	LIGATION BUFFER MIX (YELLOW: L1 VER 4)	LIGATION ENZYME MIX (YELLOW: L3 VER 4)
4.5 µL	6.0 µL	1.5 µL

*Mix by pipetting and spin down the master mix **briefly. Place on ice. Use** immediately.*

7. **Add 12 µL Ligation Master Mix to each tube of sample/adaptor. Mix thoroughly by repeatedly pipetting up and down using a large volume (100–200 µL) pipettor —** the larger bore tips used with such pipettors are necessary to sufficiently mix the Ligation Master Mix and sample.

QIAGEN Exhibit 1002
Page 758

8. Place the tubes in a pre-warmed thermal cycler programmed to run Program 1 (End Repair/Ligation; see Table 4):

   25°C – 30 min, 70°C – 10 min, hold at 4°C

9. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice.

10. Continue immediately with the Ligation Purification protocol.

E. Ligation Purification

1. Retrieve the Agencourt RNAClean XP Beads from 4°C and ensure they are at room temperature. Also remove the DNA Resuspension Buffer (DR1) from –20°C and thaw at room temperature.

2. Resuspend the beads by inverting and tapping the tube. Ensure the beads are fully resuspended before adding to the sample. After resuspending, do not spin the beads. (An excess of beads is provided; therefore, it is not necessary to recover any trapped in the cap.)

3. Add 70 µL of room-temperature Nuclease-free Water (D1) to each ligation reaction.

4. Add 80 µL (0.8 volumes) room temperature bead suspension to each sample. Mix thoroughly by pipetting 10 times.

5. Incubate at room temperature for 5 minutes.

6. Transfer the tubes to the magnet plate and let stand 5 minutes to completely clear the solution of beads.

7. Carefully remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step.

   Note: The beads should not disperse; instead, they will stay on the walls of the tubes. Significant loss of beads at this stage will impact the amount of DNA carried into Target Enrichment, so ensure beads are not removed with the binding buffer or the wash.

8. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute.

9. Remove the ethanol wash using a pipette.

10. Air dry the beads on the magnet for 10 minutes. Inspect each tube carefully to ensure that all the ethanol has evaporated. It is critical that all residual ethanol be removed prior to continuing.

11. Remove the tubes from the magnet.

QIAGEN Exhibit 1002
Page 759

12. Add 21 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature.

    Note: For the singleplex sample enrichment kit, add 33 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature.

13. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads.

14. Carefully remove 20 µL of the eluate, ensuring as few beads as possible are carried over.

    Note: For the singleplex sample enrichment kit, carefully remove 32 µL of the eluate, ensuring as few beads as possible are carried over. The samples can then be stored at –20°C or used immediately in the Target Enrichment step (Section F).

15. Combine the eluted DNA of 8 samples (combine 4 samples when using 0400-08 kit) into a single fresh 1.5 mL tube.

16. Concentrate the DNA by either using an Amicon 0.5 mL 30k MWCO spin column (Millipore UFC503024) following the manufacturers' recommendations or by vacuum concentration. If final volume falls below 32 µL after the concentration step, add water to bring the volume up to 32 µL.

17. At this point, samples can be stored at –20°C until needed or can be used immediately in the Target Enrichment step.

F. Target Enrichment

1. Remove the Target Extension Buffer Mix (clear: TX1) and Target Probe Pool (various Part Nos.) from –20°C storage.

2. Thaw TX1 and Target Probe Pool at room temperature. Mix by vortexing, spin and place on ice.

3. Prepare a master mix by combining TX1 and Target Probe Pool in a 0.5 mL capped tube, according to the volumes shown in Table 7.

Table 7. Target Enrichment Mix (volumes listed are for a single reaction)

TARGET EXTENSION BUFFER MIX (CLEAR: TX1)	TARGET PROBE POOL (CLEAR: N/A)
7 µL	10 µL

Mix by pipetting and spin down the master mix briefly. Place on ice. Use immediately.

4. Add 17 µL of Target Enrichment Master Mix to 32 µL of pooled, ligated DNA. Mix by vortexing. Spin down briefly.

QIAGEN Exhibit 1002
Page 760

*Use a heated lid in all steps to prevent evaporation. Use PCR tubes or plates that seal tightly and will not deform when the thermal cycler lid **is closed. Insufficient tube** sealing or tube deformation will result in sample loss due to evaporation during extended incubation steps.*

5. Place the tubes in a pre-warmed thermal cycler programmed to run Program 3 (Probe Hybridization and Extension; see Table 4). For alternative *Quick Hybridization and Extension* program, see below):

   95°C – 5 min, 200 cycles (80°C – 10 sec, decrease 0.1°C each cycle), hold at 60°C*, 72°C – 10 min, hold at 4°C

   *After 12–16 hours incubation at 60°C, open lid of thermal cycler, open tube caps and, **without removing the tubes from the thermal cycler**, add 1 µL of Target Extension Enzyme Mix (clear: TX2). Pipette mix up and down >5 times **with a volume equal to or greater than 40 µL. Reseal the tubes, close the thermal** cycler lid and immediately continue to the next step of the program (72°C for 10 min, hold at 4°C).

Alternative Quick Hybridization and Extension: You may instead choose to run the quick protocol to complete the target enrichment and library **preparation in a single day and fit more closely within your sample processing** pipeline. We recommend using a larger amount of input gDNA/ds-cDNA for the Quick Protocol. Lower input amounts will likely exhibit lower on-target reads and elevated duplication rates.

5. Place the tubes in a pre-warmed thermal cycler programmed to run Program 3 (Probe Hybridization and Extension; see Table 4):

   95°C – 5 min, 200 cycles (80°C – 10 sec, decrease 0.1°C each cycle), hold at 60°C*, 72°C – 10 min, hold at 4°C

   *After a minimum of 2 hours at 60°C, open lid of thermal cycler, open tube caps and, **without removing tubes from the thermal cycler, add 1 µL of** Target Extension Enzyme Mix (clear: TX2). Pipette mix up and down >5 **times with a volume equal to or greater than 40 µL. Reseal the tubes, close** the thermal cycler lid and immediately continue to the next step of the **program (72°C for 10 min, hold at 4°C).**

6. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice.
7. Continue immediately with the Enriched Library Purification protocol.

G. Enriched Library Purification

Note: this procedure involves two sequential bead purifications

QIAGEN Exhibit 1002
Page 761

V. Protocol

The purification beads should be removed from 4°C and left at bench top to reach room temperature well before the start of purification.

1. Retrieve the Agencourt RNAClean XP Beads from 4°C and ensure they are at room temperature before using. Also remove the DNA Resuspension Buffer (DR1) from –20°C and thaw at room temperature.
2. Resuspend the beads by inverting and tapping the tube. Ensure the beads are fully resuspended before adding to the sample. After resuspending, do not spin the beads. (An excess of beads is provided; therefore, it is not necessary to recover any trapped in the cap.)
3. Add 50 µL of room-temperature Nuclease-free Water (D1) to each Target Extension reaction.
4. Add 80 µL (0.8 volumes) room temperature bead suspension to each sample. Mix thoroughly by pipetting 10 times.
5. Incubate at room temperature for 5 minutes.
6. Transfer the tubes to the magnet plate and let stand 5 minutes to completely clear the solution of beads.
7. Carefully remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step.

   Note: The beads should not disperse; instead, they will stay on the walls of the tubes. Significant loss of beads at this stage will impact the amount of DNA carried into Library Amplification, so ensure beads are not removed with the binding buffer or the wash.
8. With the plate still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute.
9. Remove the ethanol wash using a pipette.
10. Air dry the beads on the magnet for a minimum of 10 minutes. Inspect each tube carefully to ensure that all the ethanol has evaporated. It is critical that all residual ethanol be removed prior to continuing.
11. Remove the tubes from the magnet.
12. Add 100 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature.
13. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads.
14. Carefully remove 100 µL of the eluate to a new tube.
15. Add 80 µL (0.8 volumes) room temperature bead suspension to the 100 µL eluate. Mix thoroughly by pipetting 10 times.
16. Incubate at room temperature for 5 minutes.
17. Transfer the tubes to the magnet and let stand 5 minutes to completely clear the solution of beads.

QIAGEN Exhibit 1002
Page 762

18. Carefully remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step.

    Note: The beads should not disperse; instead, they will stay on the walls of the tubes. Significant loss of beads at this stage will impact the amount of DNA carried into Library Amplification, so ensure beads are not removed with the binding buffer or the wash.

19. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute.

20. Remove and discard ethanol wash.

21. Air dry the beads on the magnet for at least 10 minutes at room temperature. Inspect each tube carefully to ensure that all the ethanol has evaporated. It is critical that all residual ethanol be removed prior to continuing.

22. Remove the tubes from the magnet.

23. Add 25 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature.

24. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads.

25. Carefully remove 24 µL of the eluate, ensuring as few beads as possible are carried over.

26. At this point, samples can be stored at –20°C until needed or can be used immediately for qPCR (if being performed) or in the Library Amplification protocol.

H. qPCR (Optional)

When using the Ovation Target Enrichment System for the first time with a new probe set or sample type, we recommend performing a qPCR assay in order to determine the number of cycles necessary for library amplification. See Appendix B for guidelines.

I. Library Amplification

1. Remove the Library Amplification Buffer Mix (red: P1) Amplification Primer Mix (red: P2), Amplification Enzyme Mix (red: P3) and Amplification Enhancer Mix (red: P6) from –20°C storage.

2. Spin down P3 and place on ice.

3. Thaw P1 and P2 at room temperature. Mix each by vortexing, spin and place on ice.

4. Prepare a master mix by combining P1, P2, P3, P6 and water in a 0.5 mL capped tube, according to the volumes shown in Table 8.

QIAGEN Exhibit 1002
Page 763

Mix by pipetting and spin down the master mix ***briefly. Place on ice. Use*** *immediately.*

Table 8. Amplification Master Mix (volumes listed are for a single reaction)

AMPLIFICATION BUFFER MIX (RED: P1 VER 4)	AMPLIFICATION PRIMER MIX (RED: P2 VER 14)	AMPLIFICATION ENZYME MIX (RED: P3 VER 2)	AMPLIFICATION ENHANCER MIX (RED: P6)	WATER (GREEN: D1)
20 µL	8 µL	2 µL	4 µL	44 µL

5. Add 78 µL of Library Amplification Master Mix to 22 µL of purified DNA from step G.25 above.
6. Mix well by pipetting.
7. Place the tubes in a pre-warmed thermal cycler programmed to run Program 4 (Library Amplification; see Table 4):

   37°C – 10 min, 95°C – 3 min, 15-30 cycles* (95°C – 30 sec, 62°C – 15 sec, 72°C – 20 sec), 72°C – 2 min, hold at 4°C
   *number of cycles determined by qPCR
8. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice.
9. Continue immediately with the Amplified Library Purification protocol.

J. Amplified Library Purification

Note: this procedure involves two sequential bead purifications

1. Retrieve the Agencourt RNAClean XP Beads from 4°C and ensure they are at room temperature before using. Also remove the DNA Resuspension Buffer (DR1) from –20°C and thaw at room temperature.
2. Resuspend the beads by inverting and tapping the tube. Ensure the beads are fully resuspended before adding to the sample. After resuspending, do not spin the beads. (An excess of beads is provided; therefore, it is not necessary to recover any trapped in the cap.)
3. Add 80 µL (0.8 volumes) room temperature bead suspension to each sample. Mix thoroughly by pipetting 10 times.
4. Incubate at room temperature for 5 minutes.
5. Transfer the tubes to the magnet plate and let stand 5 minutes to completely clear the solution of beads.
6. Carefully remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step.

QIAGEN Exhibit 1002
Page 764

Note: The beads should not disperse; instead, they will stay on the walls of the tubes. Significant loss of beads at this stage will impact the library yield, so ensure beads are not removed with the binding buffer or the wash.

7. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute.
8. Remove and discard the ethanol wash.
9. Air dry the beads on the magnet for at least 10 minutes at room temperature. Inspect each tube carefully to ensure that all the ethanol has evaporated. It is critical that all residual ethanol be removed prior to continuing.
10. Remove the tubes from the magnet.
11. Add 100 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature.
12. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads.
13. Carefully remove 100 µL of the eluate to a new tube, ensuring as few beads as possible are carried over.
14. Add 80 µL (0.8 volumes) room temperature bead suspension to the 100 µL eluate. Mix thoroughly by pipetting 10 times.
15. Incubate at room temperature for 5 minutes.
16. Transfer the tubes to the magnet and let stand 5 minutes to completely clear the solution of beads.
17. Carefully remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step.

    Note: The beads should not disperse; instead, they will stay on the walls of the tubes. Significant loss of beads at this stage will impact the library yield, so ensure beads are not removed with the binding buffer or the wash.
18. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute.
19. Remove and discard ethanol wash.
20. Air dry the beads on the magnet for at least 10 minutes at room temperature. Inspect each tube carefully to ensure that all the ethanol has evaporated. It is critical that all residual ethanol be removed prior to continuing.
21. Remove the tubes from the magnet.
22. Add 25 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature.

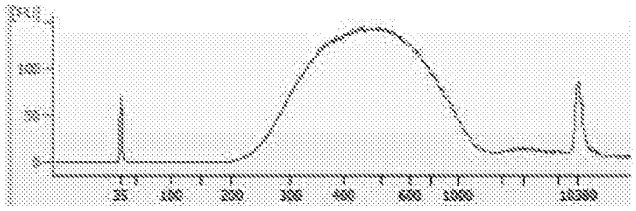
QIAGEN Exhibit 1002
Page 765

23. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads.
24. Carefully remove 24 µL of the eluate, ensuring as few beads as possible are carried over.
25. The library is now ready for quantification and sequencing and can be stored at –20°C until needed.

K. Quantitative and Qualitative Assessment of the Library

1. Run 0.5–1 µL of the sample on a Bioanalyzer High Sensitivity DNA Chip (Agilent Technologies). Typical fragment distributions are shown in Figures 4 and 5, below.

Figure 4. Fragment distribution when 1 µL of a target enriched library is loaded into a Bioanalyzer High Sensitivity DNA Chip.



Important: Concentration calculation of broad libraries should be performed using a library size of 450 bp to avoid underestimating the library concentration. We also recommend initially loading approximately 60% of the Illumina recommended input onto the flowcell to avoid potential over-clustering. The amount of library loaded can be optimized in subsequent sequencing runs.

2. Validate the library as described in the most recent Illumina user guides for DNA library construction, e.g., Genomic DNA Sample Prep Manual.

QIAGEN Exhibit 1002
Page 766

VI. Technical Support

For help with any of our products, please contact NuGEN Technical Support at 650.590.3674 (direct) or 888.654.6544, option 2 (toll-free, U.S. only). Send faxes to 888.296.6544 (toll-free) or email techserv@nugen.com.

In Europe contact NuGEN at +31(0)135780215 (Phone) or +31(0)135780216 (Fax) or email europe@nugen.com.

In all other locations, contact your NuGEN distributor for technical support.

QIAGEN Exhibit 1002
Page 767

VII. Appendix

A. Normalization of Libraries

When gDNAs are not well characterized/quantitated or of variable quality, it is important to determine library ready materials prior to pooling for the enrichment reaction if one is concerned about achieving near equal representation of all samples. Library ready materials are DNA fragments that have the Forward Adaptor ligated onto the 5′ end of the fragment and are of sufficient length and quality (no cross linked interference) to generate a qPCR product when used in combination with an Alu PCR primer.

The procedure below has been used successfully in our laboratories to balance libraries from divergent DNA sources.

Note: The qPCR reagents are not provided for this optional step and must be purchased separately.

qPCR Quantification of Ligation Products

1. Follow the Ovation Target Enrichment protocol through the adaptor ligation and bead purification steps.
2. Dilute the bead-purified ligation products 1000-fold and use as template in a qPCR reaction set up with the following pair of primers.

   P1 forward primer (specific to the ligation adaptor): 5′–CTCTTTCCCTACACGACGCTCTTCCGAT –3′

   Alu reverse primer (specific to Alu sequences in gDNA): 5′–GAGGCTGAGGCAGGAGAATCG –3′
3. Prepare a qPCR reaction mixture by combining the materials listed in Table 9. The qPCR dye can be either Eva Green or SYBR Green I.

Table 9. qPCR Reaction Mixture (volumes listed are for a single reaction)

5X MyTaq BUFFER	P1 FORWARD PRIMER, 10 µM	Alu REVERSE PRIMER, 10 µM	QPCR DYE AT 1X FINAL REACTION CONC.	MyTaq HS DNA POLYMERASE	DILUTED LIGATION PRODUCTS	WATER
2 µL	0.4 µL	0.4 µL	0.5 µL	0.1 µL	3.0 µL	3.6 µL

*Mix by pipetting and spin down the master mix **briefly. Place on ice. Use** immediately.*

QIAGEN Exhibit 1002
Page 768

4. Perform qPCR using the following protocol:
5. 95°C – 3 min, 30 cycles * (95°C – 15 sec, 60°C – 60 sec)
6. To create a standard curve for quantification, use a serial dilution of the bead purified ligation products (1:10, 1:100, 1:1000, 1:10000) from a reference library or library derived from high-quality DNA.

B. Guidelines for Optional qPCR

When using the Ovation Target Enrichment System for the first time with a new probe set or sample type, we recommend performing a qPCR assay after protocol step H.25 in order to determine the number of cycles necessary for subsequent library amplification.

1. Remove the Library Amplification Buffer Mix (red: P1) Amplification Primer Mix (red: P2), Amplification Enzyme Mix (red: P3) and Amplification Enhancer Mix (red: P6) from –20°C storage.
2. Spin down P3 and place on ice.
3. Thaw P1 and P2 at room temperature. Mix each by vortexing, spin and place on ice.
4. Prepare a master mix by combining P1, P2, P3, P6, water and dye (Eva Green or SYBR Green reagent) in a 0.5 mL capped tube, according to the volumes shown in Table 10.

Table 10.qPCR Master Mix (volumes listed are sufficient for 2 reactions and a small overage)

AMPLIFICATION BUFFER MIX (RED: P1 VER 4)	AMPLIFICATION PRIMER MIX (RED: P2 VER 14)	AMPLIFICATION ENZYME MIX (RED: P3 VER 2)	AMPLIFICATION ENHANCER MIX (RED: P6)	QPCR DYE AT 1X FINAL REACTION CONC.	WATER (GREEN: D1)
5 µL	2 µL	0.5 µL	1 µL	X	To total volume of 20 µL

Mix by pipetting and spin down the master mix briefly. Place on ice. Use immediately.

5. Add 8 µL of qPCR Master Mix to 2 µL of purified library DNA (or water for the negative control) in a qPCR plate and mix well.
6. Perform qPCR using the following protocol:

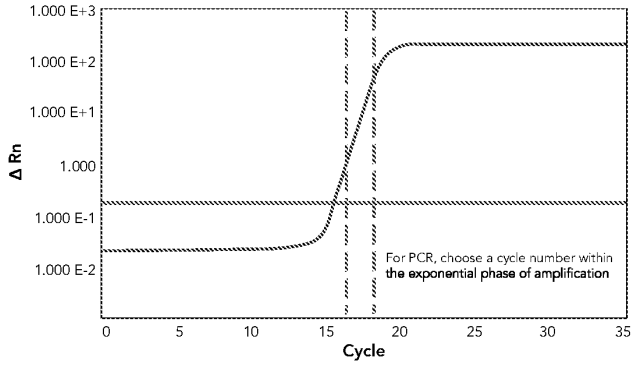
   37°C – 10 min, 95°C – 3 min, 35 cycles * (95°C – 30 sec, 62°C – 15 sec, 72°C – 20 sec)

The cycle number used for subsequent library amplification should be within the exponential phase of the amplification plot (17 or 18 cycles in Figure 5 as an example). There is sufficient reagent included to perform a single test qPCR reaction for each of the multiplexed libraries.

QIAGEN Exhibit 1002
Page 769

VII. Appendix

Figure 5. Stylized qPCR amplification plot.



C. Barcode Sequences and Guidelines for Multiplex Experiments

The Ovation Target Enrichment System 32-reaction kit contains 32 different barcodes and the 96-reaction kit contains 96 different barcodes that can be used for multiplex sequencing to interrogate several independent libraries on a single lane of the Illumina NGS platforms. Barcode sequences and multiplex guidelines for adaptors used in Ovation Target Enrichment Systems can be found in Table 11.

All barcode sequences are separated by an edit distance of three. For further details on the barcode design strategy, please refer to Faircloth BC, Glenn TC (2012) Not All Sequence Tags Are Created Equal: Designing and Validating Sequence Identification Tags Robust to Indels. *PLoS ONE* 7(8): e42543. doi:10.1371/journal.pone.0042543.

QIAGEN Exhibit 1002
Page 770

VII. Appendix

Table 11.The 8-reaction kit contains 8 barcodes shown in plate positions A01–H01. The 32-reaction kit contains the 32 barcodes shown in plate positions A01–H04. The 96-reaction kit contains 96 barcodes in plate positions as shown. Barcodes are color-balanced in sets of 8 by column on the plate. (i.e. A01–H01 is a color-balanced set of barcodes.

PLATE POSITION	BARCODE SEQUENCE	PLATE POSITION	BARCODE SEQUENCE	PLATE POSITION	BARCODE SEQUENCE
A01	CACGTCTA	A05	ATGGCGAT	A09	CAAGAAGC
B01	AGCTAGTG	B05	AACGCCTT	B09	CTGATGAG
C01	ACTATCGC	C05	GTAAGGTG	C09	CTCGAACA
D01	GCGTATCA	D05	TGTCGACT	D09	TCGACAAG
E01	ACTCTCCA	E05	ACTCTGAG	E09	AGTGCATC
F01	CGTCCATT	F05	GATGGAGT	F09	TGGCTACA
G01	AGCCGTAA	G05	CTAGCTCA	G09	GCACACAA
H01	GAGTAGAG	H05	CTGTACCA	H09	GCATAGTC
A02	ACGTCGTT	A06	CCTGTCAA	A10	AACACGCT
B02	GTCCTGTT	B06	GGTCGTAT	B10	TTCACGGA
C02	AGAAGCCT	C06	CGCTGATA	C10	TGCTGTGA
D02	GAAGATCC	D06	TAGCTTCC	D10	CCTCGAAT
E02	TAGCTGAG	E06	CAAGTCGT	E10	TGTGGCTT
F02	ACGTCCAA	F06	GTCTCATC	F10	CTGTGGTA
G02	CACACATC	G06	ACCAAGCA	G10	TCACTCGA
H02	CGGATCAA	H06	AGTCAGGT	H10	ACTCCTAC
A03	TCAGCCTT	A07	TATCGCGA	A11	CCACAACA
B03	AAGGCTCT	B07	TAGCAGGA	B11	CCGCTTAA
C03	TGTTCCGT	C07	AGAAGGAC	C11	GTGGTATG
D03	GGAATGTC	D07	TGAGCTGT	D11	GGTGTACA
E03	CATCCAAG	E07	CAGAGTGA	E11	TCTAGGAG
F03	GTCAACAG	F07	AGGTTCCT	F11	TGGAAGCA
G03	TCGCTATC	G07	AGACCTTG	G11	AACACCAC
H03	AGCCTATC	H07	CTTCCTTC	H11	CATACGGA
A04	TCGGATTC	A08	CAGGTTCA	A12	CTCTCAGA
B04	CGGAGTAT	B08	ACTGGTGT	B12	AAGCTGGT
C04	GAACCTTC	C08	GGATTCAC	C12	GTGTCCTT
D04	AGAGGATG	D08	CACGATTC	D12	AGGTCTGT
E04	ACGCTTCT	E08	AGACATGC	E12	CATTCGTC
F04	CACAGGAA	F08	GACACAGT	F12	CTCACCAA
G04	ACGAATCC	G08	CCAGTTGA	G12	GACTACGA
H04	CCTTCCAT	H08	CATGGATC	H12	ATACGCAG

QIAGEN Exhibit 1002
Page 771

VII. Appendix

D. Copy Number Alteration/Variation Analysis

The Ovation Target Enrichment Systems for DNA enable accurate measurement of Copy Number Alteration/Variation using the method described below. If you have any questions with respect to how to perform this method or interpret the results, please contact NuGEN Technical Support (techserv@nugen.com).

The general workflow for performing CNV analysis is diagrammed below, and requires the following files and programs:

Input files and formats:

- Forward sequencing reads, 100–150 bp (FASTQ)
- Index read information that contains 6-bp, or "N6" randomer for duplicate identification (FASTQ)
- probePlus300 bed file containing gene symbol or region annotation, provided by NuGEN (BED)
- Normalization factors file for each probe region (TXT)

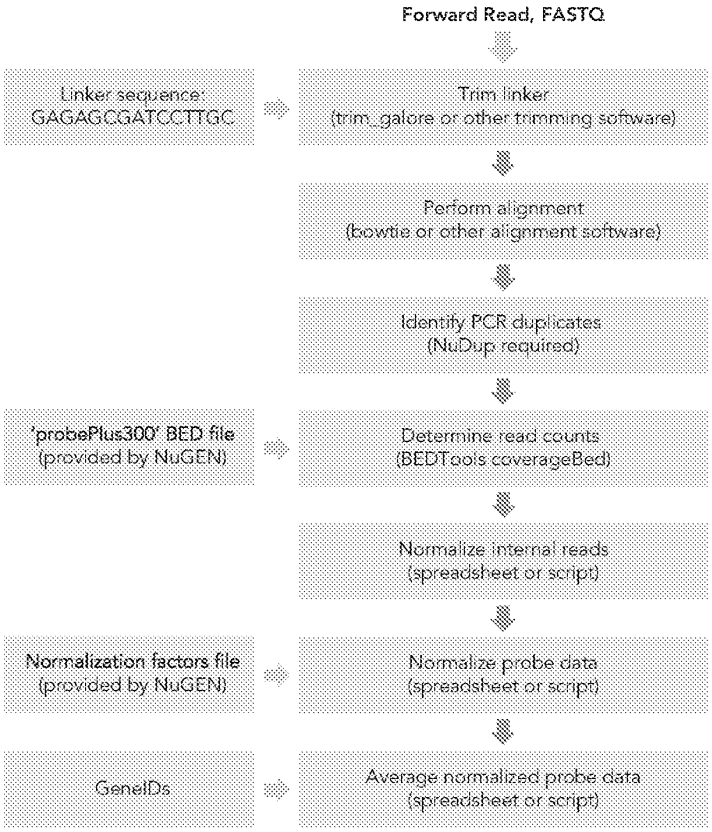
Software required:

- Python running in a Linux-based environment, version 2.7 or later
- Trimming software to remove probe linker (fastx_trimmer, trim_galore, cutadapt, bbduk)
- Read aligner (bowtie2, bwa)
- N6 PCR duplicate identification tool (NuDup, NuGEN BaseSpace App)
- BEDTools suite
- Spreadsheet program for simple arithmetic across probeset (Excel, OpenOffice)

QIAGEN Exhibit 1002
Page 772

VII. Appendix

Figure 6. Workflow for performing CNV analysis.



Analysis Workflow

Please follow the steps below to compute CNV information at the gene level. Single end or paired end datasets may be used for CNV quantitation.

1. Sequencing and N6 capture

   Please see Section III.E for details on how to properly set up a sequencing run to capture the N6 information from the Index read. This read information will be required to identify PCR duplicates and enable accurate quantitation of CNVs.

2. Trim linker

   Libraries generated with this system incorporate a 18-bp linker region that should be removed from all reads. As an example, the freely available program 'trim galore' can be used as follows to remove the linker sequence:

QIAGEN Exhibit 1002
Page 773

trim_galore -a GAGAGCGATCCTTGC R1_mysample.fq &> trim_log.txt

This command will remove any occurrence of the linker sequence and all downstream (3') bases, which will be derived from the adaptor. Removal of the linker sequence will enable more accurate alignment and quantitation.

3. Perform alignment

   Alignment can be performed using a number of different programs. In most cases, it is recommended to align to the entire reference genome (not just the target region). It is also recommended to only align reads that map two or fewer times to the genome to ensure accurate quantitation. An example command to perform alignment is:

   bowtie2 --sensitive -x hg19.bowtie2 -U R1_mysample_trimmed.fq -S R1_mysample_trimmed_bowtie2.sam &> bowtie2Report.txt

4. Identify PCR duplicates

   PCR duplicate identification requires use of NuGEN's software tool 'NuDup', which is available for download (http://nugentechnologies.github.io/nudup/). The software requires use of the 14-bp index read in order to identify PCR duplicates. An example of the command-line usage is:

   python nudup.py –f index_read_mysample.fq –o R1_mysample_trimmed_bowtie2 R1_mysample_trimmed_bowtie2.sam

   This will result in two output files; one that contains all reads with the duplicates marked (R1_mysample_trimmed_bowtie2 .sorted.markdup.bam), and a second file with the duplicate reads removed (R1_mysample_trimmed_bowtie2.sorted.dedup.bam). The read with the highest mapping quality is kept as the non-duplicate read.

5. Determine read counts for each probe region

   Use of the BEDTools Suite routine 'coverageBed' with the NuGEN-supplied "probePlus300" file will provide quantitation of the number of reads that fall within a 300 bases of the 3' end of each probe. This routine can be run using a command such as:

   coverageBed -abam R1_mysample_trimmed_bowtie2.sorted.dedup.bam -b probePlus300.bed > R1_mysample_trimmed_bowtie2.sorted.dedup.bam_coverage_probePlus300.txt

   The output from this routine is the probePlus300 bed file entry with the following appended columns:

   COL7: The number of reads that overlapped (by at least one base pair) the probePlus300 interval

   COL8: The number of bases in the region that had non-zero coverage

   COL9: The length of the region (always 300)

QIAGEN Exhibit 1002
Page 774

COL10: The fraction of bases in the region that had non-zero coverage

Values from COL7 are used in the next step of the process.

6. Normalize internal reads

   The output of coverageBedmust first be normalized for total reads in the probe regions. This is easily accomplished using a spreadsheet program such as Microsoft Excel. Divide each value in COL 7 by the SUM(COL7) to normalize for fraction of total reads.

7. Normalize probe data

   The read-normalized output from step 6 should then be adjusted using the normalization factors file provided by NuGEN. Normalization is accomplished by dividing the read normalized values by the corresponding probe normalization factor from the normalization factors file. This step provides normalization of probe signal due to small variations in probe efficiency for each probeset. The number should then be multiplied by 2, as we are assuming the normalization values correspond to 2 copy counts.

8. Average normalized probe data

   A single copy number value per gene is calculated by averaging the normalized values from Step 7 for each gene. Note that the number of probe regions for each gene can vary depending upon the size of the gene. This step is also easily accomplished using a spreadsheet program or a simple script.

QIAGEN Exhibit 1002
Page 775

VII. Appendix

E. Frequently Asked Questions (FAQs)

Q1. What is provided in the Ovation Target Enrichment System?
The custom design, target enrichment probes and all reagents necessary to create custom target-enriched, singleplex or multiplex libraries from genomic DNA ready for sequencing on the Illumina platform are included in the system.

Q2. Do I need any additional reagents or instruments?
Any laboratory that is set up for sequencing on the Illumina platform should have all the necessary reagents and labware to perform the library preparation and target enrichment. (see section II.B for a list of additional equipment, reagents and labware).

Q3. What are the available kit sizes?
The custom kits are available as 32 or 96 rxn size with the corresponding number of unique barcodes.

Q4. What information do I need to provide NuGEN for a custom design?
A member of the NuGEN Technical Services team will work with you ensure the correct information is provided for a custom design. RefSeq IDs or genomic coordinates will generally suffice.

Q5. What is the turnaround time of a custom probe design?
If the targets are provided in the appropriate format, it should take no more than a couple of days for the design to be complete. For non-human sequence this process may take longer in order for the correct genome sequence to be identified and added to the design pipeline. We estimate 3–4 weeks between obtaining customer approval on a finalized design and shipping the custom kit.

Q6. What is the size of the probe?
The exact sizes of the probes is proprietary information. We do provide the genomic coordinates and 50 bases of sequence information that includes the probe.

Q7. How do you recommend shearing the DNA?
We generally recommend Covaris focused acoustic shearing, but initial trials have suggested that other methods (fragmentase) should also work.

Q8. Can the workflow be automated?
The Ovation Target Enrichment System features a simple add and incubate workflow, so it can be automated. Contact your NuGEN Account Executive for more details if you wish to automate the system.

Q9. How well does gDNA from FFPE material work in the system?
FFPE material works well in the system and works best with a probe design optimized to accommodate the smaller expected fragment sizes of FFPE genomic DNA. Because the DNA is modified and degraded, we suggest starting with a larger amount of input material (more than 100 ng).

QIAGEN Exhibit 1002
Page 776

VII. Appendix

Q10. Can the Ovation Target Enrichment System be used on cell-free DNA (plasma DNA)?
Cell-free DNA might require an optimized probe design that addresses smaller fragments, similar to FFPE material. If there is sufficient plasma DNA as input, this may be possible, but we have not yet performed the experiment.

Q11. Can the Ovation Target Enrichment System be used on single cell level?
Genomic DNA from a single cell would likely require a preamplification step prior to the library prep and enrichment. We have not yet performed this study.

Q12. Can regions with pseudogenes be enriched?
Yes. Probes can be designed for regions that target the differences between genes and pseudogenes.

Q13. How does the probe design cover long target regions?
For larger target regions, probes are tiled within the target region in order to make sure that the entire target region is accurately represented.

Q14. Can the system accommodate long reads?
Forward reads up to 150 bp should work without a problem. Longer reads can be performed but may require a larger average fragment size and the read may extend into the probe.

Q15. Can I do paired end sequencing of an Ovation Target Enrichment System library?
Yes. For most applications single end reads are sufficient and advised. You may want to perform paired end sequencing when looking specifically for rearrangements like inversions or translocations. However, due to the probe extension technology employed by the target enrichment system, the first 68 bases of reverse read sequence are probe- and linker-derived and may require trimming prior to alignment.

Q16. What library size should I use to calculate the library concentration?
For target enrichment libraries constructed from high quality DNA or FFPE DNA, use the average fragment size calculated by the Agilent Bioanalyzer or other methods of size determination.

QIAGEN Exhibit 1002
Page 777

F. Update History

This document, the Ovation Target Enrichment System User Guide (M01372 v8) is an update to address the following topics:

Description	Section	Page(s)
Correction to Figure 6, CNV Analysis Workflow	VII.D.	36

NuGEN Technologies, Inc.

Headquarters USA

201 Industrial Road, Suite 310
San Carlos, CA 94070 USA
Toll Free Tel: 888.654.6544
Toll Free Fax: 888.296.6544
custserv@nugen.com
techserv@nugen.com

Europe

P.O. Box 109
9350 AC Leek
The Netherlands
Tel: +31-13-5780215
Fax: +31-13-5780216
europe@nugen.com

For our international distributors contact information, visit our website

www.nugen.com

© 2013–2016 NuGEN Technologies, Inc. All rights reserved. The Encore®, Ovation® and Applause® families of products and methods of their use are covered by several issued U.S. and International patents and pending applications (www.nugen.com). NuGEN, Ovation, SPIA, Ribo-SPIA, Applause, Encore, Prelude, Mondrian and Imagine More From Less are trademarks or registered trademarks of NuGEN Technologies, Inc. Other marks appearing in these materials are marks of their respective owners.

For research use only.

PLOS | ONE

RESEARCH ARTICLE

An Efficient Method for Identifying Gene Fusions by Targeted RNA Sequencing from Fresh Frozen and FFPE Samples

Jonathan A. Scolnick, Michelle Dimon[¤a], I-Ching Wang, Stephanie C. Huelga, Douglas A. Amorese*

NuGEN Technologies, Inc., San Carlos, California, United States of America

¤a Current address: Google, Inc., Mountain View, California, United States of America
* damorese@nugen.com

CrossMark
click for updates

OPEN ACCESS

Citation: Scolnick JA, Dimon M, Wang I-C, Huelga SC, Amorese DA (2015) An Efficient Method for Identifying Gene Fusions by Targeted RNA Sequencing from Fresh Frozen and FFPE Samples. PLoS ONE 10(7): e0128916. doi:10.1371/journal.pone.0128916

Editor: Keping Xie, The University of Texas MD Anderson Cancer Center, UNITED STATES

Received: January 5, 2015

Accepted: May 1, 2015

Published: July 1, 2015

Copyright: © 2015 Scolnick et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All sequencing files are available from the SRA database (accession number SRP051118).

Funding: NuGEN Technologies provided funding for this work. The funder provided support in the form of salaries for all authors but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Abstract

Fusion genes are known to be key drivers of tumor growth in several types of cancer. Traditionally, detecting fusion genes has been a difficult task based on fluorescent *in situ* hybridization to detect chromosomal abnormalities. More recently, RNA sequencing has enabled an increased pace of fusion gene identification. However, RNA-Seq is inefficient for the identification of fusion genes due to the high number of sequencing reads needed to detect the small number of fusion transcripts present in cells of interest. Here we describe a method, Single Primer Enrichment Technology (SPET), for targeted RNA sequencing that is customizable to any target genes, is simple to use, and efficiently detects gene fusions. Using SPET to target 5701 exons of 401 known cancer fusion genes for sequencing, we were able to identify known and previously unreported gene fusions from both fresh-frozen and formalin-fixed paraffin-embedded (FFPE) tissue RNA in both normal tissue and cancer cells.

Introduction

Cancer cells frequently contain chromosomal rearrangements that lead to the formation of fusion genes expressed in the cell [1]. These fusion genes can act as drivers for cell growth. For example, the Philadelphia Chromosome rearrangement, originally identified in chronic myelogenous leukemia [2] (CML), is the result of a translocation between chromosomes 9 and 22 which leads to the expression of a fusion gene combining the BCR and ABL kinases [3]. While the BCR-ABL fusion gene leads to uncontrolled cell growth, when the fusion gene is identified in a patient, CML can be treated successfully with tyrosine kinase inhibitors.

Gene fusions are typically identified in cells by fluorescent *in situ* hybridization (FISH), a technique in which selected regions of chromosomes are fluorescently labeled through the hybridization of specific oligonucleotide probes. Aclinician must then visually identify two chromosomal regions that have rearranged in a known pattern. FISH has many problems as a technique including being a difficult, low throughput procedure, required knowledge of both

QIAGEN Exhibit 1002
Page 779

Competing Interests: JS, IW, SH and DA are current employees of NuGEN Technologies and MD was a paid consultant to NuGEN Technologies. NuGEN Technologies also provided the funding for this research and markets the Ovation Target Enrichment Fusion Panel described in the manuscript. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

gene fusion partners, low spatial resolution, the availability of the necessary fluorescent probes and the need for highly trained personnel to decide if a fusion event has taken place [4]. These difficulties apply to both clinical and research laboratories, thus limiting the potential understanding of gene fusion events in human biology.

One recent alternative to FISH when studying gene fusions is high throughput sequencing. In particular, RNA sequencing (RNA-Seq) has been used to identify gene fusions that are transcribed into RNAs within various cells [5]. Software has been specifically developed to identify gene fusion events in RNA-Seq data [6] and many previously unknown gene fusions have been identified this way. However, while RNA-Seq is a powerful tool for identifying gene fusion transcripts, it is currently cost prohibitive to sequence RNA from a tumor sample, in which only a small fraction of the total cells are expressing the gene fusion, and obtain enough sequencing reads to identify those fusion events. Here we describe an innovative new assay for gene fusion discovery and verification. Based on Single Primer Enrichment Technology (SPET, Fig 1A), we target particular RNAs for sequencing, thus reducing the number of uninformative sequencing reads and increasing the sensitivity of gene fusion detection compared to standard RNA-Seq methods. The SPET based assay is easy to use, has low RNA input requirements and can be used with RNA from formalin fixed, paraffin embedded (FFPE) tissue, which is important for clinically relevant samples. In addition, the assay is fully customizable to target any gene or set of genes in any genome. To show the value of this technology, we have created a panel of probes to target 401 human genes based on gene fusion events found in the Wellcome

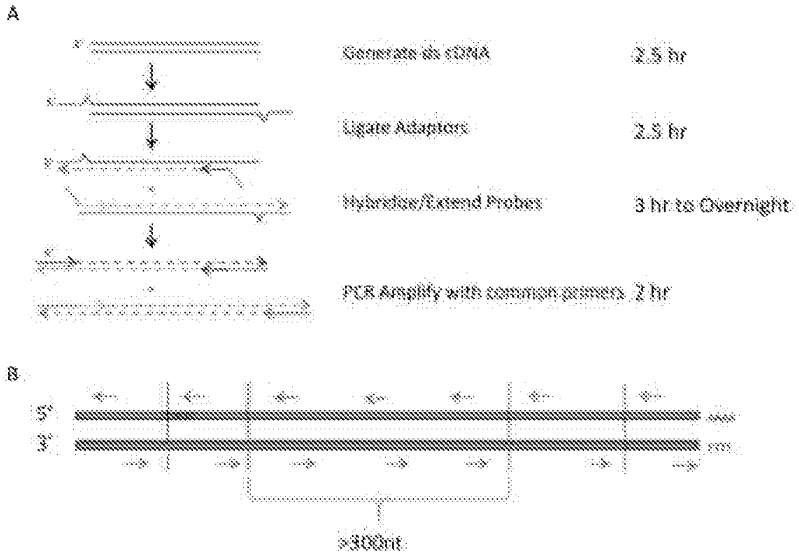


Fig 1. Description of the Ovation Target Enrichment System. (A) Experimental steps of the assay and time required for each step. Adaptors (green) are ligated on to generated double stranded cDNA (ds-cDNA). Probes (shown in red and yellow) are hybridized to target cDNA and extended with a polymerase (dashed grey lines). All probes have common tail sequences (blue), which are used as priming sites along with adaptor sequences in subsequent library amplification PCR steps. (B) Example of probe positioning across different exons in a full length double stranded cDNA. Each exon (demarked by blue vertical lines) will have probes (green arrows; arrow points in the 3' direction) designed to hybridize near the predicted exon-exon junctions. Exons larger than 300 nucleotides (nt) may have additional probes tiled along the length of the exon to obtain more complete sequence coverage. Probes are designed against both strands of the cDNA to enable identification of gene fusions when only one of the pair of genes is targeted. Translation start sequence (ATG) and poly A tail are labeled.

doi:10.1371/journal.pone.0128916.g001

QIAGEN Exhibit 1002
Page 780

Trust databases [1], Chimerdb 2.0 [7] and The Cancer Genome Atlas (http://cancergenome.nih.gov/) (full list in S1 Table) and have used this probe panel to identify known and previously unidentified gene fusion events in cancer cells lines and human whole brain RNA.

Results

Targeted RNA-Sequencing

In order to test the ability of the Ovation Fusion Panel Target Enrichment System assay to effectively target RNAs of interest, we ran the assay using 100 ng of total RNA from the H2228, lung adenocarcinoma cell line to see what percentage of sequencing reads were derived from the targeted genes. The resulting library was sequenced on an Illumina MiSeq with paired end 75bp reads. 92.5% of reads mapped back to the genome with 96% of those reads mapping to the targeted RNAs. These results indicate a strong selection for targeted RNAs by the Ovation Fusion Panel Target Enrichment System. Comfortable that we were selecting almost exclusively for the sequences we were interested in, we next looked to identify gene fusions.

Gene Fusion Identification in H2228 Cells

A key feature of our assay is the ability to identify gene fusions even when only one of the partners has been targeted. This is achieved by through the use of Single Primer Enrichment Technology (SPET), which only requires the hybridization of a single primer in order to obtain sequencing information from the downstream RNA. H2228 cells carry one allele of the *ALK* gene that is fused to *EML4* and the other allele is fused to *PTPN3*. In our targeting panel we have probes targeting *ALK* and *EML4*, but not *PTPN3* so this served as a good test for identifying fusion transcripts for which only one of the genes was directly targeted. The sequencing data were analyzed for the presence of gene fusions (see materials and methods) and we were able to identify both the predicted *ALK-EML4* and the *ALK-PTPN3* gene fusions with 512 and 20 reads respectively for each fusion out of a total of 4.5 million reads.

Targeted RNA Sequencing Identifies More Fusions than Standard RNA-Seq

We next wanted to compare our results to those of standard RNA-Seq experiments to test how many sequencing reads were necessary to identify gene fusion events when using the Ovation Fusion Panel Target Enrichment System relative to a non-targeted RNA-Seq approach. Universal Human Reference RNA (UHR) contains RNA from ten different cancer cell lines. Many standard RNA-Seq experiments have been performed using UHR RNA and these experiments have identified a number of gene fusion events [8]. Our 401 gene panel contains probes that target two of the fusions previously shown to exist in UHR RNA, *BCR-ABL* and *NUP214-XKR3*. 100 ng of UHR total RNA was used to prepare sequencing libraries with Ovation Fusion Panel Target Enrichment System. Libraries were sequenced on either the Illumina MiSeq or HiSeq and analyzed with SOAPFuse [6]. Our initial tests utilizing 1.6M sequencing reads resulted in the identification of both *BCR-ABL* and *NUP214-XKR3* fusion transcripts. In addition, we also identified an *FGFR1-WHSC1L1* gene fusion, which has been reported as a recurring fusion in breast cancer [9], but has not been reported in UHR RNA. Deeper sequencing of another library to 8.7M reads on the Illumina HiSeq led to the further identification of another previously unreported UHR gene fusion, *MLLT10* fused to *PICALM*. Having identified a set of gene fusions in UHR RNA, we compared the number of fusion reads in these genes detected by SOAPFuse in our datasets to those of a publically available RNA-Seq dataset that consists of 125M paired end sequencing reads [10]. As seen in Fig 2, when comparing 1.6M

QIAGEN Exhibit 1002
Page 781

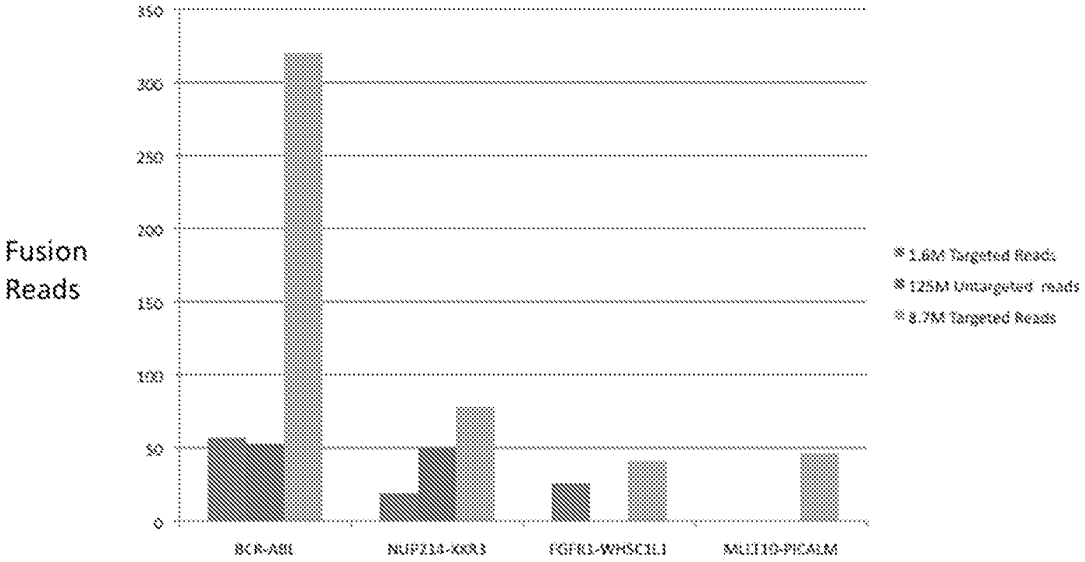


Fig 2. Ovation Fusion Panel Target Enrichment System identifies known and novel gene fusion events in Universal Human Reference RNA. The number of sequencing reads determined to be derived from gene fusions in two different targeted sequencing libraries (blue and green) compared to the events identified in a standard, untargeted RNA-Seq library. The untargeted library (red) consists of 125 million total sequencing reads while the targeted libraries consist of 1.6 and 8.7 million sequencing reads.

doi:10.1371/journal.pone.0128916.g002

Ovation Fusion Panel Target Enrichment System sequencing reads (blue bars) to 125M standard RNA-Seq sequencing reads (red bars), fusion read counts are comparable for the *BCR-ABL* fusion. The 1.6M targeted sequencing reads has fewer reads for *NUP214-XKR3*, which is expected because our targeting panel only contains probes for *NUP214* and not *XKR3*. The standard RNA-Seq does not contain any detectable reads for *FGFR1-WHSC1L1* nor any reads for *MLLT10-PICALM*, while the targeted sequencing approach described here identified both of these fusion events. Both of these previously undescribed fusions were verified by RT-PCR and Sanger sequencing. In summary, by using a targeted sequencing approach, we were able to identify known gene fusion events in UHR RNA while using only 1.3% of the sequencing reads necessary to identify approximately the same number of fusion reads in a standard RNA sequencing experiment. In addition we were able to identify previously undocumented gene fusion events using the targeted approach that were not identified in standard RNA sequencing, even though the targeted sequencing had 14–78X fewer sequencing reads. The inability of standard RNA sequencing to identify fusion transcripts in a cell line that has been studied in detail suggests that even samples for which deep RNA sequencing data already exists should be tested again specifically for the presence of fusion transcripts.

Human Tissue Gene Fusions

We next sought to test our fusion panel assay on RNA derived from human tissue rather than cultured cells. For this, we utilized human brain reference RNA, a pooled RNA sample derived from 23 healthy donors. Brain tissue has been shown to contain multiple genomic lesions [11] and so we wondered if any of these genomic lesions would lead to the detectable expression of gene fusions. As with UHR RNA, 100 ng of brain reference RNA was used to generate targeted sequencing libraries with Ovation Fusion Panel Target Enrichment System, which were

QIAGEN Exhibit 1002
Page 782

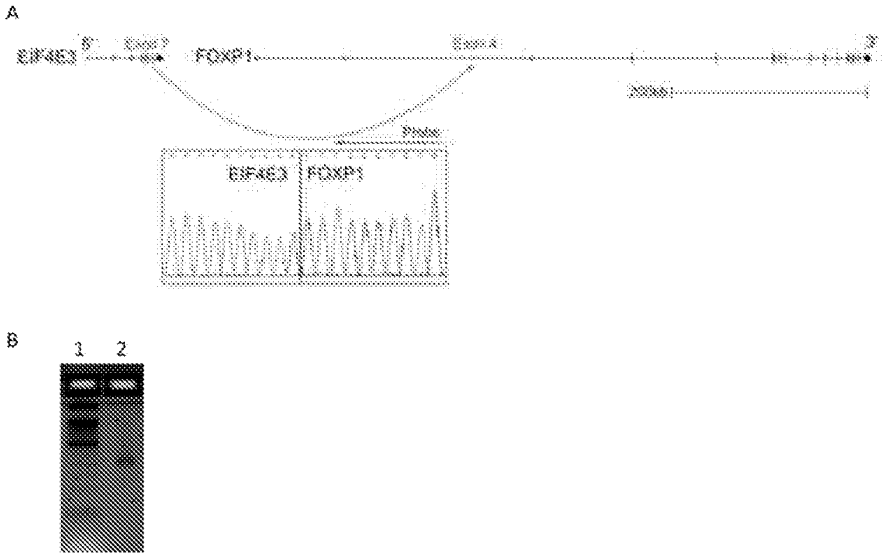


Fig 3. EIF4E3-FOXP1 fusion transcript. (A) Structure of the fusion transcript based on UCSC genome browser tracks. Red arrow indicates the in frame fusion of EIF4E3 exon7 to FOXP1 exon 4. The Sanger sequencing trace below indicates the sequence of the fusion point, while the blue arrow over the sanger sequence trace indicates the 3' end of the targeting probe. (B) RT-PCR result using PCR primers indicated described in S2 for detecting this fusion transcript. Lane 1: RT-PCR product showing the correct 209bp size, Lane 2: 50bp ladder.

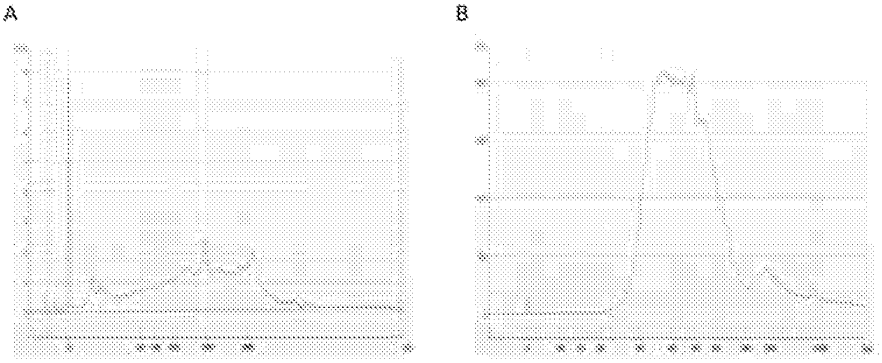
doi:10.1371/journal.pone.0128916.g003

sequenced on an Illumina HiSeq with paired-end, 100base reads. Approximately 7.2M reads were obtained and analyzed for the presence of gene fusions with SOAPFuse. Upon analysis, we identified two potential gene fusions, *ATP5L-MLL* and *EIF4E3-FOXP1*. The *EIF4E3-FOXP1* fusion (Fig 3) again shows the power of only requiring one gene be targeted out of a pair of fused genes as our targeting primers were only designed to *FOXP1*. To our knowledge, this is the first report of *EIF4E3* being fused to *FOXP1* and the fact that this fusion transcript, along with *ATP5L-MLL*, is found in normal tissue suggests that fusion transcripts may occur more often than currently appreciated.

Gene Fusion Detection from an FFPE RNA Sample

Many primary tissue samples are preserved by formalin fixation followed by paraffin embedding (FFPE). This common procedure is known to fragment nucleic acids making it difficult to retrieve sequence data from them. We extracted RNA from HCC1937 primary breast ductal carcinoma cells that had been formalin-fixed and paraffin-embedded (Bioanalyzer trace of RNA in Fig 4A). We used 50 ng and 100 ng of the FFPE RNA in the Ovation Fusion Panel Target Enrichment System assay (resulting library Bioanalyzer trace Fig 4B). Following sequencing 1.2M and 1.6M reads respectively, we were able to identify the known *NFIA-EHF* gene fusion [12] even though we only targeted NFIA and did not have probes targeting EHF in this assay. We again compared our targeted data to a public dataset [13] of 24.6 million untargeted RNA sequencing reads from fresh RNA and found that the untargeted reads identified 11 fusion reads, while the targeted RNA sequencing produced 12 and 20 fusion reads for the 50 ng and 100 ng inputs respectively.

QIAGEN Exhibit 1002
Page 783



C

FFPE RNA Input	Total Reads	% Reads Aligned	% Reads On Target
50ng	1.2M	95.9	94.7
100ng	1.7M	94.7	94.7

Fig 4. HCC1937 Breast ductal carcinoma RNA. (A) FFPE RNA Bioanalyzer trace. (B) Bioanalyzer trace of sequencing library derived from 100 ng of RNA input shown in A. (C) Sequencing metrics for targeted RNA show that FFPE RNA is efficiently targeted.

doi:10.1371/journal.pone.0128916.g004

Discussion

We have applied our customizable SPET technology, previously employed for SNP detection within genomic DNA samples [13], to develop an assay for identification of gene fusion transcripts through targeted RNA sequencing. The Ovation Fusion Panel Target Enrichment System only requires a single primer to hybridize to a fusion transcript in order to identify both gene partners, allowing any potential fusions involving targeted genes to be interrogated in a single tube assay. We have developed a panel to target 401 genes known to form fusion transcripts and have tested the assay on RNA from cell lines, fresh tissue and FFPE samples. In each case we have identified known fusion transcripts and have additionally identified previously unreported fusion transcripts in UHR and normal brain RNA. Furthermore, we show that using a targeted sequencing approach is an efficient mechanism for enhancing the detection of gene fusions compared to standard RNA sequencing suggesting that even samples that have been deeply sequenced by standard RNA sequencing methods would benefit from a targeted approach to identifying gene fusion events.

Materials and Methods

Single Primer Enrichment Technology for Gene Fusion Detection

Total RNA was obtained from various sources. RNA derived from the commercial cell line NCI-H2228 (ATCC CRL-5935) was a gift from Susana Ortiz, M.D., Ph.D. (University of California, San Francisco). Universal Human Reference RNA (UHR) (Agilent Technologies) and Human Brain RNA (Ambion) were purchased and used without modification. IRB approval was not required as the human RNA samples used were reference samples obtained from cadavers and purchased from a commercial source. No identifying patient information was obtained by the authors. FFPE sections of cell line HCC1937 were purchased from Alamak

QIAGEN Exhibit 1002
Page 784

Bioscience. 10 micron sections were cut and RNA was extracted from these samples with the Recoverall Total Nucleic Acid Extraction Kit (Ambion). Total RNA was used as input into the Ovation Fusion Panel Target Enrichment System to generate targeted sequencing libraries following the manufacturer's recommended protocol in the User Guide and described in Fig 1 (PN # 9103 and 9104, NuGEN Technologies, http://www.nugen.com/nugen/index.cfm/products/ovation/ovation-fusion-panel-tes/). Briefly, utilizing the contents of the Ovation Fusion Panel Target Enrichment System kit, 50 or 100 ng of total RNA was reverse transcribed into cDNA. The unpurified cDNA was then made double stranded by the addition of a DNA polymerase. The double stranded cDNA (ds-cDNA) underwent end repair to blunt the ends followed by barcoded adapter ligation. Ligated samples were bead purified and hybridizated with targeting probes starting at 95°C followed by a slow ramp down to 60°C overnight. Following probe annealing, a DNA polymerase was added to the mixture which was heated to 72°C to extend the probes through the target region (see Fig 1). Unextended probes were removed by bead purification and the resulting products were amplified by PCR to create a sequence ready library. Sequencing libraries were quantitated by Bioanalyzer (Agilent Technologies) and Kapa Library Quantification Kit (KAPA Biosytems). Libraries were diluted to 10nM and sequenced on either the Illumina MiSeq or HiSeq platforms as indicated in main text as paired-end reads, either 75 or 100 bases in length (MiSeq and HiSeq respectively).

Sequence Data Analysis

Resulting sequences were analyzed in the following ways: 1. To determine the reads that were derived from targeted RNA sequences, forward reads were mapped to the Hg19 reference genome using STAR v2.4.0d [15] only allowing unique alignment and allowing chimeric reads. Bedtools intersect [16] was then used to count the number of sequenced reads that were derived from targeted transcripts using default settings, and 2. paired-end sequence data were analyzed for the presence of gene fusions with SOAPFuse v1.22 software [8] with default settings.

For publically available UHR RNA-Seq data, we downloaded SRX333353 [10] and for HCC1937 we downloaded SRX317716 [13].

Fusion Transcript Validation

Selected putative gene fusions were verified by designing PCR primers around the predicted fusion site. Reverse transcription PCR was used to amplify the predicted fusion gene junctions from the same starting RNA material as was used for sequencing. Resulting PCR products were purified and Sanger sequenced to verify the fusion junction. 200 ng of total RNA input was used in the Ovation cDNA Module For Target Enrichment to generate the cDNA used for the PCR reaction (included as part of the Ovation Fusion Panel Target Enrichment kit). PCR was performed in 50 ul reactions using standard Taq buffer, 0.5mM dNTPs and standard Taq Polymerase (NEB) with 0.4 uM each primer. The PCR reaction was carried out with the following program: 95°C, 3 minutes, followed by 30 cycles of 95°C, 15 seconds, 57°C, 20 seconds and 72°C, 10 seconds. See S2 Table for primer sequences.

Probe Selection

401 genes were selected for inclusion in the initial gene fusion targeting panel. All refseq exons were input into the proprietary probe selection software with the goal of designing at least two probes targeting each exon. For exons smaller than 300 bases probes were designed so that the 3' end of the probe would hybridize as near as possible to the known exon boundary, with one probe designed to hybridize to each strand of the double stranded cDNA. For exons larger than 300 bases, additional probes were tiled across the exon at 300 base intervals in order to

QIAGEN Exhibit 1002
Page 785

maximize coverage across these exons. Placing probes near the exon boundaries ensures that extension products will cross at least one exon-exon junction, thus increasing the probability of extending across a gene fusion junction. These design parameters resulted in the generation of 17,999 unique targeting probes across 5769 exons.

Supporting Information

S1 Table. List of Ovation Fusion Panel Target Enrichment System targeted genes.
(TXT)

S2 Table. PCR primers used for validation of gene fusion events and Sanger sequence of fusions. Forward and reverse primer sequences used in RT-PCR validation. For Sanger sequences, | denotes the breakpoint between sequence aligning to one gene versus the other in the fusion pair.
(TXT)

Acknowledgments

The authors would like to acknowledge Katy Planey for her bioinformatics support.

Author Contributions

Conceived and designed the experiments: JS DA MD. Performed the experiments: JS. Analyzed the data: JS MD ICW SH. Wrote the paper: JS DA.

References

1. Forbes S, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, et al. (2014) COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res Advanced Access. cancer.ac.sanger.uk
2. Rowly JD (1973) Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature 243(5405):290–3. PMID: 4126434
3. Heisterkamp N, Stam K, Groffen J, De Klein A, Grosveld G (1985) Structural organization of the *bcr* gene and its role In the Ph' translocation. Nature 315:758–761. PMID: 2989703
4. Liehr T (2010) Fluorescence In Situ Hybridization (FISH)–Quality Issues In Cytogenetics; p. 315–320. In Kristoffersson U., Schmidtke J. and Cassiman J. (Eds.), Quality Issues In Clinical Genetic Services, Springer Netherlands, Houten, Netherlands.
5. Ozsolak F, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12:87–98. doi: 10.1038/nrg2934 PMID: 21191423
6. Jia W, Qui K, He M, Song P, Zhou Q, Zhou F, et al. (2013) SOAPFuse: an algorithm for identifying fusion transcripts from paired-end RNA-seq data. Genome Biology 14:R12 doi: 10.1186/gb-2013-14-2-r12 PMID: 23409703
7. Kim P, Yoon S, Kim N, Lee S, Ko M, Lee H, et al. (2009) ChimerDB 2.0 –a knowledgebase of fusion genes updated. Nucleic Acids Res. 38:D81–D85. doi: 10.1093/nar/gkp982 PMID: 19906715
8. Kim D, Salzberg S (2011) TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biology 12:R72. doi: 10.1186/gb-2011-12-8-r72 PMID: 21835007
9. Chinnaiyan A, Kumar-Sinha C, Robinson D, Kalyana-Sundaram S (2011)Recurrent gene fusions in breast cancer. US patent US20130096021.
10. Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, et al. (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Research 14:R95.
11. McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, et al. (2013) Mosaic copy number variation in human neurons. Science 342:632–637. doi: 10.1126/science.1243472 PMID: 24179226

QIAGEN Exhibit 1002
Page 786

12. Ha KCH, Lalonde E, Li L., Cavallone L, Natrajan R, Lambros MB, et al. (2011) Identification of gene fusion transcripts by transcriptome sequencing in *BRCA1*- mutated breast cancers and cell lines. BMC Med Genomics 4:75. doi: 10.1186/1755-8794-4-75 PMID: 22032724

13. Daemen A, Griffith OL, Heiser LM, Wang NJ, Enache OM, Sanborn Z, et al. (2013) Modeling precision treatment of breast cancer. *Genome Biology* 14(10):R110.

14. Durinck S, Stawiski E, Pavia-Jimenez A, Modrusan Z, Kapur P, Jaiswal BS, et al. (2014) Spectrum of diverse genomic alterations define non-clear cell renal carcinoma subtypes. Nature Genetics. doi: 10.1038/ng.3146

15. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. (2012) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. doi: 10.1093/bioinformatics/bts635 PMID: 23104886

16. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. doi: 10.1093/bioinformatics/btq033 PMID: 20110278

QIAGEN Exhibit 1002
Page 787


NIH Public Access
Author Manuscript
Nat Methods. Author manuscript; available in PMC 2009 November 1.

Published in final edited form as:
Nat Methods. 2009 May ; 6(5): 315–316. doi:10.1038/nmeth.f.248.

Massively parallel exon capture and library-free resequencing across 16 individuals

Emily H. Turner[1], **Choli Lee**[1], **Sarah B. Ng**[1], **Deborah A. Nickerson**[1], and **Jay Shendure**[1]
1*Department of Genome Sciences, University of Washington, Seattle, WA 98195*

To the Editor: The adoption of molecular inversion probes (MIPs) to massively parallel exon capture[1] has been limited by representational and allelic bias. We describe modifications enabling simultaneous amplification and accurate shotgun sequencing of 50,000 exons. We also prove the scalability and accuracy of direct sequencing of MIP amplicons, which circumvents all shotgun library construction, while resequencing 1 megabase of coding sequence across 16 human genomes with >99% HapMap concordance.

MIPs represent a scalable technology previously applied to ~10,000-plex SNP genotyping[2]. Its adaptation to massively parallel exon capture demonstrated extensive multiplexing, straightforward scalability, high specificity, and low DNA input requirements[1]. However, two crippling deficiencies were encountered. First, only ~10,000 of 55,000 targeted exons were detectably amplified, with highly variable representation. Second, heterozygous alleles were not equally sampled, substantially impairing variant detection. Resolution of these deficiencies was a clear prerequisite for MIP-based exon capture to be broadly useful.

We modified the original protocol to enhance capture efficiency (Supplementary Methods online). 55,000 array-derived 100-mers were amplified and converted to single-stranded 70-mer MIPs as previously described[1], and again applied to 55,000-plex exon capture on genomic DNA (HapMap NA12248). Key changes included lengthening hybridization and gap-fill incubations, and increasing MIP and ligase concentrations. Post-capture PCR amplicons were linearly concatenated and processed to a standard shotgun sequencing library[3].

These modifications yielded a remarkable improvement in capture efficiency and allelic sampling while maintaining high specificity. Of 18 million uniquely mapped[4] reads obtained by Illumina sequencing, 99% overlapped with one of the 55,000 targets. 91% of targets were detectably captured (50,080 of 55,000), compared to 18% previously described1. Representational uniformity improved markedly as well (Fig. 1a). Allelic sampling of heterozygous variants showed a dramatic improvement as compared to previous data (Fig. 1b,c), now matching our expectation of a distribution converging to 0.5 with increasing coverage. An assessment of variant calling accuracy (including Sanger-based confirmation of a subset of novel variants) is provided in Supplementary Note 1 online.

Shotgun library construction remains a key bottleneck for scaling second-generation sequencing[5] to thousands of samples, as mechanical fragmentation and gel-based size-selection are challenging to automate. With longer read-lengths (e.g. Illumina GA-2; 76+ base-pairs), we realized direct sequencing of MIP-derived amplicons would enable targeted resequencing without shotgun library construction (Fig. 2).

Correspondence to: Jay Shendure (E-mail: shendure@u.washington.edu).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

To test this, gDNA from 16 individuals were subjected to targeted capture with 13,000 MIPs targeting 13,000 exons (Supplementary Data 1 online; subset of 55,000 with greater design constraints). Minor additional protocol changes were made (e.g. reducing input to 750 nanograms), but the primary change was the introduction of direct resequencing. Following capture, two multi-template PCRs (per individual) appended Illumina adaptors in either orientation. Amplicons from each individual were pooled 1:1 and directly sequenced with 76-bp single-end reads in one lane (Supplementary Fig. 1 online).

Performance of capture was highly consistent across the 16 individuals (Supplementary Table 1 online). On average, 8.4 million quality-filtered reads were collected per individual, of which 90.4% were confidently mapped[4]. The proportion of mapped reads was higher with direct sequencing than shotgun sequencing (90.4% vs. 56.8%), primarily because shotgun libraries are contaminated by the common MIP linker. Captures were highly specific with >99% of mapped reads aligning to one of 13,000 targets (Supplementary Fig. 2 online). Representational uniformity was further improved with ~98% of all targets captured per reaction, ~58% within a 10-fold, and ~88% within a 100-fold range (Fig. 1a), approaching the performance of array-based capture[6]. The relative capture efficiencies of individual MIPs were reproducible (average pairwise rank-order correlation coefficient of 0.92).

As the useful read-length for variant calling is 56 bases (Fig. 2), the maximum target length accessible with bi-directional sequencing was 112 bp. Because the lengths of the 13,000 targets ranged from 100 to 191 bp, we focused our analysis on accessible coding bases within targets (~1.4 of ~1.7 Mb), with the expectation that direct sequencing of longer capture products will be possible as read-lengths increase.

Performance of variant calling was highly reproducible across the 16 individuals (Supplementary Table 2 online). The overall concordance to HapMap genotypes was high for both homozygous (99.8%; n = 21,346) and heterozygous genotypes (99.3%; n = 3,622). While ~98% of targets were captured in each sample, only 75% of accessible target bases (~1.0 of ~1.4 Mb) were covered sufficiently for genotype calling. We estimate that 2x, 3x, and 4x increased sequence depth would increase the call-rate from 75% to 85%, 89%, and 91% respectively, i.e. with diminishing returns. Achieving greater completeness will likely require either further protocol optimizations or independent targeting of under-covered targets. Alternatively, we estimate that performing 2, 3 or 4 capture reactions and sequencing each at equivalent depth would increase the call-rate to 90%, 92%, and 94%, respectively. 54% of accessible target bases (~0.8 Mb of ~1.4 Mb) were sufficiently covered for variant calling in all 16 samples.

Variants (Supplementary Data 2 online; 593 per individual on average) were compared to dbSNP. The average proportion of annotated variants was higher for Yoruba (87%; n = 10) than Europeans (95%; n = 4) and Asians (94%, n = 2), consistent with greater diversity and poorer historical ascertainment in Africans. Across 16 individuals, 779 novel variants were identified. In contrast with previously annotated variants, most novel variants were observed just once across 32 chromosomes (Supplementary Fig. 3 online). Comparison of variant genotypes called here to genotypes generated independently by hybrid capture and sequencing (manuscript in preparation) validated 99% of all variants (n = 3703) and 94% (n = 303) of novel variants called in both data-sets, consistent with a low false discovery rate.

Our results establish the accuracy, reproducibility, and scalability of array-derived MIPs for massively parallel exon capture and resequencing. Our simplified protocol involving direct sequencing of MIP amplicons has important advantages over alternative methodologies for targeted capture: (a) concurrent targeting of at least 50,000 exons per reaction; (b) sub-microgram input DNA requirements; (c) a single synthesis of array-derived MIP precursors

QIAGEN Exhibit 1002
Page 789

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

can potentially support thousands of capture reactions; (d) small solution-based capture reactions that are highly scalable with no requirement for shotgun library construction at any stage.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported in part by grants from the National Institutes of Health / National Heart Lung and Blood Institute (R01 HL094976 to D.N. and J.S.) and the National Institutes of Health / National Human Genome Research Institute (R21 HG004749 to J.S.). E.T. is supported by a training fellowship from the National Institutes of Health / National Human Genome Research Institute (T32 HG00035). S.N. is supported by the Agency for Science, Technology and Research, Singapore. For technical assistance with sequencing, we thank K. Bertucci and S. da Ponte. For assistance with genotyping data, we thank J. Smith, G. Cooper, and I. Stanway. For helpful discussions, we thank H. Ji, J. Li, K. Zhang, G. Church, J. Teer, J. Egertson, and G. Porreca. For assistance with oligo libraries, we thank E. Leproust and W. Woo.

References

1. Porreca GJ, et al. Nat. Methods 2007;4:931–936. [PubMed: 17934468]
2. Hardenbol P, et al. Genome Res 2005;15:269–275. [PubMed: 15687290]
3. Bentley DR, et al. Nature 2008;456:53–59. [PubMed: 18987734]
4. Li H, Ruan J, Durbin R. Genome Res 2008;18:1851–1858. [PubMed: 18714091]
5. Shendure J, Ji H. Nat. Biotechnol 2008;26:1135–1145. [PubMed: 18846087]
6. Albert TJ, et al. Nat Methods 2007;4(11):903–905. [PubMed: 17934467]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

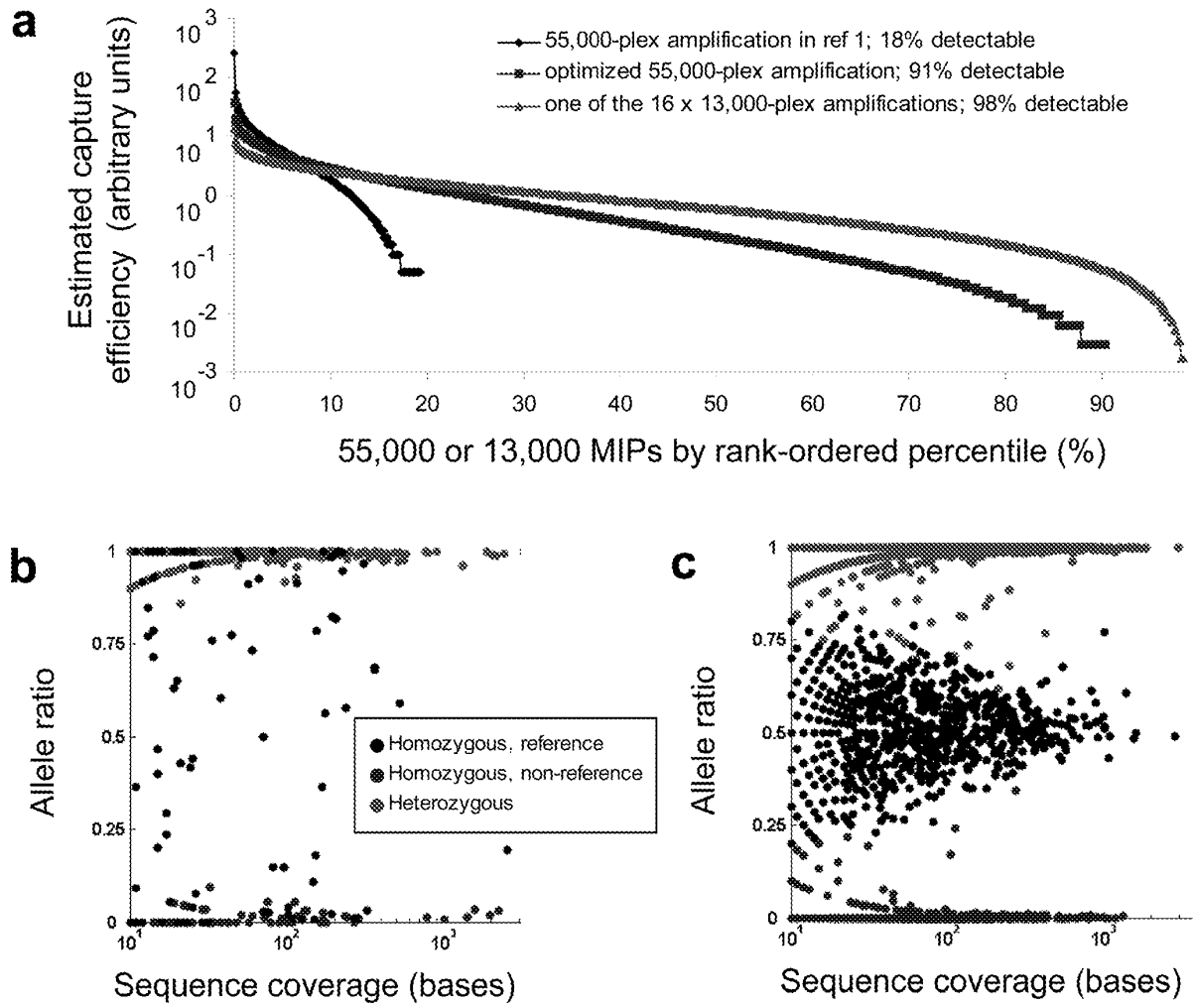


Figure 1. Optimizations yield reduced representational and allelic bias

(**a**) The uniformity of capture across 55,000 or 13,000 targeted exons is illustrated, estimated by the relative depth of sequence coverage of each target. (**b,c**) Sequencing-based allele ratios at positions of common variation are plotted for previously described[1] (**b**) and current (**c**) protocols. Each point represents a position where targeted resequencing data from a 55,000-plex amplification intersects with a HapMap genotype for this individual (NA12248). The allele ratio is the frequency with which the reference allele was observed in sequence data. Genotypes are indicated according to HapMap data.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

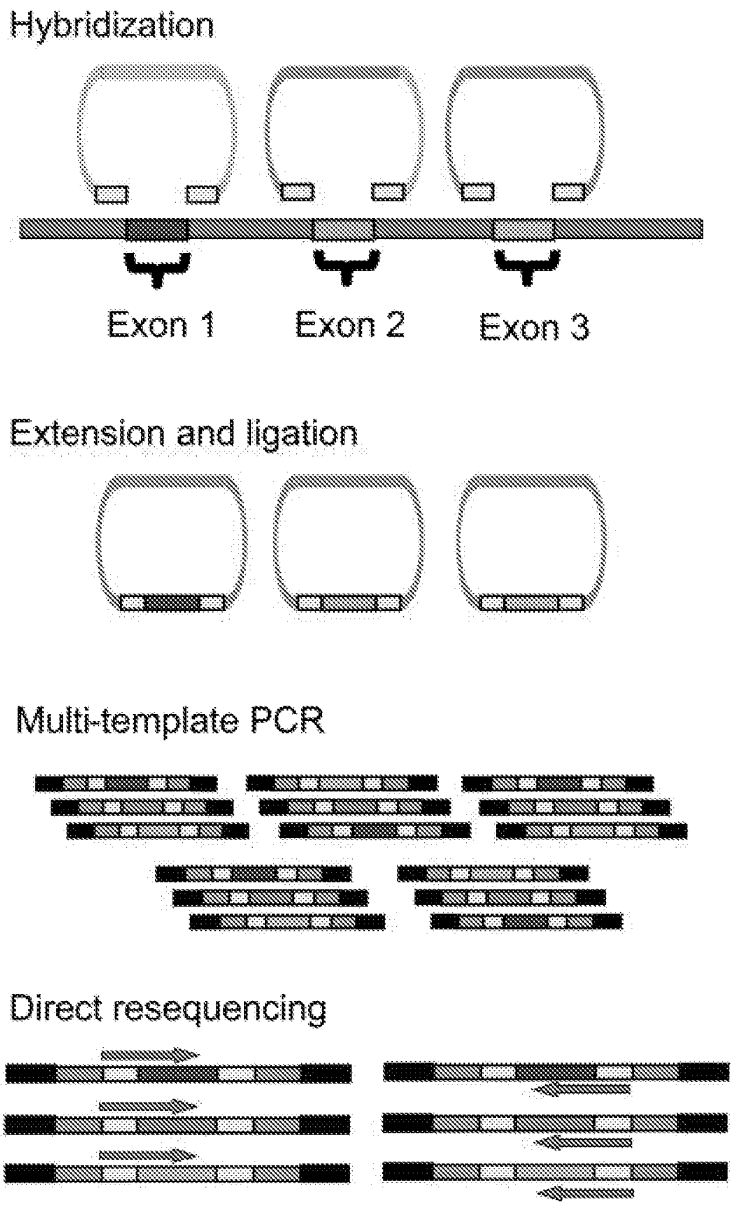


Figure 2. Schematic of MIP-based exon capture and direct resequencing

Thousands of MIP precursors can be cost-effectively produced by release from programmable arrays. Each MIP is an oligonucleotide that includes a common linker (green) flanked by target-specific sequences (yellow) that hybridize immediately upstream and downstream of its target. Addition of polymerase and ligase results in copying of targets and conversion to a circular form. Illumina adaptors (black) are appended by a multi-template, inverse PCR reaction, resulting in amplicons that can be directly sequenced. All 76 bp sequencing reads (grey arrows) include 20 bases of targeting arm sequence and an additional 56 bases of within-target sequence.

QIAGEN Exhibit 1002
Page 792

Electronic Acknowledgement Receipt

EFS ID:	16377918
Application Number:	13980987
International Application Number:	PCT/US12/22448
Confirmation Number:	8781
Title of Invention:	STEM-LOOP COMPOSITE RNA-DNA ADAPTOR-PRIMERS: COMPOSITIONS AND METHODS FOR LIBRARY GENERATION, AMPLIFICATION AND OTHER DOWNSTREAM MANIPULATIONS
First Named Inventor/Applicant Name:	Nurith KURN
Customer Number:	21971
Filer:	Jeffrey J. Seidel/Ashley Agacki
Filer Authorized By:	Jeffrey J. Seidel
Attorney Docket Number:	25115-747.831
Receipt Date:	22-JUL-2013
Filing Date:	
Time Stamp:	15:21:17
Application Type:	U.S. National Stage under 35 USC 371

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$1260
RAM confirmation Number	1994
Deposit Account	232415
Authorized User	

File Listing:

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

1	Transmittal of New Application	371.pdf	336079 fe517e9568e683f5ee7bee9542cf0498d9c7b7c1	no	3
Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	37211 ae7d08d570d36500460f64f858ae30e63621df99	no	2
Warnings:					
Information:					
			Total Files Size (in bytes):	373290	

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.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 August 2012 (02.08.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/103154 A1

(51) International Patent Classification:
C12P 19/34 (2006.01)

(21) International Application Number:
PCT/US2012/022448

(22) International Filing Date:
24 January 2012 (24.01.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/435,730 24 January 2011 (24.01.2011) US

(71) Applicant *(for all designated States except US)*: **NUGEN TECHNOLOGIES, INC.** [US/US]; 201 Industrial Road, Suite 310, San Carlos, CA 94070 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KURN, Nurith** [IL/US]; 2876 Ramona Street, Palo Alto, CA 94306 (US). **STAPLETON, Mark** [US/US]; 3170 Hambletonian Lane, Walnut Creek, CA 94598 (US). **WANG, Shenlong** [CA/US]; 8248 Saturn Park Drive, San Ramon, CA 94582 (US).

(74) Agents: **SKUBATCH, Maya** et al.; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, CA 94304-1050 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

WO 2012/103154 A1

(54) **Title**: STEM-LOOP COMPOSITE RNA-DNA ADAPTOR-PRIMERS: COMPOSITIONS AND METHODS FOR LIBRARY GENERATION, AMPLIFICATION AND OTHER DOWNSTREAM MANIPULATIONS

(57) **Abstract**: The present invention concerns manipulation of nucleic acid molecules, such as amplification, or adaptor-ligation. RNA/DNA chimeric stem-loop adaptor/primers are utilized during target nucleic acid manipulation. Methods controlling the activation of stem-loop oligonucleotides are within the scope of the invention. Methods combining multiple nucleic acid reactions in one reaction volume are envisioned in various embodiments of the invention.

Figure 1

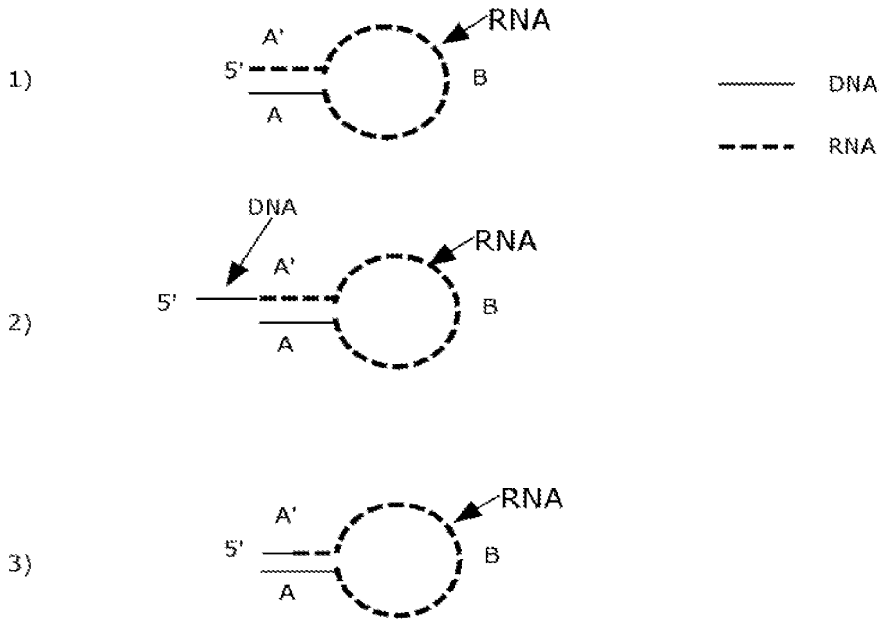


Figure 2a

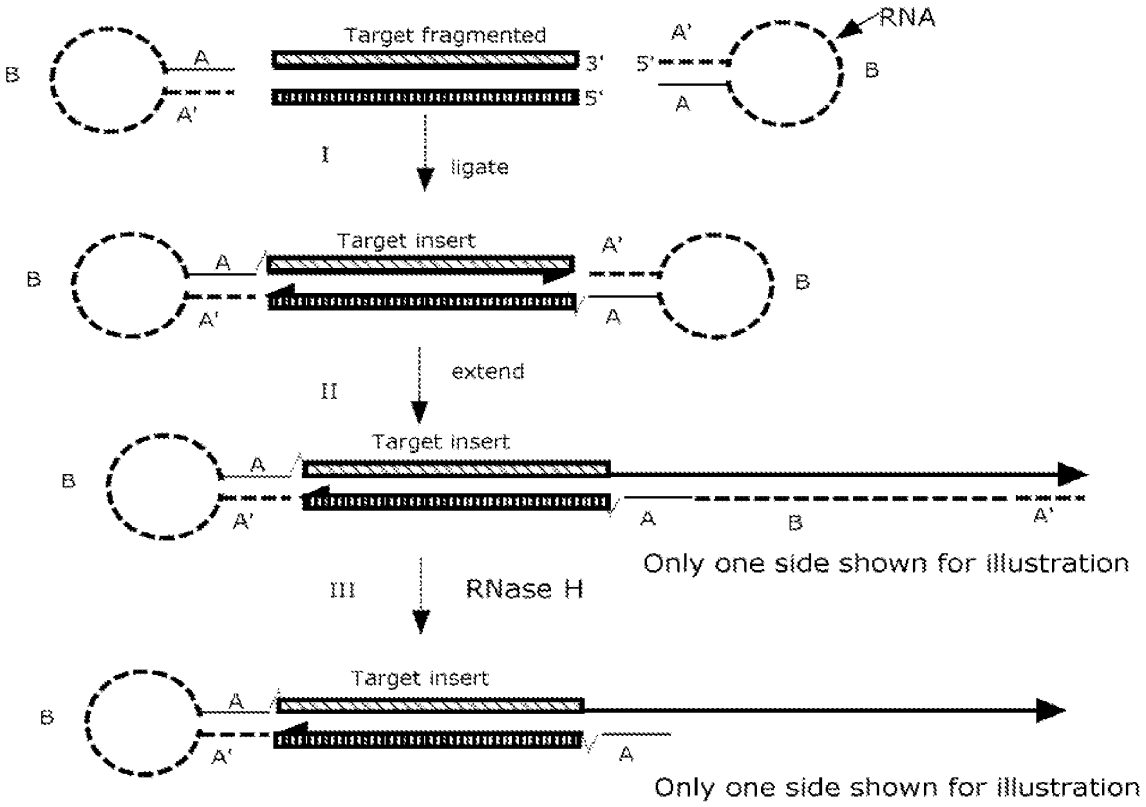

Target fragmented
3'
5'
A'
5'
RNA
B
A
A'
A
B
I
ligate
Target insert
A
A'
A'
A
B
B
II
extend
Target insert
A
A'
A
B
A'
Only one side shown for illustration
III
RNase H
Target insert
A
A'
A
Only one side shown for illustration

Figure 2b

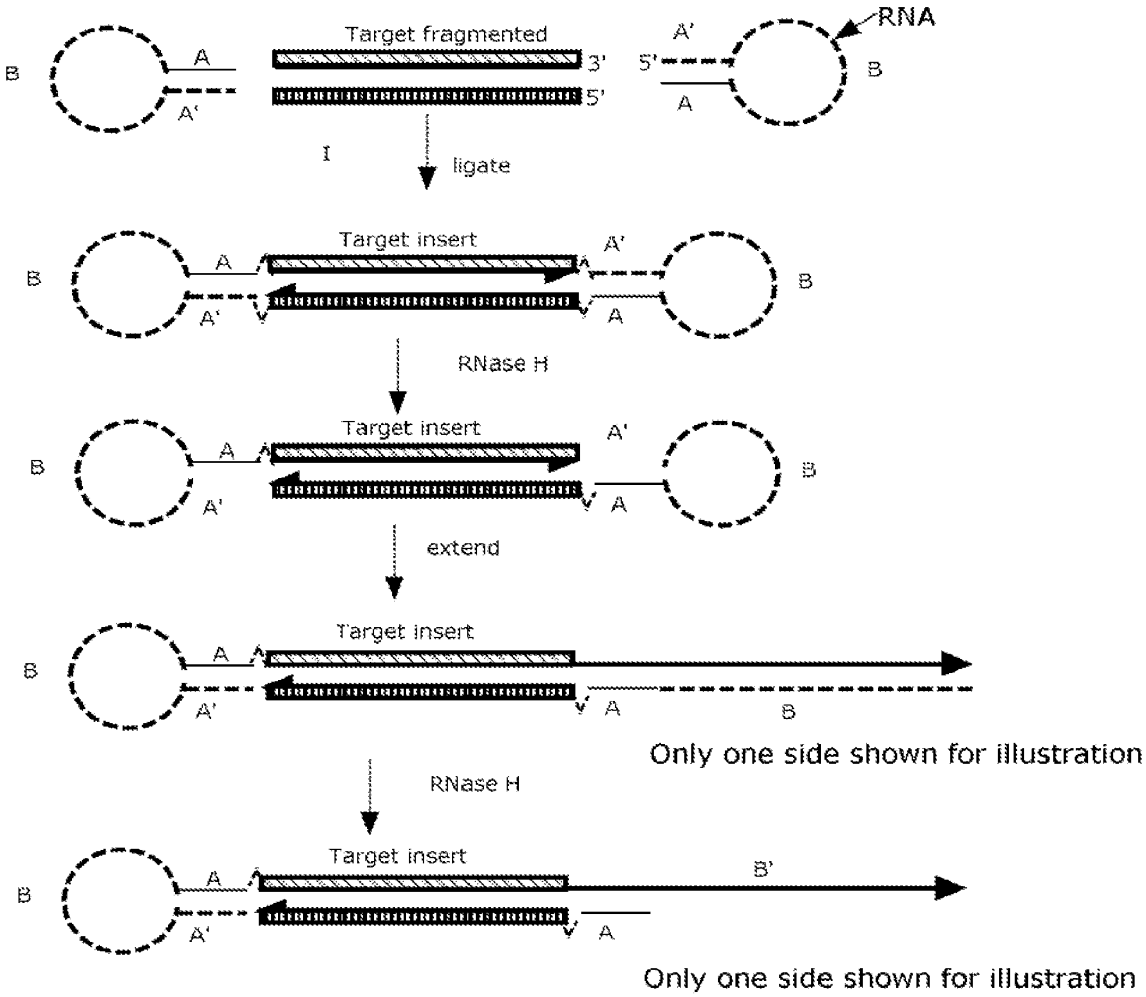

Target fragmented
RNA
A'
5'
B
A
3'
5'
A'
A
I
ligate
Target insert
A'
B
A
A'
A
B
RNase H
Target insert
A'
B
A
A'
A
B
extend
Target insert
B
A
A'
A
B
Only one side shown for illustration
RNase H
Target insert
B'
B
A
A'
A
Only one side shown for illustration

Figure 3

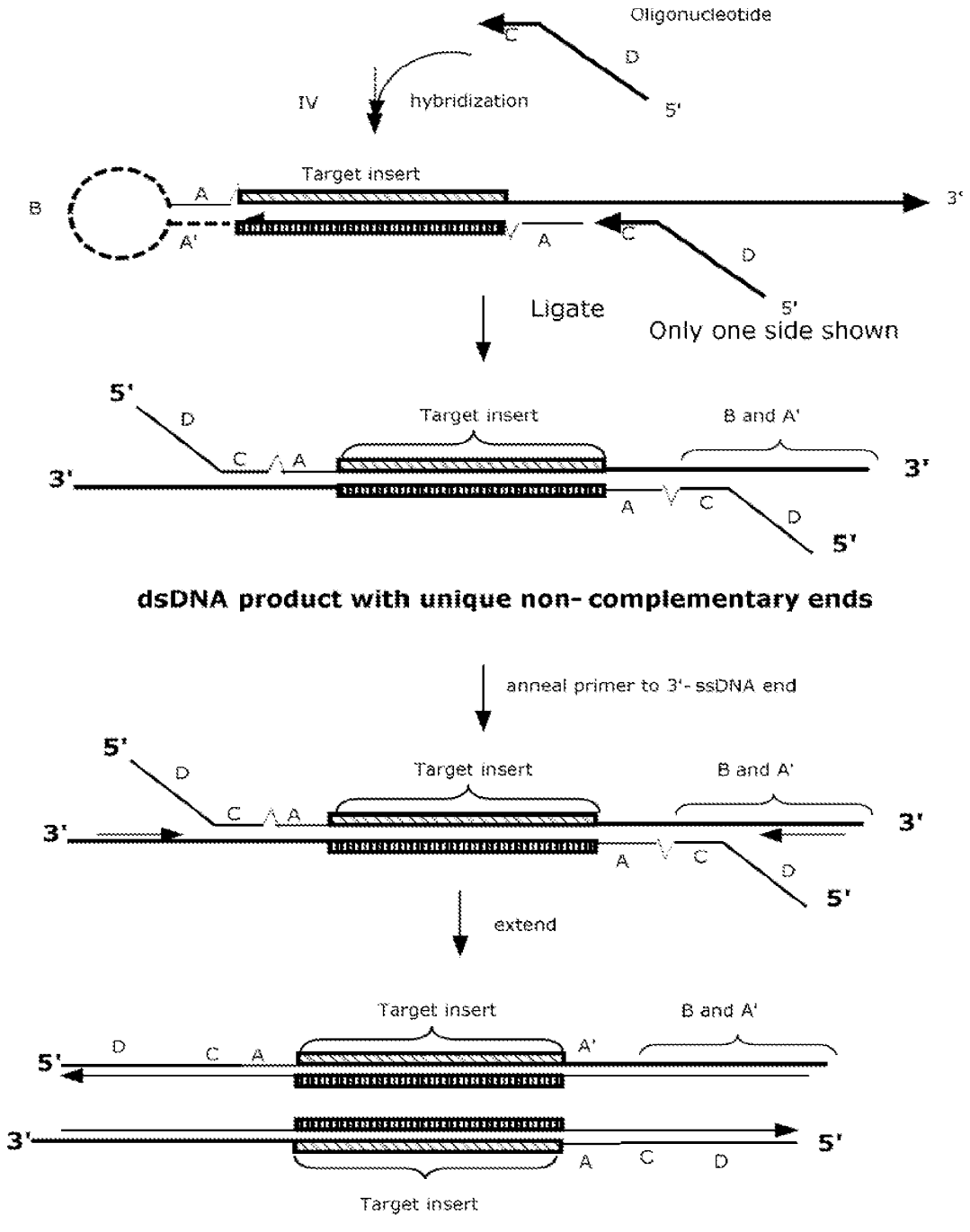

Oligonucleotide
C
D
5'
IV
hybridization
Target insert
B
A
A'
3'
A
C
D
5'
Ligate
Only one side shown
5'
D
C
A
Target insert
B and A'
3'
3'
A
C
D
5'
dsDNA product with unique non- complementary ends
anneal primer to 3'- ssDNA end
5'
D
C
A
Target insert
B and A'
3'
3'
A
C
D
5'
extend
Target insert
B and A'
D
C
A
A'
5'
3'
A
C
D
5'
Target insert

Figure 4a

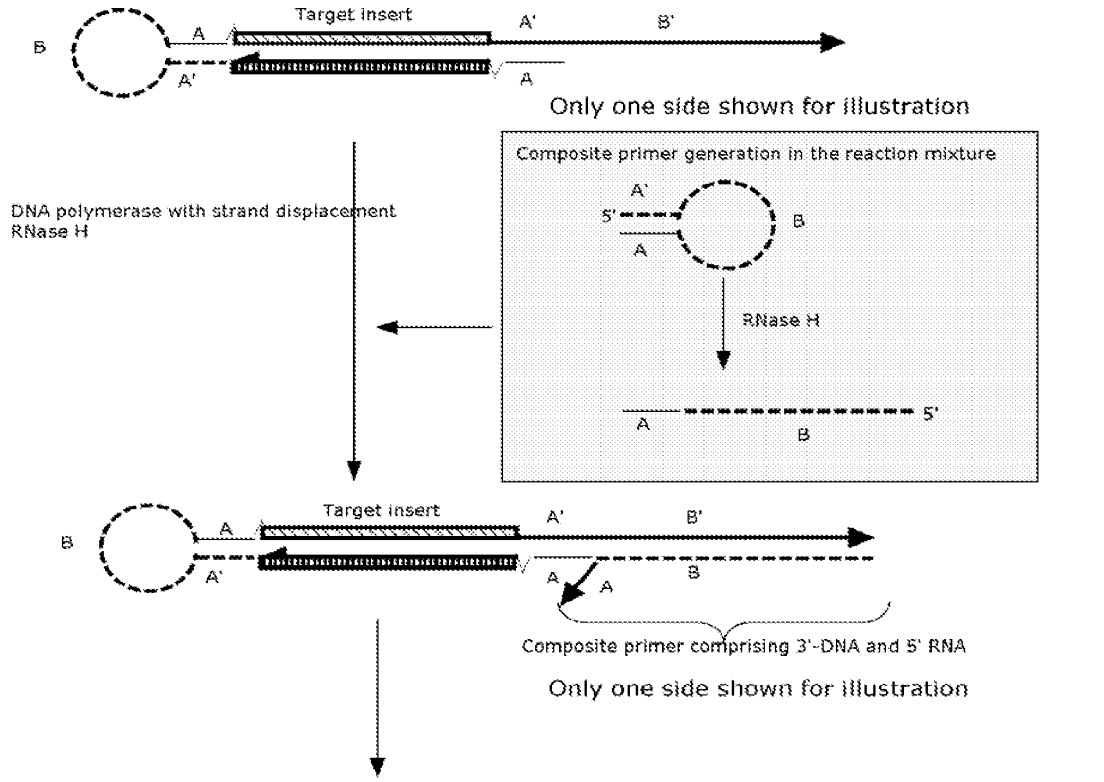

Target insert
A'
B'
B
A
A'
A
Only one side shown for illustration
Composite primer generation in the reaction mixture
A'
5'
B
A
RNase H
A
5'
B
DNA polymerase with strand displacement
RNase H
Target insert
A'
B'
B
A
A'
A
A
B
Composite primer comprising 3'-DNA and 5' RNA
Only one side shown for illustration
SPIA amplification
Both target insert strands are amplified
Generate copies of sense and antisense strands comprising sequence A or A' at the 5'-region

Figure 4b

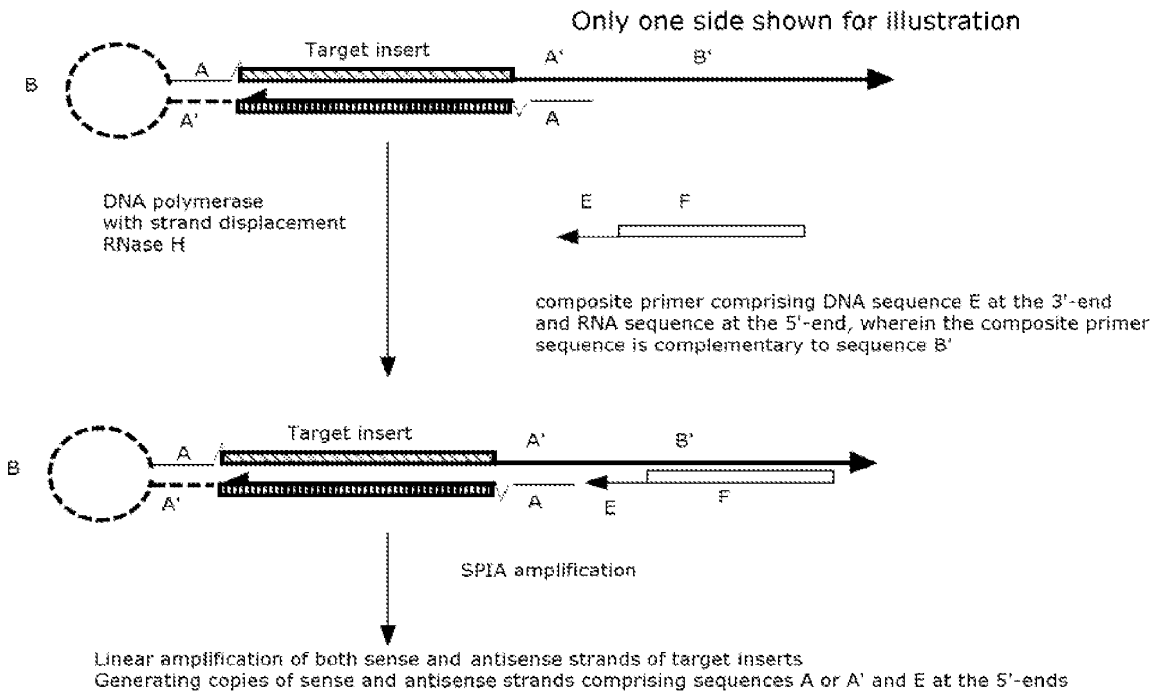

Only one side shown for illustration
Target insert
A
A'
B'
B
A'
A
DNA polymerase
with strand displacement
RNase H
E
F
composite primer comprising DNA sequence E at the 3'-end
and RNA sequence at the 5'-end, wherein the composite primer
sequence is complementary to sequence B'
Target insert
A
A'
B'
B
A'
A
E
F
SPIA amplification
Linear amplification of both sense and antisense strands of target inserts
Generating copies of sense and antisense strands comprising sequences A or A' and E at the 5'-ends

Figure 5

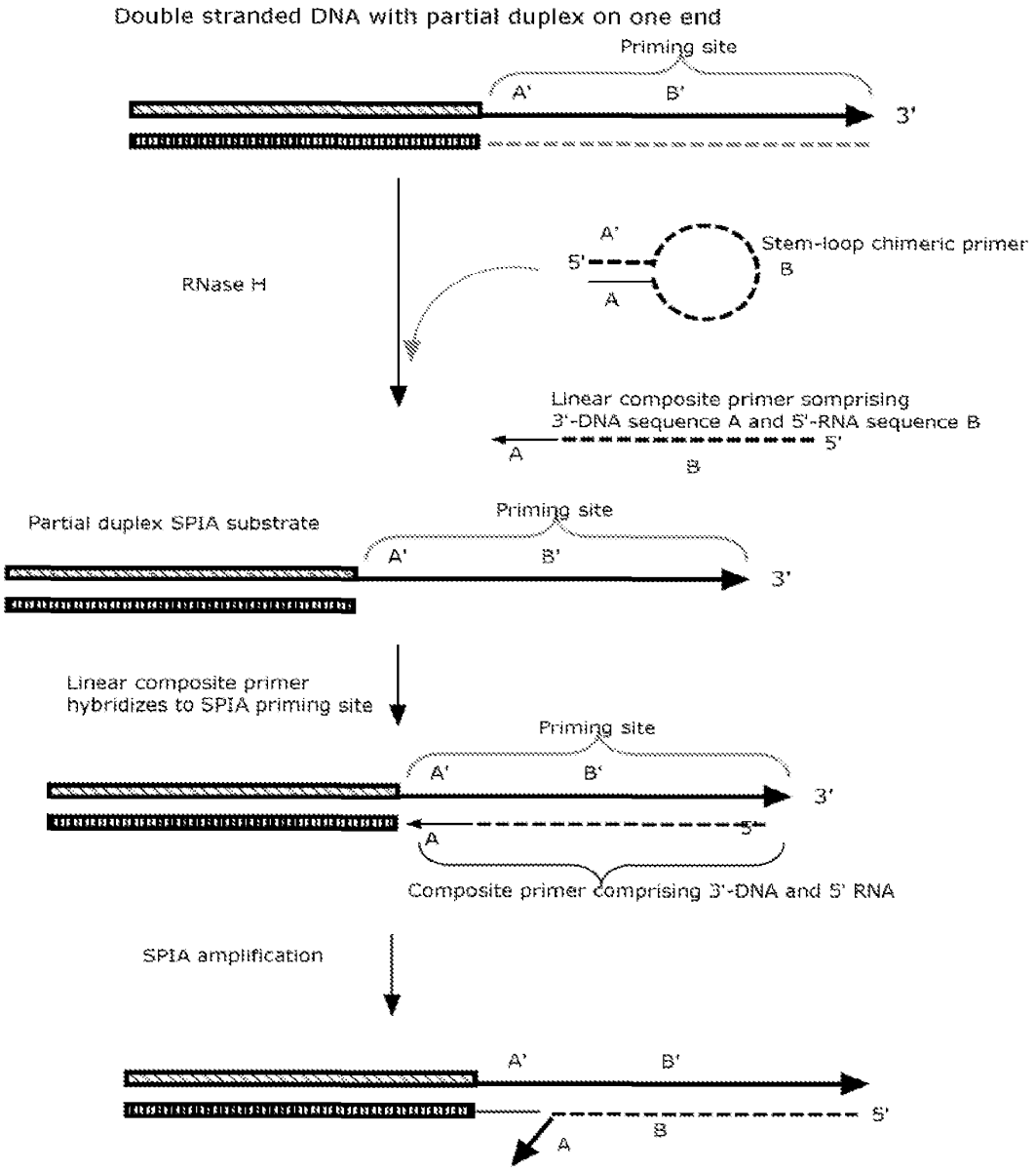

Double stranded DNA with partial duplex on one end
Priming site
A'
B'
3'
RNase H
A'
5'
Stem-loop chimeric primer
B
A
Linear composite primer somprising
3'-DNA sequence A and 5'-RNA sequence B
A
5'
B
Partial duplex SPIA substrate
Priming site
A'
B'
3'
Linear composite primer
hybridizes to SPIA priming site
Priming site
A'
B'
3'
A
5'
Composite primer comprising 3'-DNA and 5' RNA
SPIA amplification
A'
B'
5'
B
A

Figure 6

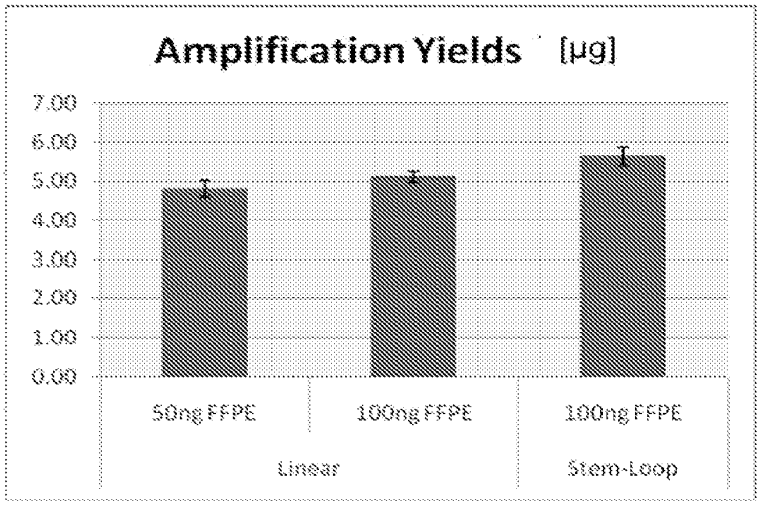
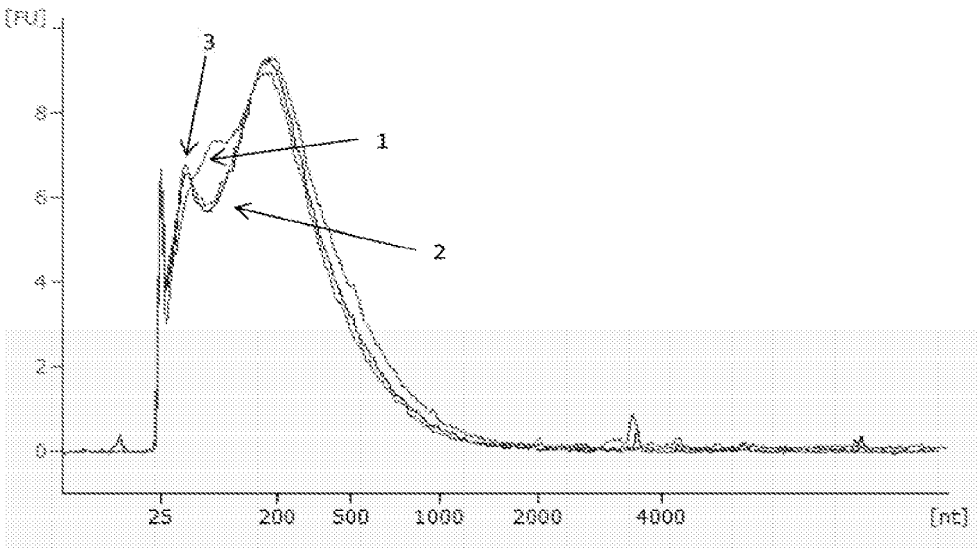

Amplification Yields [µg]
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
50ng FFPE
100ng FFPE
100ng FFPE
Linear
Stem-Loop

Figure 7



STEM-LOOP COMPOSITE RNA-DNA ADAPTOR-PRIMERS: COMPOSITIONS AND METHODS FOR LIBRARY GENERATION, AMPLIFICATION AND OTHER DOWNSTREAM MANIPULATIONS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 61/435,730, filed January 24, 2011, which application is incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0002] Advances in the study of biological molecules have been led, in part, by improvement in technologies used to characterizee the molecules or their biological reactions. In particular, the study of the nucleic acids DNA and RNA has benefited from developing technologies used for manipulating and analyzing nucleic acids, such as amplification, quantification, genotyping, and sequence analysis. Expanding use of nucleic acid based techniques demand employment of new methods and composition that exceed speed, efficiency, range and/or automation of current techniques. The instant application describes methods and compositions that improve currently available techniques for nucleic acid manipulation.

SUMMARY OF THE INVENTION

[0003] In one aspect, the invention provides methods of preparing a target nucleic acid molecule for further processing, comprising: attaching an RNA/DNA chimeric stem-loop oligonucleotide having a portion of its stem and at least a portion of its loop comprise of RNA to said target nucleic acid molecule thereby producing a stem-loop-target complex; and contacting said stem-loop-target complex with an RNase that cleaves RNA from a RNA/DNA heteroduplex, thereby preparing the stem-loop-target complex for further processing. In some embodiments, the RNase is RNase H. In some embodiments, further processing comprises nucleic acid sequencing. In some embodiments, sequencing comprises single molecule sequencing. In some embodiments, sequencing comprises sequencing by synthesis. In some embodiments, further processing comprises nucleic acid amplification. In some embodiments, a portion of the stem-loop oligonucleotide is further used to prime an amplification reaction. In some embodiments, the amplification is isothermal amplification. In some embodiments, the amplification is PCR. In some embodiments, the nucleic acid is produced by fragmenting a longer nucleic acid. In some embodiments, the target nucleic acid is blunt ended. In some embodiments, the target nucleic acid comprises a 3' overhang. In some embodiments, the target nucleic acid comprises a 5' overhang. In some embodiments, the stem-loop oligonucleotide further

comprises a 3' overhang. In some embodiments, the stem-loop oligonucleotide further comprises a 5' overhang. In some embodiments, the stem-loop oligonucleotide overhang and nucleic acid overhang are at least partially complementary to each other. In some embodiments, the attaching step comprises attaching a first RNA/DNA chimeric stem-loop oligonucleotide to a first end of said target nucleic acid and a second RNA/DNA chimeric stem-loop oligonucleotide to a second end of said target nucleic acid. In some embodiments, said first and said second RNA/DNA chimeric stem-loop oligonucleotides are the same. In some embodiments, the attaching step comprises the use of a ligase. In some embodiments, the ligase is T4 ligase. In some embodiments, the stem-loop oligonucleotide lacks 5' phosphate. In some embodiments, the target nucleic acid is further extended along the cleaved stem-loop oligonucleotide sequence, thereby forming a target nucleic acid extension product, wherein a portion of the stem-loop oligonucleotide acts as a template. In some embodiments, the stem-loop oligonucleotide that is in duplex with the nucleic acid extension is further cleaved using an RNase that cleaves RNA from a RNA/DNA heteroduplex, thereby forming a priming ready construct. In some embodiments, the priming ready construct is further amplified. In some embodiments, the amplification comprises single primer isothermal amplification (SPIA). In some embodiments, at least one loop of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group. In some embodiments, at least one stem of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group. In some embodiments, at least one overhang of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe

complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

[0004] In another aspect, the invention provides methods for amplifying a nucleic acid molecule, comprising: contacting in a single mixture an RNA/DNA stem-loop adaptor having a portion of its stem and at least a portion of its loop comprise of RNA; a double-stranded nucleic acid with a 3' overhang; RNase that cleaves RNA from a RNA/DNA heteroduplex; a DNA polymerase with strand displacement activity. Some embodiments, further comprise forming an extension product by priming the double-stranded nucleic acid at its 3' overhang using at least a portion of the stem-loop oligonucleotide. In some embodiments, the loop-RNA in the extension product is further cleaved by the RNase, thereby forming a second double stranded nucleic acid with at least one 3' overhang. In some embodiments, the loop-RNA in the extension product is further cleaved by the RNase, thereby forming a second double stranded nucleic acid with at least one 3' overhang. In some embodiments, the RNase comprises RNase H. Some embodiments further comprise at least a second amplification mixture with at least a second stem-loop oligonucleotide, wherein each stem-loop oligonucleotide further comprises a barcode sequence. In some embodiments, the 3' overhang is produced by forming a product, wherein the stem-loop oligonucleotide is attached to a double stranded nucleic acid by one strand; extending the double stranded nucleic acid along the stem-loop oligonucleotide; and cleaving the stem-loop oligonucleotide using an RNase that cleaves RNA sequences in an RNA/DNA heteroduplex. In some embodiments, the 3' overhang is produced during single primer isothermal amplification.

[0005] In yet another aspect, the invention provides kits for amplifying nucleic acids, comprising: a stem-loop oligonucleotide, wherein at least a portion of the stem comprises RNA and wherein at least a portion of the loop comprises RNA; an RNase that cleaves RNA sequences in an RNA/DNA duplex; a ligase; and a DNA polymerase.

[0006] In a further aspect, the invention provides methods of amplifying a double stranded target nucleic acid molecule, comprising: (a) attaching a strand of a first stem-loop oligonucleotide to said double stranded target nucleic acid molecule thereby producing an oligonucleotide-attached nucleic acid molecule; and (b) cleaving part of the stem-loop oligonucleotide using an RNase; and using a second stem-loop oligonucleotide as a primer in amplifying the double stranded target nucleic acid molecule. In some embodiments, the second stem-loop oligonucleotide is an exact copy of the first stem-loop oligonucleotide. In some embodiments, the RNase comprises RNase H. In some

embodiments, the stem-loop oligonucleotide comprises an RNA/DNA heteroduplex in a stem. In some embodiments, a strand of the oligonucleotide-attached nucleic acid molecule is extended. In some embodiments, the extension yields a second RNA/DNA heteroduplex. In some embodiments, cleaving part of the stem-loop oligonucleotide comprises cleaving the second RNA/DNA heteroduplex. In some embodiments, cleaving part of the stem-loop oligonucleotide generates a priming ready construct. In some embodiments, the priming ready construct comprises a single stranded region that is 3' to the double stranded target nucleic acid. In some embodiments, the amplification comprises PCR. In some embodiments, the amplification comprises SPIA. In some embodiments, the attaching comprises ligation. In some embodiments, T4 ligase is utilized in ligation. In some embodiments, the stem-loop oligonucleotide lacks 5' phosphate. In some embodiments, the target nucleic acid is produced by fragmenting a longer nucleic acid. In some embodiments, the target nucleic acid is blunt ended. In some embodiments, the target nucleic acid comprises a 3' overhang. In some embodiments, the target nucleic acid comprises a 5' overhang. In some embodiments, the stem-loop oligonucleotide further comprises a 3' overhang. In some embodiments, the stem-loop oligonucleotide further comprises a 5' overhang. In some embodiments, the stem-loop oligonucleotide overhang and nucleic acid overhang are at least partially complementary to each other. In some embodiments, at least one loop of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group. In some embodiments, at least one stem of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group. In some embodiments, at least one overhang of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe

complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

INCORPORATION BY REFERENCE

[0007] All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0009] Figure 1 illustrates examples of stem-loop oligonucleotides according to embodiments of the invention.

[0010] Figure 2 A-B illustrate two different methods of ligating stem-loop oligonucleotides to a double stranded target nucleic acid followed by nucleic acid extension and RNA cleavage generating priming ready constructs.

[0011] Figure 3 illustrates a method for appending disparate sequences to the two ends of a target nucleic acid according to the embodiments of the invention.

[0012] Figure 4 A-B illustrate methods for performing single primer isothermal amplification (SPIA) using stem-loop oligonucleotide (A) or linear (B) primers according to embodiments of the invention.

[0013] Figure 5 illustrates the use of stem-loop oligonucleotides as pro-primers and primers according to some embodiments of the invention.

[0014] Figure 6 illustrates in μg, the yields of whole genome amplification (WGA) reactions utilizing either the stem-loop chimeric amplification primer or the corresponding linear chimeric amplification primer.

[0015] Figure 7 illustrates the BioAnalyzer Profile of whole genome amplification (WGA) products demonstrating the florescent units as a function of size in nucleotides. Line 1, stem-loop primer with 100ng FFPE sample; line 2, linear primer with 50 ng FFPE sample; and line 3, linear primer with 100ng FFPE sample.

DEFINITIONS

[0016] The term "nucleic acid" refers to a nucleotide polymer, and unless otherwise limited, includes known analogs of natural nucleotides that can function in a similar manner (e.g., hybridize) to naturally occurring nucleotides.

[0017] The terms "polynucleotide", "nucleotide", "nucleotide sequence", "nucleic acid" and "oligonucleotide" are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three dimensional structure, and may perform any function, known or unknown. The following are non limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, intergenic DNA, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), small nucleolar RNA, ribozymes, complementary DNA (cDNA), which is a DNA representation of mRNA, usually obtained by reverse transcription of messenger RNA (mRNA) or by amplification; DNA molecules produced synthetically or by amplification, genomic DNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Polynucleotide sequences, when provided, are listed in the 5' to 3' direction, unless stated otherwise.

[0018] The term nucleic acid encompasses double- or triple-stranded nucleic acids, as well as single-stranded molecules. In double- or triple-stranded nucleic acids, the nucleic acid strands need not be coextensive (i.e., a double-stranded nucleic acid need not be double-stranded along the entire length of both strands).

[0019] The term nucleic acid also encompasses any chemical modification thereof, such as by methylation and/or by capping. Nucleic acid modifications can include addition of chemical groups that incorporate additional charge, polarizability, hydrogen bonding, electrostatic interaction, and functionality to the individual nucleic acid bases or to the nucleic acid as a whole. Such modifications may include base modifications such as 2'-position sugar modifications, 5-position pyrimidine modifications, 8-position purine

modifications, modifications at cytosine exocyclic amines, substitutions of 5-bromo-uracil, backbone modifications, unusual base pairing combinations such as the isobases isocytidine and isoguanidine, and the like.

[0020] More particularly, in certain embodiments, nucleic acids, can include polydeoxyribonucleotides (containing 2-deoxy-D-ribose), polyribonucleotides (containing D-ribose), and any other type of nucleic acid that is an N- or C-glycoside of a purine or pyrimidine base, as well as other polymers containing normucleotidic backbones, for example, polyamide (e.g., peptide nucleic acids (PNAs)) and polymorpholino (commercially available from the Anti-Virals, Inc., Corvallis, Oreg., as Neugene) polymers, and other synthetic sequence-specific nucleic acid polymers providing that the polymers contain nucleobases in a configuration which allows for base pairing and base stacking, such as is found in DNA and RNA. The term nucleic acid also encompasses linked nucleic acids (LNAs), which are described in U.S. Pat. Nos. 6,794,499, 6,670,461, 6,262,490, and 6,770,748, which are incorporated herein by reference in their entirety for their disclosure of LNAs.

[0021] The nucleic acid(s) can be derived from a completely chemical synthesis process, such as a solid phase-mediated chemical synthesis, from a biological source, such as through isolation from any species that produces nucleic acid, or from processes that involve the manipulation of nucleic acids by molecular biology tools, such as DNA replication, PCR amplification, reverse transcription, or from a combination of those processes.

[0022] As used herein, the term "target polynucleotide" or "target nucleic acid" refer to a nucleic acid molecule or polynucleotide in a starting population of nucleic acid molecules having a target sequence whose presence, amount, and/or nucleotide sequence, or changes in these, are desired to be determined. In general, a target polynucleotide is a double-stranded nucleic acid molecule, and may be derived from any source of or process for generating double-stranded nucleic acid molecules.

[0023] As used herein, the term "target sequence" refers generally to a nucleic acid sequence on a single strand of nucleic acid. The target sequence may be a portion of a gene, a regulatory sequence, genomic DNA, cDNA, RNA including mRNA, miRNA, and rRNA, or others. The target sequence may be a target sequence from a sample or a secondary target such as a product of an amplification reaction. The "target sequence" can be embedded in a molecule that includes the nucleotide sequence of a target nucleic acid, such as, for example, the amplification product obtained by amplifying a target nucleic acid or the cDNA produced upon reverse transcription of an RNA target nucleic acid.

[0024] A "nucleotide probe" or "probe" refers to a polynucleotide used for detecting or identifying its corresponding target polynucleotide in a hybridization reaction. Thus, a probe is hybridizable to one or more target polynucleotides. Probes can be perfectly complementary to one or more target polynucleotides in a sample, or contain one or more nucleotides that are not complemented by a corresponding nucleotide in the one or more target polynucleotides in a sample.

[0025] As used herein, the term "complementary" refers to the capacity for precise pairing between two nucleotides. I.e., if a nucleotide at a given position of a nucleic acid is capable of hydrogen bonding with a nucleotide of another nucleic acid, then the two nucleic acids are considered to be complementary to one another at that position. Complementarity between two single-stranded nucleic acid molecules may be "partial," in which only some of the nucleotides bind, or it may be complete when total complementarity exists between the single-stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands.

[0026] "Hybridization" and "annealing" refer to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson Crick base pairing, Hoogstein binding, or in any other sequence specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi stranded complex, a single self hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR or other amplification reactions, or the enzymatic cleavage of a polynucleotide by a ribozyme. A first sequence that can be stabilized via hydrogen bonding with the bases of the nucleotide residues of a second sequence is said to be "hybridizable" to said second sequence. In such a case, the second sequence can also be said to be hybridizable to the first sequence.

[0027] In general, a "complement" of a given sequence is a sequence that is fully or substantially complementary to and hybridizable to the given sequence. In general, a first sequence that is hybridizable to a second sequence or set of second sequences is specifically or selectively hybridizable to the second sequence or set of second sequences, such that hybridization to the second sequence or set of second sequences is preferred (e.g. thermodynamically more stable under a given set of conditions, such as stringent conditions commonly used in the art) to hybridization with non-target sequences during a

hybridization reaction. Typically, hybridizable sequences share a degree of sequence complementarity over all or a portion of their respective lengths, such as between 25%-100% complementarity, including at least about 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, and 100% sequence complementarity.

[0028] The term "hybridized" as applied to a polynucleotide refers to a polynucleotide in a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson Crick base pairing, Hoogstein binding, or in any other sequence specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi stranded complex, a single self hybridizing strand, or any combination of these. The hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme. A sequence hybridized with a given sequence is referred to as the "complement" of the given sequence.

[0029] "Specific hybridization" refers to the binding of a nucleic acid to a target nucleotide sequence in the absence of substantial binding to other nucleotide sequences present in the hybridization mixture under defined stringency conditions. Those of skill in the art recognize that relaxing the stringency of the hybridization conditions allows sequence mismatches to be tolerated.

[0030] As used herein, the term "complementary" refers to the capacity for precise pairing between two nucleotides. I.e., if a nucleotide at a given position of a nucleic acid is capable of hydrogen bonding with a nucleotide of another nucleic acid, then the two nucleic acids are considered to be complementary to one another at that position. Complementarity between two single-stranded nucleic acid molecules may be "partial," in which only some of the nucleotides bind, or it may be complete when total complementarity exists between the single-stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands.

[0031] The term "oligonucleotide" is generally used to refer to a nucleic acid that is relatively short, generally shorter than 200 nucleotides, more particularly, shorter than 100 nucleotides, most particularly, shorter than 50 nucleotides. Typically, oligonucleotides are single-stranded DNA molecules.

[0032] The term "primer" refers to an oligonucleotide that is capable of hybridizing (also termed "annealing") with a nucleic acid and serving as an initiation site for nucleotide (RNA or DNA) polymerization under appropriate conditions (i.e., in the presence of four different nucleoside triphosphates and an agent for polymerization, such as DNA or RNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. The appropriate length of a primer depends on the intended use of the primer, but primers are typically at least 7 nucleotides long and, more typically range from 10 to 30 nucleotides, or even more typically from 15 to 30 nucleotides, in length. Other primers can be somewhat longer, e.g., 30 to 50 nucleotides long. In this context, "primer length" refers to the portion of an oligonucleotide or nucleic acid that hybridizes to a complementary "target" sequence and primes nucleotide synthesis. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template but must be sufficiently complementary to hybridize with a template. The term "primer site" or "primer binding site" refers to the segment of the target nucleic acid to which a primer hybridizes. A construct with presenting a primer binding site is often referred to as a "priming ready construct" or "amplification ready construct".

[0033] A primer is said to anneal to another nucleic acid if the primer, or a portion thereof, hybridizes to a nucleotide sequence within the nucleic acid. The statement that a primer hybridizes to a particular nucleotide sequence is not intended to imply that the primer hybridizes either completely or exclusively to that nucleotide sequence.

[0034] As used herein, "expression" refers to the process by which a polynucleotide is transcribed into mRNA and/or the process by which the transcribed mRNA (also referred to as "transcript") is subsequently being translated into peptides, polypeptides, or proteins. The transcripts and the encoded polypeptides are collectively referred to as "gene product." If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell.

DETAILED DESCRIPTION OF THE INVENTION

[0035] The practice of the present invention employs, unless otherwise indicated, conventional techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics and recombinant DNA, which are within the skill of the art. See Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR

BIOLOGY (F. M. Ausubel, et al. eds., (1987)); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R.I. Freshney, ed. (1987)).

[0036] The present invention relates to methods and compositions relating to the use of stem-loop oligonucleotides as adaptors and/or primers. In various embodiments, the stem-loop oligonucleotides are chimeras of RNA and DNA. In some embodiments, the stem-loop oligonucleotide is activated through the cleavage of an RNA/DNA heteroduplex stem (e.g. by degradation of RNA by RNase H). In some embodiments, the stem-loop oligonucleotide comprises an RNA sequence in at least a portion of its loop and/or in its stem. Methods of replication and amplification using the stem-loop oligonucleotides described herein can also generate RNA/DNA heteroduplexes. Further, methods and compositions in various embodiments of the invention relate to the use of a stem-loop oligonucleotide as a dual purpose agent acting as an adaptor and primer at once.

[0037] In some embodiments, the present invention allows for the amplification of molecules having at least one double stranded region by using adaptors that reduce the propensity to form adaptor dimers. In certain aspects, the present invention provides an inert oligonucleotide (e.g. a stem-loop oligonucleotide) for attachment to a double stranded molecule such that it renders the oligonucleotide-ligated molecule capable of being modified, such as amplified, for example by strand displacement, single primer isothermal amplification (SPIA) or polymerase chain reaction. For example, upon cleavage of the inert adaptor, the oligonucleotide becomes active and suitable for providing at least in part one or more sequences employable for adaptor dependent nucleic acid extension or amplification. For another example, during SPIA amplification or polymerase chain reaction, non-cleaved adaptor/primer (e.g. stem-loop composite oligonucleotide) cannot prime effectively due to duplex formation along its stem. Its use as a primer (before it is cleaved) can be mildly or strictly limited due to the internal stem formation competing with target hybridization.

[0038] In addition to the advantages provided by the adaptors/primers of the invention, the invention further provides novel conditions for modification of nucleic acid molecules with the composite stem-loop adaptors, and subsequent amplification with said activated oligonucleotides. In most cases in the art, the process of nucleic acid analysis involves multiple enzymatic reactions that are performed in a sequential manner, frequently with

intermediate purification steps between at least some of the reactions. For example, preparation of a library from genomic or viral DNA may involve fragmentation, purification, "polishing" (end-repair) of fragments, a second purification or heat inactivation, adaptor ligation and subsequent extension reactions in separate steps that often need a considerable amount of time, equipment use and human intervention, creating major obstacles for high throughput and diagnostic-type applications. The process becomes more complicated when library preparation involves additional enzymatic steps, such as DNA or library digestion with methylation-sensitive or methylation-specific endonucleases (for example, in the preparation of Methylome libraries or restriction nuclease cleavage within the adaptor sequence to produce sticky ends for DNA cloning.

[0039] In specific aspects, the present invention introduces a concept of multiplexing two or more enzymatic processes in one reaction, and teaches how to optimize a highly multiplexed enzymatic process. In particular embodiments, the present invention is directed to compositions and methods for simultaneous processing of DNA molecules with a combination of enzymes in a one-step-one-tube reaction and producing either a collection of molecules suitable for further processing, for example for amplification, or in some cases, the methods and compositions of the invention result in amplified DNA molecules. In addition, the methods and compositions of the present invention provides composite adaptor/primer oligonucleotides and methods of using the same that eliminate the possible formation of stable adaptor dimers, thus eliminating potential difficulties regularly introduced by other commonly employed methods and reagents. In particular embodiments, the methods of the invention can be easily applied to any type of fragmented double stranded DNA including but not limited to, for example, free DNA isolated from plasma, serum, and/or urine; apoptotic DNA from cells and/or tissues; DNA fragmented enzymatically in vitro (for example, by DNase I and/or restriction endonuclease); and/or DNA fragmented by mechanical forces (hydro-shear, sonication, nebulization, etc.). Additional suitable methods and compositions of producing nucleic acid molecules comprising stem-loop oligonucleotides are further described in detail in US Patent No. 7,803,550, which is herein incorporated by reference in its entirety.

[0040] In other embodiments, the invention can be easily applied to any high molecular weight double stranded DNA including, for example, DNA isolated from tissues, cell culture, bodily fluids, animal tissue, plant, bacteria, fungi, viruses, etc.

[0041] Disclosed are compositions and methods for the generation of libraries using chimeric RNA-DNA stem loop oligonucleotides from any nucleic acid molecules. In various embodiments, the chimeric stem-loop oligonucleotides comprise heteroduplex stems. The stem-loop oligonucleotides described herein can be used as adaptors, primers or both. In various embodiments, the nucleic acid molecules are double stranded DNA molecules. In some embodiments, the unique structure of the stem loop chimeric oligonucleotides provide for the generation of partial duplex DNA from fragments of the sample dsDNA comprising a dsDNA fragment insert and 3'-ssDNA at one or both ends, which comprise a sequence complementary to a sequence of the stem-loop composite chimeric oligonucleotide.

[0042] The partial duplex product thus generated can be a substrate for amplification, for example, PCR or isothermal amplification. In some embodiments, the amplification reaction employs the same chimeric RNA-DNA stem-loop oligonucleotide or a different stem-loop oligonucleotide as a primer. For example, an isothermal amplification reaction comprising a chimeric RNA-DNA stem-loop oligonucleotide or a portion thereof, DNA polymerase with strand displacement, and RNase H can be performed for nucleic acid amplification using as substrate said partial duplex product.

[0043] In some embodiments, the disclosed stem-loop chimeric adaptor-pro-primer, is used as a primer for amplification, for example isothermal amplification of nucleic acids. Such pro-primer molecules can be activated to a primer in the reaction mixture. Primer activation can be carried out by an RNase targeting RNA in a RNA/DNA duplex, such as, RNase H. The RNase can be included as a component of the amplification reaction mixture. The concentration of the pro-primer, the amount of RNase activity and other reaction conditions can be adjusted to achieve activation of the pro-primer at a desired rate. For example, the pro-primer can be activated in the same reaction mixture, at a rate that is comparable to the overall duration of the amplification reaction allowing for gradual activation of stem-loop oligonucleotides over time. The pro-primer chimeric RNA-DNA stem-loop oligonucleotide can be useful for the isothermal amplification, for example SPIA, of nucleic acids employing any other method for the generation of substrates for SPIA amplification such as described for RNA amplification or whole genome amplification, as described in US. Patent No.s 6,251,639 and 6,946,251, which are both herein incorporated by reference in their entirety. Primer activation from pro-primer stem-loop chimeric oligonucleotide serves as a "slow release" of an active primer in the reaction. In general, formation of side products, for example due to high primer concentration in the initial period of an amplification reaction, can be mitigated leading

to increase in desired amplification products (specificity) and higher yield, for example, formation of stem-loop oligonucleotide dimers, which are formed from the attachment of two stem-loop oligonucleotides, can be targeted by RNase activity cleaving the RNA in the RNA/DNA duplex of the stem, thus releasing remaining portions of the two oligonucleotides from each other.

Stem-loop chimeric RNA-DNA adaptor/pro-primer:

[0044] In various embodiments, a chimeric/composite stem-loop oligonucleotide comprises any oligonucleotide having a sequence that can be joined to a target polynucleotide. In many embodiments, at least a portion of the oligonucleotide sequence is known. Stem-loop oligonucleotides can comprise DNA, RNA, nucleotide analogues, non-canonical nucleotides, labeled nucleotides, modified nucleotides, or combinations thereof. Stem-loop oligonucleotides can comprise one or more of single-stranded and double-stranded regions, giving rise to stems, loops and overhangs. For example, a stem-loop oligonucleotide can have one or more loops, stems and optionally one or more overhangs. In a single strand stem-loop oligonucleotide, multiple overhangs can be achieved by lack of complementarity in the 5' and 3' distal regions. Any one of stems, loops, and overhangs can comprise DNA or RNA sequences. In general, a partial-duplex oligonucleotide comprises one or more single-stranded regions and one or more double-stranded regions. Double-stranded compositions can comprise two separate oligonucleotides hybridized to one another (also referred to as an "oligonucleotide duplex"), and hybridization may leave one or more blunt ends, one or more 3' overhangs, one or more 5' overhangs, one or more bulges resulting from mismatched and/or unpaired nucleotides, or any combination of these. In some embodiments, a single-stranded adaptor comprises two or more sequences that are able to hybridize with one another. When two such hybridizable sequences are contained in a single-stranded oligonucleotide, hybridization yields a hairpin structure or a stem-loop structure. When two hybridized regions of an oligonucleotide are separated from one another by a non-hybridized region, a "bubble" structure results. Oligonucleotide comprising a bubble structure can consist of a single strand comprising internal hybridizations, or may comprise two or more strands hybridized to one another. Internal sequence hybridization, such as between two hybridizable sequences in a strand, can produce a double-stranded structure in a single-stranded adaptor oligonucleotide. Oligonucleotides of different kinds can be used in combination, such as a hairpin oligonucleotide, stem-

loop oligonucleotide and a double-stranded oligonucleotide, or oligonucleotides of different sequences. Hybridizable sequences of the adaptors and primers described herein may or may not include one or both ends of the oligonucleotide. When neither of the ends are included in the hybridizable sequences, both ends are "free" or "overhanging." When only one end is hybridizable to another sequence in the adaptor, the other end forms an overhang, such as a 3' overhang or a 5' overhang. When both the 5'-terminal nucleotide and the 3'-terminal nucleotide are included in the hybridizable sequences, such that the 5'-terminal nucleotide and the 3'-terminal nucleotide are complementary and hybridize with one another, the end is referred to as "blunt." Different adaptors can be joined to target polynucleotides in sequential reactions or simultaneously. For example, the first and second adaptors can be added to the same reaction. Adaptors can be manipulated prior to combining with target polynucleotides. For example, terminal phosphates can be added or removed.

[0045] Stem-loop oligonucleotides, according to the various embodiments of the invention vary in size. Stem-loop oligonucleotides are typically at least 10 nucleotides long and, more typically range from 14 to 60 nucleotides, or even more typically from 18, 21, 24, 27 to 40, 45 or 50 nucleotides, in length. Other stem-loop oligonucleotides can be somewhat longer, e.g., 50 to 80 nucleotides long or longer. Stem-loop oligonucleotides can have any of the length ranges capped by any of the limits above, e.g. 18 to 60 nucleotides.

[0046] In some embodiments, one of the hybridizable sequences in a single-stranded stem-loop structure comprises RNA. For example, a stem-loop oligonucleotide can comprise a 5' end comprising sequence A' and a 3' end comprising sequence A, where A is hybridizable to A', one of A or A' comprises DNA, and the other of A or A' comprises RNA. Examples of such stem-loop oligonucleotides are illustrated in Figure 1. Similarly, a stem-loop oligonucleotide can comprise a 5' end comprising sequence B and a 3' end comprising sequence B', where B is hybridizable to B', one of B or B' comprises DNA, and the other of B or B' comprises RNA. In some embodiments, one of A or A' consists entirely of DNA, and/or one of A or A' consists entirely of RNA. In some embodiment, one of B or B' consists entirely of DNA, and/or one of B or B' consists entirely of RNA. Sequence A can be the same as or different from sequence B and/or B'. Sequence A' can be the same as or different from sequence B and/or B'. In some embodiments, the stem comprising RNA (e.g. A, A', B, or B') further comprises one or more terminal DNA residues (e.g. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or more terminal DNA residues), such that the sequence comprising RNA is flanked by DNA residues at both

ends (i.e. both the 5' end and the 3' end of the sequence comprising RNA). Hybridization of a sequence comprising RNA to a sequence comprising DNA creates an RNA-DNA heteroduplex. In some embodiments, RNA is cleaved by an enzyme that cleaves RNA from an RNA-DNA heteroduplex, such as enzymes comprising ribonuclease activity. Preferably, the enzyme comprising ribonuclease activity cleaves ribonucleotides in an RNA/DNA heteroduplex regardless of the identity and type of nucleotides adjacent to the ribonucleotide to be cleaved. It is preferred that the ribonuclease cleaves independent of sequence identity. Examples of suitable enzymes comprising ribonuclease activity for the methods and compositions of the invention are well known in the art, including ribonuclease H (RNase H) and enzymes comprising RNase H activity, e.g., Hybridase. In some embodiments, cleavage of RNA from an RNA-DNA heteroduplex removes all double-stranded character from a single-stranded stem-loop oligonucleotide, such that extension by a polymerase that uses the oligonucleotide as template requires no strand displacement step or strand displacement activity. In some embodiments, one or both ends of a stem-loop oligonucleotide comprising RNA in its stem are joined to a target polynucleotide, such that cleavage of the RNA from the RNA-DNA hetero duplex produces a 5' overhang or a 3' overhang. In some embodiments, a stem-loop oligonucleotide comprises a tag/barcode sequence. Reactions of the stem-loop oligonucleotide with a target nucleic acid, such as ligation, extension, priming etc., can impart the target oligonucleotide tagged with the barcode sequence. Blunt or sticky ended stem-loop oligonucleotides can be joined with blunt or sticky ended target polynucleotides. In some embodiments, an end comprising a 5' overhang produced by cleavage of RNA from an RNA-DNA heteroduplex is filled in by the extension of the produced 3' end using the 5' overhang as template.

[0047] Various exemplary configurations of the stem-loop chimeric RNA-DNA adaptor/pro-primer are shown in Figure 1. In some embodiments, the stem-loop chimeric RNA-DNA oligonucleotides are composed of sequences that are complementary at their distal ends (A and A') and form a DNA-RNA heteroduplex stem. In various embodiments, the stem-loop structures comprise a single stranded loop region (B). All or part of the loop region can comprise RNA (loop-RNA). The complement of the DNA sequence at the 3'-region as well as the RNA loop sequence can be removed exposing a priming site for nucleic acid amplification. In some embodiments, an isothermal amplification comprising an RNA-DNA chimeric primer utilizes said priming site. In some embodiments the RNA/DNA chimeric primer is a stem-loop chimeric pro-primer. In

some embodiments the stem-loop chimeric adaptor and the stem-loop chimeric pro-primer are the same.

[0048] In some embodiments, the 3' or 5'-region of the chimeric RNA-DNA stem loop oligonucleotide of the current invention comprises a 3' or 5'-single stranded overhang, comprising DNA or RNA or combination of RNA and DNA. The 3' or 5'-single stranded overhangs may be utilized for hybridization to sticky-ended template fragments, such as those generated by fragmentation of the template dsDNA by restriction enzymes, for example, or when fragments are tailed by a single nucleotide, such as for example in the case of AT cloning procedures as described in detail in US Patent Pub. US 20100167954 and Marchuk et al. (Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Marchuk et al., *Nucleic Acids Res.* 1991 March 11; 19(5):1154), which are both incorporated herein in their entirety. In these cases, the overhang sequence can be designed to be sufficiently complementary to a 3' or 5'-overhang on a template fragment allowing hybridization under the reaction conditions.

[0049] The sequence in the 3' or 5'-region comprising the stem of the stem-loop chimeric oligonucleotide of the invention may be all RNA or may comprise RNA and DNA nucleotides. The length of the RNA sequence in the stem region can be designed to optimize the rate of cleavage of the RNA portion by an RNase H that cleaves the RNA sequences in a DNA-RNA heteroduplex. In some embodiments the chimeric stem-loop oligonucleotide adaptor and pro-primer are the same. In some embodiments, the stem sequence is designed for optimal performance of the chimeric primer generated by cleavage of the RNA portion of the stem heteroduplex by RNase H.

[0050] In various embodiments of the invention, a plurality of stem-loop oligonucleotides are used as adaptors and/or primers. For example, a first stem-loop oligonucleotide and a second stem-loop oligonucleotide can be used to attach to two different ends of a target nucleic acid. For another example, a first stem-loop oligonucleotide and a second stem-loop oligonucleotide can be used to prime from two different ends of a target nucleic acid. For yet another example, a first stem-loop oligonucleotide can be used as an adaptor and a second stem-loop oligonucleotide can be used as a primer. A first and a second stem-loop oligonucleotide can be 25%-100% similar to each other, including at least about 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, and 100% sequence similarity. A first and a second stem-loop oligonucleotide can be 25%-100% complementarity, including at least about 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%,

65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, and 100% sequence complementarity.

[0051] In some embodiments, a second stem-loop oligonucleotide interacts with a product generated using a first stem-loop oligonucleotide. For example, a first stem-loop oligonucleotide can be attached to a target nucleic acid via the oligonucleotide's 3' end and the target nucleic acid can be extended complementary to the attached oligonucleotide, as described below (Figure 2a). A portion of the attached oligonucleotide can be cleaved generating a priming ready construct. A second stem-loop oligonucleotide that is hybridizable to the priming ready construct, such as one that has a sequence similarity with the first stem-loop oligonucleotde, can further prime an amplification reaction. For another example, a first stem-loop oligonucleotide can be attached to a target nucleic acid via the oligonucleotide's 5' end. Upon cleavage of a heteroduplex stem region, a priming ready construct can be generated comprising a portion of the stem-loop oligonucleotide in a 3' overhang. A second stem-loop oligonucleotide, that is hybridizable to the priming ready construct, such as one that has a sequence complementarity with the first stem-loop oligonucleotde, can further prime an amplification reaction.

[0052] The sequence of the stem-loop chimeric RNA-DNA adaptor/pro-primer may further comprise one or more unique identifier sequences, or molecular barcodes, which will be useful for multiplex amplification and analysis of pooled samples.

[0053] Further features and uses of adaptor oligonucleotides are described in US Patent Publications US 2011-0319290 and US 2011-0105364, which are both herein incorporated by reference in their entirety.

Library generation using stem-loop chimeric RNA-DNA adaptor/pro-primer

[0054] In various embodiments, library generation of target nucleic acids is carried out by ligation of stem-loop chimeric RNA-DNA adaptor/pro-primer to the ends of the said target nucleic acids. The target nucleic acids can be dsDNA or fragments thereof. Cleavage of the RNA portion of the heteroduplex DNA-RNA stem can be achieved using an RNase targeting said RNA, such as RNase H. Replication of the released stem and or loop sequences that is attached to a first strand of the target nucleic acid can be achieved, for example, by extending a complementary strand along the released sequences. The released sequences can be directly adjacent to the 3'-DNA sequence of the stem or include the same. Synthesis of a DNA sequence along the stem-loop oligonucleotide sequences, for example by DNA polymerase, can yield a DNA-RNA heteroduplex comprising one or more RNA sequences from the stem-loop oligonucleotide and the synthesized DNA that is complementary. Cleavage of the RNA sequence(s) in the RNA-DNA heteroduplex, for example with RNase H, can result in a 3' overhang that is

complementary to at least a portion of the stem-loop oligonucleotide. In some embodiments, the RNA sequence(s) forming a duplex with the synthesized DNA originate from the loop of the stem-loop oligonucleotide. In some embodiments the DNA polymerase is a DNA dependent DNA polymerase with RNA dependent DNA polymerase activity. In some embodiment the DNA dependent DNA polymerase and the RNA dependent DNA polymerase are two different enzymes.

[0055] Figures 2a and 2b illustrate a series of reaction steps for the generation of nucleic acid libraries, according to some embodiments of the invention. The two figures differ in the first step of attachment of the stem-loop chimeric RNA-DNA adaptor/ pro-primer to the fragments of the dsDNA target, yielding two different kinds of stem-loop-target complexes. Adaptor attachment through ligation may be carried out by ligation to one of the strands of the dsDNA fragment (Figure 2a), resulting in a stem-loop-target complex with a nick or gap or to both strand of the dsDNA fragment (Figure 2b), resulting in a stem-loop-target complex with strand continuity, without a nick or a gap between the stem-loop oligonucleotide and the target. The nature of the ligation may depend on the nature of the ends of the dsDNA and the stem-loop adaptor as well as the choice of the ligase activity. In many cases, ligation depends on the phosphorylation on the 5' end. In these cases, ligation to both strands of the dsDNA can be achieved using a phosphorylated 5'-end on both the target nucleic acid and the stem-loop chimeric RNA-DNA adaptor/pro-primer. When only one of the components comprises a 5'-phosphate, only one of the strands of the target nucleic acids can be ligated to the adaptor/pro-primer. In some embodiments, a strand of the target nucleic acid is ligated to the double stranded portion of the stem-loop. When one end, e.g. the 5'-end, of the chimeric stem-loop RNA-DNA adaptor/pro-primer comprises RNA, employment of a ligase capable of ligating a DNA sequence to an RNA sequence can be beneficial. The ligase can further comprise DNA-DNA ligation activity allowing for a DNA end of the stem-loop oligonucleotide to ligate to a DNA target nucleic acid.

[0056] After ligation to either a single strand of the target nucleic acid, for example a double stranded target nucleic acid (Figure 2a) or to both strands (Figure 2b), a complementary sequence can be synthesized, for example using 3'-extension along the stem-loop oligonucleotide sequence. Cleaving the stem of the adaptor using an RNase, for example RNase H, can linearize the stem loop structure and release it from one of the strand of the target nucleic acid at the same time. The stem loop sequences can be replicated by a DNA polymerase through extension of a free 3'-end of the target nucleic acid, forming a target nucleic acid extension product. Target nucleic acid-adaptor constructs with a single strand attachment can be replicated without the activity of

QIAGEN Exhibit 1002
Page 823

an RNase. Extension of the fragment strand along the ligated chimeric DNA-RNA oligonucleotide can be carried out with a DNA polymerase comprising DNA- and RNA-dependent DNA polymerase, such as a reverse transcriptase or DNA polymerase comprising both activities, such as Bst, or a combination thereof. The target nucleic acid extension product resulting from the extension may comprise a DNA-RNA heteroduplex comprising an RNA sequence originating from the adaptor, e.g. from the loop, from any other adaptor sequence, or from a combination adaptor locations, and a DNA extension product. The RNA portion of the heteroduplex can be cleaved by an RNase, such as RNase H, in the same reaction mixture combining the product of the extension by DNA polymerase and RNase H. In various embodiments, the product of this reaction comprises a library of partial duplexes comprising dsDNA inserts of the target nucleic acids flanked by 3'-single stranded DNA extensions (or overhangs) comprising sequences complementary to the A and B sequences of the stem-loop chimeric RNA-DNA adaptor/pro-primer (Figure 2a and 2b).

[0057] The partial duplex products described above are suitable for further manipulation, for example generation of libraries comprising dsDNA fragments and 5'- and/or 3'-single stranded DNA overhangs as illustrated in Figure 3 or isothermal amplification using the chimeric RNA-DNA primers, as illustrated in Figures 4a and 4b.

Generation of libraries of dsDNA fragments with 5'- and 3'-single stranded non-complementary DNA comprising defined sequences

[0058] In various embodiments, libraries comprising dsDNA fragment inserts and known overhang sequences at one or both ends are generated from the partial duplexes described above. A partial duplex with a 3' overhang can be combined with an oligonucleotide comprising a 3'-sequence sufficiently complementary to a sequence of the single stranded 3' overhang portion of the partial duplex and a 5'-tail sequence. The complementary sequence can be directly adjacent to the double stranded region. The 5' tail sequence can be any desired sequence that is unable to efficiently hybridize to the 3' overhang based on lack of complementarity, under conditions which allow hybridization of the 3' end of the oligonucleotide to the partial duplex . The 3'-hybridized portion of the hybridized oligonucleotide can be ligated to the 5' end of the partial duplex downstream (Figure 3), thereby generating terminal Y (or forked) products comprising a dsDNA portion and non-complementary 3'- and 5'-single stranded sequences at a given end. Methods and compositions relating to terminal Y products are further described in US Patent No. 7,741,463, which is herein incorporated in its entirety.

[0059] A library comprising nucleic acids with greater dsDNA content or fully double stranded (ds) nucleic acids with desired sequences at both ends can be generated. The desired sequences can

be different on each end of a strand. An oligonucleotide comprising a sequence complementary to a single stranded 3'-end of a terminal Y product like the ones described above can be hybridized to the terminal Y product. Extension of the hybridized oligonucleotide (primer) by DNA polymerase can generate dsDNA products comprising a target insert and a distinct sequence at each end, as shown in Figure 3. This process provides for the generation of libraries with asymmetric end sequences without the need to append/ligate a different sequence to each end. The described process of appending sequences is symmetric with increased efficiency for the library generation.

[0060] Further amplification of the various libraries described herein is possible, when desired, by various methods including, for example, PCR employing suitable forward and reverse primers or isothermal amplification/SPIA with one or more suitable primer(s). In some embodiments, the amplification primers may further comprise tails that are not hybridizable to the unique sequences of the library, thus extending the common sequences used in the library as may be desired for further manipulation of the libraries, such as for sequencing.

Isothermal amplification using chimeric RNA-DNA primers

[0061] Partial duplex products comprising dsDNA insert of the fragmented dsDNA template described above (Figures 2a and 2b), can be used as substrates for isothermal amplification employing chimeric RNA-DNA primers (SPIA), as shown in Figures 4a and 4b. Amplification may be carried out with stem-loop chimeric RNA-DNA pro-primer or a linear, non stem-loop chimeric RNA-DNA primer. The chimeric primer may be the stem-loop chimeric adaptor/pro-primer employed for the generation of the said partial duplex.

[0062] Amplification can be carried out in reaction mixtures comprising a partial duplex, a chimeric RNA-DNA primer, DNA polymerase with strand displacement and RNase, such as RNase H, which cleaves RNA in a RNA-DNA duplex. The chimeric RNA-DNA primer may be generated in the reaction mixture from the stem-loop chimeric RNA-DNA pro-primer through cleavage of the RNA portion of the stem RNA-DNA heteroduplex. The generation of an active primer from the pro-primer in the reaction mixture reduces the potential accumulation of undesired products that may be generated in the presence of high concentration of active primer. Active primer can be generated from the stem loop chimeric RNA-DNA pro-primer by cleavage of the RNA in the RNA-DNA heteroduplex stem portion by RNase H as shown in Figure 4a (insert).

[0063] The sequence of the chimeric RNA-DNA primer can comprise a sequence that is the same as sequence A and or B (the loop sequence of the stem-loop chimeric pro-primer; Figure 4a), or may comprise sequences hybridizable to the complement of the loop sequence (Figure 4b).

[0064] Figure 5 illustrates the use of stem-loop oligonucleotides as pro-primers and primers according to some embodiments of the invention. Accordingly, a composite amplification primer can be generated in the amplification reaction mixture from a stem-loop chimeric pro-primer. The amplification reaction mixture can comprise one or more target partial duplex nucleic acid(s), for example a partial duplex DNA with a 3' overhang as illustrated in the figure, one or more chimeric stem-loop pro-primer(s), one or more DNA polymerase(s) with strand displacement activity, and one or more RNases targeting RNA in RNA/DNA heteroduplexes, for example RNase H. The RNA portion of the RNA/DNA heteroduplex at the stem of the chimeric stem-loop pro-primer can be cleaved by RNase H to generate, for example, a linear composite primer comprising a 3'-DNA and 5'-RNA. The linearized amplification primer can hybridize to a 3'-single stranded DNA portion (overhang) of a target partial duplex and can be extended by the DNA polymerase with strand displacement activity. The RNA portion of the hybridized primer in a heteroduplex can be cleaved by RNase H to free a portion of the primer binding site. A second linear composite amplification primer can hybridize to the freed primer binding site, and can be extended along the target DNA strand. The previously synthesized primer extension product (amplification product) can be displaced by the newly extended primer. Repeated cycles of primer hybridization, primer extension by strand displacement DNA polymerase, and cleavage of the RNA portion of the hybridized primer can generate multiple copies of a target nucleic acid (Figure 5).

Target polynucleotides

[0065] In various embodiments of the invention, nucleic acid are used as substrates for further manipulation. The input nucleic acid can be DNA, or complex DNA, for example genomic DNA. The input DNA may also be cDNA. The cDNA can be generated from RNA, e.g., mRNA. The input DNA can be of a specific species, for example, human, rat, mouse, other animals, plants, bacteria, algae, viruses, and the like. The input nucleic acid also can be from a mixture of genomes of different species such as host-pathogen, bacterial populations and the like. The input DNA can be cDNA made from a mixture of genomes of different species. Alternatively, the input nucleic acid can be from a synthetic source. The input DNA can be mitochondrial DNA. The input DNA can be cell-free DNA. The cell-free DNA can be obtained from, e.g., a serum or plasma sample. The input DNA can comprise one or more chromosomes. For example, if the input DNA is from a human, the DNA can comprise one or more of chromosome 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, X, or Y. The DNA can be from a linear

QIAGEN Exhibit 1002
Page 826

or circular genome. The DNA can be plasmid DNA, cosmid DNA, bacterial artificial chromosome (BAC), or yeast artificial chromosome (YAC). The input DNA can be from more than one individual or organism. The input DNA can be double stranded or single stranded. The input DNA can be part of chromatin. The input DNA can be associated with histones. The methods described herein can be applied to high molecular weight DNA, such as is isolated from tissues or cell culture, for example, as well as highly degraded DNA, such as cell-free DNA from blood and urine and/or DNA extracted from formalin-fixed, paraffin-embedded tissues, for example.

[0066] The different samples from which the target polynucleotides are derived can comprise multiple samples from the same individual, samples from different individuals, or combinations thereof. In some embodiments, a sample comprises a plurality of polynucleotides from a single individual. In some embodiments, a sample comprises a plurality of polynucleotides from two or more individuals. An individual is any organism or portion thereof from which target polynucleotides can be derived, non-limiting examples of which include plants, animals, fungi, protists, monerans, viruses, mitochondria, and chloroplasts. Sample polynucleotides can be isolated from a subject, such as a cell sample, tissue sample, or organ sample derived therefrom, including, for example, cultured cell lines, biopsy, blood sample, or fluid sample containing a cell. The subject may be an animal, including but not limited to, an animal such as a cow, a pig, a mouse, a rat, a chicken, a cat, a dog, etc., and is usually a mammal, such as a human. Samples can also be artificially derived, such as by chemical synthesis. In some embodiments, the samples comprise DNA. In some embodiments, the samples comprise genomic DNA. In some embodiments, the samples comprise mitochondrial DNA, chloroplast DNA, plasmid DNA, bacterial artificial chromosomes, yeast artificial chromosomes, oligonucleotide tags, or combinations thereof. In some embodiments, the samples comprise DNA generated by primer extension reactions using any suitable combination of primers and a DNA polymerase, including but not limited to polymerase chain reaction (PCR), reverse transcription, and combinations thereof. Where the template for the primer extension reaction is RNA, the product of reverse transcription is referred to as complementary DNA (cDNA). Primers useful in primer extension reactions can comprise sequences specific to one or more targets, random sequences, partially random sequences, and combinations thereof. Reaction conditions suitable for primer extension reactions are known in the art. In general, sample polynucleotides

comprise any polynucleotide present in a sample, which may or may not include target polynucleotides.

[0067] Methods for the extraction and purification of nucleic acids are well known in the art. For example, nucleic acids can be purified by organic extraction with phenol, phenol/chloroform/isoamyl alcohol, or similar formulations, including TRIzol and TriReagent. Other non-limiting examples of extraction techniques include: (1) organic extraction followed by ethanol precipitation, e.g., using a phenol/chloroform organic reagent (Ausubel et al., 1993), with or without the use of an automated nucleic acid extractor, e.g., the Model 341 DNA Extractor available from Applied Biosystems (Foster City, Calif.); (2) stationary phase adsorption methods (U.S. Pat. No. 5,234,809; Walsh et al., 1991); and (3) salt-induced nucleic acid precipitation methods (Miller et al., (1988), such precipitation methods being typically referred to as "salting-out" methods. Another example of nucleic acid isolation and/or purification includes the use of magnetic particles to which nucleic acids can specifically or non-specifically bind, followed by isolation of the beads using a magnet, and washing and eluting the nucleic acids from the beads (see e.g. U.S. Pat. No. 5,705,628). In some embodiments, the above isolation methods may be preceded by an enzyme digestion step to help eliminate unwanted protein from the sample, e.g., digestion with proteinase K, or other like proteases. See, e.g., U.S. Pat. No. 7,001,724. If desired, RNase inhibitors may be added to the lysis buffer. For certain cell or sample types, it may be desirable to add a protein denaturation/digestion step to the protocol. Purification methods may be directed to isolate DNA, RNA, or both. When both DNA and RNA are isolated together during or subsequent to an extraction procedure, further steps may be employed to purify one or both separately from the other. Sub-fractions of extracted nucleic acids can also be generated, for example, purification by size, sequence, or other physical or chemical characteristic. In addition to an initial nucleic isolation step, purification of nucleic acids can be performed after any step in the methods of the invention, such as to remove excess or unwanted reagents, reactants, or products.

Fragmentation methods

[0068] In some embodiments, sample polynucleotides are fragmented into a population of fragmented insert DNA molecules of one or more specific size range(s). In some embodiments, fragments are generated from at least about 1, 10, 100, 1000, 10000, 100000, 300000, 500000, or more genome-equivalents of starting DNA. Fragmentation may be accomplished by methods known in the art, including chemical, enzymatic, and

QIAGEN Exhibit 1002
Page 828

mechanical fragmentation. In some embodiments, the fragments have an average length from about 10 to about 10,000 nucleotides. In some embodiments, the fragments have an average length from about 50 to about 2,000 nucleotides. In some embodiments, the fragments have an average length from about 100-2,500, 10-1,000, 10-800, 10-500, 50-500, 50-250, or 50-150 nucleotides. In some embodiments, the fragments have an average length less than 500 nucleotides, such as less than 400 nucleotides, less than 300 nucleotides, less than 200 nucleotides, or less than 150 nucleotides. In some embodiments, the fragmentation is accomplished mechanically comprising subjecting sample polynucleotides to acoustic sonication. In some embodiments, the fragmentation comprises treating the sample polynucleotides with one or more enzymes under conditions suitable for the one or more enzymes to generate double-stranded nucleic acid breaks. Examples of enzymes useful in the generation of polynucleotide fragments include sequence specific and non-sequence specific nucleases. Non-limiting examples of nucleases include DNase I, Fragmentase, restriction endonucleases, variants thereof, and combinations thereof. For example, digestion with DNase I can induce random double-stranded breaks in DNA in the absence of Mg++ and in the presence of Mn++. In some embodiments, fragmentation comprises treating the sample polynucleotides with one or more restriction endonucleases. Fragmentation can produce fragments having 5' overhangs, 3' overhangs, blunt ends, or a combination thereof. In some embodiments, such as when fragmentation comprises the use of one or more restriction endonucleases, cleavage of sample polynucleotides leaves overhangs having a predictable sequence. In some embodiments, the method includes the step of size selecting the fragments via standard methods such as column purification or isolation from an agarose gel. Combination of fragmentation methods can be utilized, such as a combination enzymatic and chemical methods. In a particular example, an abasic site can be generated, e.g. using a glycosylase (Uracil-DNA glycosylase, Thymine-DNA glycosylase etc.), and the abasic site can be cleaved using a chemical method, such as by contacting the abasic site with dimethylethylenediamine (DMED).

[0069] In some embodiments, the 5' and/or 3' end nucleotide sequences of fragmented DNA are not modified prior to ligation with one or more adaptor oligonucleotides. For example, fragmentation by a restriction endonuclease can be used to leave a predictable overhang, followed by ligation with one or more adaptor oligonucleotides comprising an overhang complementary to the predictable overhang on a DNA fragment. In another example, cleavage by an enzyme that leaves a predictable blunt end can be followed by ligation of

blunt-ended DNA fragments to adaptor oligonucleotides comprising a blunt end. In some embodiments, the fragmented DNA molecules are blunt-end polished (or "end repaired") to produce DNA fragments having blunt ends, prior to being joined to adaptors. The blunt-end polishing step may be accomplished by incubation with a suitable enzyme, such as a DNA polymerase that has both 3' to 5' exonuclease activity and 5' to 3' polymerase activity, for example T4 polymerase. In some embodiments, end repair is followed by an addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more nucleotides, such as one or more adenine, one or more thymine, one or more guanine, or one or more cytosine, to produce an overhang. DNA fragments having an overhang can be joined to one or more adaptor oligonucleotides having a complementary overhang, such as in a ligation reaction. For example, a single adenine can be added to the 3' ends of end repaired DNA fragments using a template independent polymerase, followed by ligation to one or more adaptors each having a thymine at a 3' end. In some embodiments, adaptor oligonucleotides can be joined to blunt end double-stranded DNA fragment molecules which have been modified by extension of the 3' end with one or more nucleotides followed by 5' phosphorylation. In some cases, extension of the 3' end may be performed with a polymerase such as for example Klenow polymerase or any of the suitable polymerases provided herein, or by use of a terminal deoxynucleotide transferase, in the presence of one or more dNTPs in a suitable buffer containing magnesium. In some embodiments, target polynucleotides having blunt ends are joined to one or more adaptors comprising a blunt end. Phosphorylation of 5' ends of DNA fragment molecules may be performed for example with T4 polynucleotide kinase in a suitable buffer containing ATP and magnesium. The fragmented DNA molecules may optionally be treated to dephosphorylate 5' ends or 3' ends, for example, by using enzymes known in the art, such as phosphatases.

[0070] In some embodiments, each of the plurality of independent samples comprises at least about 1pg, 10pg, 100pg, 1ng, 10ng, 20ng, 30ng, 40ng, 50ng, 75ng, 100ng, 150ng, 200ng, 250ng, 300ng, 400ng, 500ng, 1μg, 1.5μg, 2μg, or more of nucleic acid material. In some embodiments, each of the plurality of independent samples comprises less than about 1pg, 10pg, 100pg, 1ng, 10ng, 20ng, 30ng, 40ng, 50ng, 75ng, 100ng, 150ng, 200ng, 250ng, 300ng, 400ng, 500ng, 1μg, 1.5μg, 2μg, or more of nucleic acid.

[0071] In some embodiments each of the individual or plurality of samples comprises a single polynucleotide target or a single genome.

QIAGEN Exhibit 1002
Page 830

[0072] In another aspect, the invention provides compositions that can be used in the above described methods. Compositions of the invention can comprise any one or more of the elements described herein. In one embodiment, the composition comprises a plurality of target polynucleotides, each target polynucleotide comprising one or more barcode sequences selected from a plurality of barcode sequences, wherein said target polynucleotides are from two or more different samples, and further wherein the sample from which each of said polynucleotides is derived can be identified in a combined sequencing reaction with an accuracy of at least 95% based on a single barcode contained in the sequence of said target polynucleotide. In some embodiments, the composition comprises a plurality of first adaptor/primer oligonucleotides, wherein each of said first adaptor/primer oligonucleotides comprises at least one of a plurality of barcode sequences, wherein each barcode sequence of the plurality of barcode sequences differs from every other barcode sequence in said plurality of barcode sequences at at least three nucleotide positions.

Features of stem-loop oligonucleotides

[0073] Adaptors/primers can contain one or more of a variety of sequence elements, including but not limited to, one or more amplification primer annealing sequences or complements thereof, one or more sequencing primer annealing sequences or complements thereof, one or more transcriptional promoter sequences (e.g. T7, T3, SP6, etc.) or complements thereof, one or more barcode sequences, one or more common sequences shared among multiple different adaptors or subsets of different adaptors (adaptor group sequence), one or more restriction enzyme recognition sites, one or more overhangs complementary to one or more target polynucleotide overhangs, one or more probe binding sites (e.g. for attachment to a sequencing platform, such as a flow cell for massive parallel sequencing, such as developed by Illumina, Inc.), one or more random or near-random sequences (e.g. one or more nucleotides selected at random from a set of two or more different nucleotides at one or more positions, with each of the different nucleotides selected at one or more positions represented in a pool of adaptors comprising the random sequence), and combinations thereof. Two or more sequence elements can be non-adjacent to one another (e.g. separated by one or more nucleotides), adjacent to one another, partially overlapping, or completely overlapping. For example, an amplification primer annealing sequence can also serve as a sequencing primer annealing sequence. Sequence elements can be located at or near the 3' end, at or near the 5' end, or in the

interior of the adaptor oligonucleotide. In stem-loop oligonucleotides comprising stem, loop and/or overhang regions , sequence elements can be located partially or completely outside the secondary structure, partially or completely inside the secondary structure, or in between sequences participating in the secondary structure. For example, , sequence elements can be located partially or completely inside or outside the hybridizable sequences (the "stem"), including in the sequence between the hybridizable sequences (the "loop"). In some embodiments, the first adaptor/primer oligonucleotides in a plurality of first adaptor/primer oligonucleotides having different barcode sequences comprise a sequence element common among all first adaptor oligonucleotides in the plurality. In some embodiments, all second adaptor/primer oligonucleotides comprise a sequence element common among all second adaptor/primer oligonucleotides that is different from the common sequence element shared by the first adaptor/primer oligonucleotides. A difference in sequence elements can be any such that at least a portion of different adaptors /primers do not completely align, for example, due to changes in sequence length, deletion or insertion of one or more nucleotides, or a change in the nucleotide composition at one or more nucleotide positions (such as a base change or base modification). In some embodiments, an adaptor/primer oligonucleotide comprises a 5' overhang, a 3' overhang, or both that is complementary to one or more target polynucleotides. Complementary overhangs can be one or more nucleotides in length, including but not limited to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or more nucleotides in length. Complementary overhangs may comprise a fixed sequence. Complementary overhangs may comprise a random sequence of one or more nucleotides, such that one or more nucleotides are selected at random from a set of two or more different nucleotides at one or more positions, with each of the different nucleotides selected at one or more positions represented in a pool of adaptors with complementary overhangs comprising the random sequence. In some embodiments, an adaptor/primer overhang is complementary to a target polynucleotide overhang produced by restriction endonuclease digestion. In some embodiments, an adaptor overhang consists of an adenine or a thymine.

[0074] Adaptor/primer oligonucleotides can have any suitable length, at least sufficient to accommodate the one or more sequence elements of which they are comprised. In some embodiments, adaptor/primers are about, less than about, or more than about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 200, or more nucleotides in length. In some embodiments, the stem of a stem-loop adaptor/primer is about, less than about,

or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, or more nucleotides in length. Stems may be designed using a variety of different sequences that result in hybridization between the complementary regions on a stem-loop chimeric adaptor, resulting in a local region of double-stranded DNA. For example, stem sequences may be utilized that are from 15 to 18 nucleotides in length with equal representation of G:C and A:T base pairs. Such stem sequences are predicted to form stable dsDNA structures below their predicted melting temperatures of ~45°C. Sequences participating in the stem of the stem-loop adaptor can be perfectly complementary, such that each base of one region in the stem hybridizes via hydrogen bonding with each base in the other region in the stem according to Watson-Crick base-pairing rules. Alternatively, sequences in the stem may deviate from perfect complementarity. For example, there can be mismatches and or bulges within the stem structure created by opposing bases that do not follow Watson-Crick base pairing rules, and/or one or more nucleotides in one region of the stem that do not have the one or more corresponding base positions in the other region participating in the stem. Mismatched sequences may be cleaved using enzymes that recognize mismatches. The stem and/or loop of a stem-loop can comprise DNA, RNA, or both DNA and RNA. In some embodiments, the stem and/or loop of a stem-loop, or one or both of the hybridizable sequences forming the stem of a stem-loop, comprise nucleotides, bonds, or sequences that are substrates for cleavage, such as by an enzyme, including but not limited to endonucleases and glycosylases. The composition of a stem may be such that only one of the hybridizable sequences forming the stem is cleaved. For example, one of the sequences forming the stem may comprise RNA while the other sequence forming the stem consists of DNA, such that cleavage by an enzyme that cleaves RNA in an RNA-DNA duplex, such as RNase H, cleaves only the sequence comprising RNA. The stem and/or loop of a stem-loop oligonucleotide can comprise non-canonical nucleotides (e.g. uracil), and/or methylated nucleotides. In some embodiments, the loop sequence of a hairpin adaptor is about, less than about, or more than about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, or more nucleotides in length.

Barcodes

[0075] As used herein, the term "barcode" refers to a known nucleic acid sequence that allows some feature of a polynucleotide with which the barcode is associated to be identified. In some embodiments, the feature of the polynucleotide to be identified is the sample from

which the polynucleotide is derived. In some embodiments, barcodes are at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or more nucleotides in length. In some embodiments, barcodes are shorter than 10, 9, 8, 7, 6, 5, or 4 nucleotides in length. In some embodiments, barcodes associated with some polynucleotides are of different length than barcodes associated with other polynucleotides. In general, barcodes are of sufficient length and comprise sequences that are sufficiently different to allow the identification of samples based on barcodes with which they are associated. In some embodiments, a barcode, and the sample source with which it is associated, can be identified accurately after the mutation, insertion, or deletion of one or more nucleotides in the barcode sequence, such as the mutation, insertion, or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides. In some embodiments, each barcode in a plurality of barcodes differ from every other barcode in the plurality at at least three nucleotide positions, such as at least 3, 4, 5, 6, 7, 8, 9, 10, or more positions. In some embodiments, both a first adaptor/primer and a second adaptor/primer comprise at least one of a plurality of barcode sequences. In some embodiments, barcodes for second adaptor/primer oligonucleotides are selected independently from barcodes for first adaptor/primer oligonucleotides. In some embodiments, first adaptor/primer oligonucleotides and second adaptor/primer oligonucleotides having barcodes are paired, such that adaptor/primers of the pair comprise the same or different one or more barcodes. In some embodiments, the methods of the invention further comprise identifying the sample from which a target polynucleotide is derived based on a barcode sequence to which the target polynucleotide is joined. In general, a barcode comprises a nucleic acid sequence that when joined to a target polynucleotide serves as an identifier of the sample from which the target polynucleotide was derived.

[0076] In some embodiments, the plurality of barcode sequences from which barcode sequences are selected includes sequences selected from the group consisting of: AAA, TTT, CCC, GGG. In some embodiments, the plurality of barcode sequences from which barcode sequences are selected includes sequences selected from the group consisting of: AAAA, CTGC, GCTG, TGCT, ACCC, CGTA, GAGT, TTAG, AGGG, CCAT, GTCA, TATC, ATTT, CACG, GGAC, and TCGA. In some embodiments, the plurality of barcode sequences from which barcode sequences are selected includes sequences selected from the group consisting of: AAAAA, AACCC, AAGGG, AATTT, ACACG, ACCAT, ACGTA, ACTGC, AGAGT, AGCTG, AGGAC, AGTCA, ATATC, ATCGA, ATGCT, ATTAG, CAACT, CACAG, CAGTC, CATGA, CCAAC, CCCCA, CCGGT, CCTTG,

CGATA, CGCGC, CGGCG, CGTAT, CTAGG, CTCTT, CTGAA, CTTCC, GAAGC, GACTA, GAGAT, GATCG, GCATT, GCCGG, GCGCC, GCTAA, GGAAG, GGCCT, GGGGA, GGTTC, GTACA, GTCAC, GTGTG, GTTTT, TAATG, TACGT, TAGCA, TATAC, TCAGA, TCCTC, TCGAG, TCTCT, TGACC, TGCAA, TGGTT, TGTGG, TTAAT, TTCCG, TTGGC, and TTTTA.

Attachment of Adaptors

Ligation

[0077] The terms "joining" and "ligation" as used herein, with respect to two polynucleotides, such as a stem-loop adaptor/primer oligonucleotide and a target polynucleotide, refers to the covalent attachment of two separate polynucleotides to produce a single larger polynucleotide with a contiguous backbone. Methods for joining two polynucleotides are known in the art, and include without limitation, enzymatic and non-enzymatic (e.g. chemical) methods. Examples of ligation reactions that are non-enzymatic include the non-enzymatic ligation techniques described in U.S. Pat. Nos. 5,780,613 and 5,476,930, which are herein incorporated by reference. In some embodiments, an adaptor oligonucleotide is joined to a target polynucleotide by a ligase, for example a DNA ligase or RNA ligase. Multiple ligases, each having characterized reaction conditions, are known in the art, and include, without limitation NAD^+-dependent ligases including tRNA ligase, Taq DNA ligase, *Thermus filiformis* DNA ligase, Escherichia coli DNA ligase, Tth DNA ligase, *Thermus scotoductus* DNA ligase (I and II), thermostable ligase, Ampligase thermostable DNA ligase, VanC-type ligase, 9° N DNA Ligase, Tsp DNA ligase, and novel ligases discovered by bioprospecting; ATP-dependent ligases including T4 RNA ligase, T4 DNA ligase, T3 DNA ligase, T7 DNA ligase, Pfu DNA ligase, DNA ligase 1, DNA ligase III, DNA ligase IV, and novel ligases discovered by bioprospecting; and wild-type, mutant isoforms, and genetically engineered variants thereof. Ligation can be between polynucleotides having hybridizable sequences, such as complementary overhangs. Ligation can also be between two blunt ends. Generally, a 5' phosphate is utilized in a ligation reaction. The 5' phosphate can be provided by the target polynucleotide, the adaptor oligonucleotide, or both. 5' phosphates can be added to or removed from polynucleotides to be joined, as needed. Methods for the addition or removal of 5' phosphates are known in the art, and include without limitation enzymatic and chemical processes. Enzymes useful in the addition and/or removal of 5' phosphates include kinases, phosphatases, and polymerases. In some embodiments, both of the two ends joined in a ligation reaction (e.g. an adaptor end and a target polynucleotide end)

provide a 5' phosphate, such that two covalent linkages are made in joining the two ends. In some embodiments, only one of the two ends joined in a ligation reaction (e.g. only one of an adaptor end and a target polynucleotide end) provides a 5' phosphate, such that only one covalent linkage is made in joining the two ends. In some embodiments, only one strand at one or both ends of a target polynucleotide is joined to an adaptor oligonucleotide. In some embodiments, both strands at one or both ends of a target polynucleotide are joined to an adaptor oligonucleotide. In some embodiments, 3' phosphates are removed prior to ligation. In some embodiments, an adaptor oligonucleotide is added to both ends of a target polynucleotide, wherein one or both strands at each end are joined to one or more adaptor oligonucleotides. When both strands at both ends are joined to an adaptor oligonucleotide, joining can be followed by a cleavage reaction that leaves a 5' overhang that can serve as a template for the extension of the corresponding 3' end, which 3' end may or may not include one or more nucleotides derived from the adaptor oligonucleotide. In some embodiments, a target polynucleotide is joined to a first adaptor oligonucleotide on one end and a second adaptor oligonucleotide on the other end. In some embodiments, the target polynucleotide and the adaptor to which it is joined comprise blunt ends. In some embodiments, separate ligation reactions are carried out for each sample, using a different first adaptor oligonucleotide comprising at least one barcode sequence for each sample, such that no barcode sequence is joined to the target polynucleotides of more than one sample. A target polynucleotide that has an adaptor/primer oligonucleotide joined to it is considered "tagged" by the joined adaptor.

[0078] In some embodiments, joining of an adaptor/primer to a target polynucleotide produces a joined product polynucleotide having a 3' overhang comprising a nucleotide sequence derived from the adaptor/primer. In some embodiments, a primer oligonucleotide comprising a sequence complementary to all or a portion of the 3' overhang is hybridized to the overhang and extended using a DNA polymerase to produce a primer extension product hybridized to one strand of the joined product polynucleotide. The DNA polymerase may comprise strand displacement activity, such that one strand of the joined product polynucleotide is displaced during primer extension.

Pooling

[0079] In some embodiments, after joining at least one adaptor oligonucleotide to a target polynucleotide, the 3' end of one or more target polynucleotides is extended using the one or more joined adaptor oligonucleotides as template. For example, an adaptor

comprising two hybridized oligonucleotides that is joined to only the 5' end of a target polynucleotide allows for the extension of the unjoined 3'end of the target using the joined strand of the adaptor as template, concurrently with or following displacement of the unjoined strand. If both strands of an adaptor comprising two hybridized oligonucleotides are joined to a target polynucleotide such that the joined product has a 5' overhang, the complementary 3' end can be extended using the 5' overhang as template. As a further example, a stem-loop adaptor oligonucleotide can be joined to the 5' end of a target polynucleotide. While double-stranded in secondary structure, such a stem-loop adaptor remains single-stranded, and is thus a 5' overhang appended to the target polynucleotide (e.g. when the 5' end of the stem-loop adaptor is not joined to the target polynucleotide). Removal of the secondary structure, either prior to (e.g. thermal denaturing, or degradation) or concurrently with (e.g. strand displacement) the activity of a polymerase, provides a template for the extension of the 3' end of the complementary strand of the target polynucleotide. In some embodiments, the 3' end of the target polynucleotide that is extended comprises one or more nucleotides from an adaptor oligonucleotide. For target polynucleotides to which adaptors are joined on both ends, extension can be carried out for both 3' ends of a double-stranded target polynucleotide having 5' overhangs. This 3' end extension, or "fill-in" reaction, generates a complementary sequence, or "complement," to the adaptor oligonucleotide template that is hybridized to the template, thus filling in the 5' overhang to produce a double-stranded sequence region. Where both ends of a double-stranded target polynucleotide have 5' overhangs that are filled in by extension of the complementary strands' 3' ends, the product is completely double-stranded. Extension can be carried out by any suitable polymerase known in the art, such as a DNA polymerase, many of which are commercially available. DNA polymerases can comprise DNA-dependent DNA polymerase activity, RNA-dependent DNA polymerase activity, or DNA-dependent and RNA-dependent DNA polymerase activity. DNA polymerases can be thermostable or non-thermostable. Examples of DNA polymerases include, but are not limited to, Taq polymerase, Tth polymerase, Tli polymerase, Pfu polymerase, Pfutubo polymerase, Pyrobest polymerase, Pwo polymerase, KOD polymerase, Bst polymerase, Sac polymerase, Sso polymerase, Poc polymerase, Pab polymerase, Mth polymerase, Pho polymerase, ES4 polymerase, VENT polymerase, DEEPVENT polymerase, EX-Taq polymerase, LA-Taq polymerase, Expand polymerases, Platinum Taq polymerases, Hi-Fi polymerase, Tbr polymerase, Tfl polymerase, Tru polymerase, Tac polymerase, Tne

polymerase, Tma polymerase, Tih polymerase, Tfi polymerase, Klenow fragment, and variants, modified products and derivatives thereof. 3' end extension can be performed before or after pooling of target polynucleotides from independent samples.

[0080] In some embodiments, the fill-in reaction is followed by or performed as part of amplification of one or more target polynucleotides using a first primer and/or a second primer, wherein the first primer comprises a sequence that is hybridizable to at least a portion of the complement of one or more of a first adaptor/primer oligonucleotides, and further wherein the second primer comprises a sequence that is hybridizable to at least a portion of the complement of one or more of a second adaptor/primer oligonucleotides. Each of the first and second primers may be of any suitable length, such as about, less than about, or more than about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, or more nucleotides, any portion or all of which may be complementary to the corresponding target sequence (e.g. about, less than about, or more than about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, or more nucleotides). "Amplification" refers to any process by which the copy number of a target sequence is increased. Methods for primer-directed amplification of target polynucleotides are known in the art, and include without limitation, methods based on the polymerase chain reaction (PCR). Conditions favorable to the amplification of target sequences by PCR are known in the art, can be optimized at a variety of steps in the process, and depend on characteristics of elements in the reaction, such as target type, target concentration, sequence length to be amplified, sequence of the target and/or one or more primers, primer length, primer concentration, polymerase used, reaction volume, ratio of one or more elements to one or more other elements, and others, some or all of which can be altered. In general, PCR involves the steps of denaturation of the target to be amplified (if double stranded), hybridization of one or more primers to the target, and extension of the primers by a DNA polymerase, with the steps repeated (or "cycled") in order to amplify the target sequence. Steps in this process can be optimized for various outcomes, such as to enhance yield, decrease the formation of spurious products, and/or increase or decrease specificity of primer annealing. Methods of optimization are well known in the art and include adjustments to the type or amount of elements in the amplification reaction and/or to the conditions of a given step in the process, such as temperature at a particular step, duration of a particular step, and/or number of cycles. In some embodiments, an amplification reaction comprises at least 5, 10, 15, 20, 25, 30, 35, 50, or more cycles. In some embodiments, an amplification reaction comprises no more than 5, 10, 15, 20, 25, 35, 50, or more cycles. Cycles can

contain any number of steps, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more steps. Steps can comprise any temperature or gradient of temperatures, suitable for achieving the purpose of the given step, including but not limited to, 3' end extension (e.g. adaptor fill-in), primer annealing, primer extension, and strand denaturation. Steps can be of any duration, including but not limited to about, less than about, or more than about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 180, 240, 300, 360, 420, 480, 540, 600, or more seconds, including indefinitely until manually interrupted. Cycles of any number comprising different steps can be combined in any order. In some embodiments, different cycles comprising different steps are combined such that the total number of cycles in the combination is about, less that about, or more than about 5, 10, 15, 20, 25, 30, 35, 50, or more cycles. In some embodiments, amplification is performed following the fill-in reaction. Amplification can be performed before or after pooling of target polynucleotides from independent samples.

Methods
Methods of Amplification

[0081] The methods, compositions and kits described herein can be useful to generate amplification-ready products directly from genomic DNA or whole or partial transcriptome RNA for downstream applications such as massively parallel sequencing (Next Generation Sequencing methods), multiplexed quantification of large sets of sequence regions of interest, such as by high density qPCR arrays and other highly parallel quantification platforms (selective massively parallel target pre-amplification), as well as generation of libraries with enriched population of sequence regions of interest. The methods described herein can be used to generate a collection of at least 25, 50, 75, 100, 500, 1000, 2500, 5000, 10,000, 25,000, 50,000, 100,000, 500,000, or 1,000,000 amplification-ready target sequence regions of interest directly from a sample of complex DNA using a plurality of oligonucleotides.

[0082] Methods of nucleic acid amplification are well known in the art. In some embodiments, the amplification method is isothermal. In other embodiments the amplification method is linear. In other embodiments the amplification is exponential.

[0083] **SPIA Amplification**

[0084] Amplification of the sequence regions of interest employing a linear amplification method such as the single primer isothermal amplification (SPIA) can be used. SPIA enables generation of multiple copies of the strand specific sequence regions of interest

and employs a single amplification primer, thus reducing the complexity associated with multiple oligonucleotide design and manufacturing, enables the use of a generic amplification primer, and can be linear. The fidelity of quantification of the copy number of the sequence regions of interest in the complex genomic NA sample is a highly desirable feature of the presented methods of the invention.

[0085] Amplification by SPIA can occur under conditions permitting composite primer hybridization, primer extension by a DNA polymerase with strand displacement activity, cleavage of RNA from a RNA/DNA heteroduplex and strand displacement. In so far as the composite amplification primer hybridizes to the 3′-single-stranded portion (of the partially double stranded polynucleotide which is formed by cleaving RNA in the complex comprising a RNA/DNA partial heteroduplex) comprising, generally, the complement of at least a portion of the composite amplification primer sequence, composite primer hybridization may be under conditions permitting specific hybridization. In SPIA, all steps are isothermal (in the sense that thermal cycling is not required), although the temperatures for each of the steps may or may not be the same. It is understood that various other embodiments can be practiced given the general description provided above. For example, as described and exemplified herein, certain steps may be performed as temperature is changed (e.g., raised, or lowered).

[0086] Although generally only one composite amplification primer is described above, it is further understood that the SPIA amplification methods can be performed in the presence of two or more different first and/or second composite primers that randomly prime template polynucleotide. In addition, the amplification polynucleotide products of two or more separate amplification reactions conducted using two or more different first and/or second composite primers that randomly prime template polynucleotide can be combined.

[0087] The composite amplification primers are primers that are composed of RNA and DNA portions. In the amplification composite primer, both the RNA and the DNA portions are generally complementary and can hybridize to a sequence in the amplification-ready product to be copied or amplified. In some embodiments, a 3′-portion of the amplification composite primer is DNA and a 5′-portion of the composite amplification primer is RNA. The composite amplification primer is designed such that the primer is extended from the 3′-DNA portion to create a primer extension product. The 5′-RNA portion of this primer extension product in a RNA/DNA heteroduplex is susceptible to cleavage by RNase H, thus freeing a portion of the polynucleotide to the hybridization of an additional

composite amplification primer. The extension of the amplification composite primer by a DNA polymerase with strand displacement activity releases the primer extension product from the original primer and creates another copy of the sequence of the polynucleotide. Repeated rounds of primer hybridization, primer extension with strand displacement DNA synthesis, and RNA cleavage create multiple copies of the strand-specific sequence of the polynucleotide.

[0088] In some embodiments, the composite amplification primer is generated in the amplification reaction mixture from a stem-loop chimeric pro-primer. The amplification reaction mixture can comprise a target partial duplex nucleic acid, for example a target partial duplex DNA, a chimeric stem-loop pro-primer, DNA polymerase with strand displacement activity, and an RNase targeting RNA in a RNA/DNA heteroduplex, for example RNase H. The RNA portion of the RNA/DNA heteroduplex at the stem of the chimeric stem-loop pro-primer can be cleaved by RNase H to generate, for example, a linear composite primer comprising a 3'-DNA and 5'-RNA. The linearized amplification primer can hybridize to a 3'-single stranded DNA portion (overhang) of a target partial duplex and can be extended by the DNA polymerase with strand displacement activity. The RNA portion of the hybridized primer in a heteroduplex can be cleaved by RNase H to free a portion of the primer binding site. A second linear composite amplification primer can hybridize to the freed primer binding site, and can be extended along the target DNA strand. The previously synthesized primer extension product (amplification product) can be displaced by the newly extended primer. Repeated cycles of primer hybridization, primer extension by strand displacement DNA polymerase, and cleavage of the RNA portion of the hybridized primer can generate multiple copies of a target nucleic acid (Figure 5).

[0089] **Other Amplification Methods**

[0090] Some aspects of the invention comprise the amplification of polynucleotide molecules or sequences within the polynucleotide molecules. Amplification generally refers to a method that can result in the formation of one or more copies of a nucleic acid or polynucleotide molecule or in the formation of one or more copies of the complement of a nucleic acid or polynucleotide molecule. Amplifications can be used in the invention, for example, to amplify or analyze a polynucleotide bound to a solid surface. The amplifications can be performed, for example, after archiving the samples in order to analyze the archived polynucleotide.

QIAGEN Exhibit 1002
Page 841

[0091] In some aspects of the invention, exponential amplification of nucleic acids or polynucleotides is used. These methods often depend on the product catalyzed formation of multiple copies of a nucleic acid or polynucleotide molecule or its complement. The amplification products are sometimes referred to as "amplicons." One such method for the enzymatic amplification of specific double stranded sequences of DNA is polymerase chain reaction (PCR). This in vitro amplification procedure is based on repeated cycles of denaturation, oligonucleotide primer annealing, and primer extension by thermophilic template dependent polynucleotide polymerase, resulting in the exponential increase in copies of the desired sequence of the polynucleotide analyte flanked by the primers. The two different PCR primers, which anneal to opposite strands of the DNA, are positioned so that the polymerase catalyzed extension product of one primer can serve as a template strand for the other, leading to the accumulation of a discrete double stranded fragment whose length is defined by the distance between the 5′ ends of the oligonucleotide primers. Other amplification techniques that can be used in the methods of the provided invention include, e.g., AFLP (amplified fragment length polymorphism) PCR (see e.g.: Vos et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-14), allele-specific PCR (see e.g., Saiki R K, Bugawan T L, Horn G T, Mullis K B, Erlich H A (1986). Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes Nature 324: 163-166), Alu PCR, assembly PCR (see e.g., Stemmer W P, Crameri A, Ha K D, Brennan T M, Heyneker H L (1995). Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides Gene 164: 49-53), assymetric PCR (see e.g., Saiki R K supra), colony PCR, helicase dependent PCR (see e.g., Myriam Vincent, Yan Xu and Huimin Kong (2004). Helicase-dependent isothermal DNA amplification EMBO reports 5 (8): 795-800), hot start PCR, inverse PCR (see e.g., Ochman H, Gerber A S, Hartl D L. Genetics. 1988 November; 120(3):621-3), in situ PCR, intersequence-specific PCR or IS SR PCR, digital PCR, linear-after-the-exponential-PCR or Late PCR (see e.g., Pierce K E and Wangh L T (2007). Linear-after-the-exponential polymerase chain reaction and allied technologies Real-time detection strategies for rapid, reliable diagnosis from single cells Methods Mol. Med. 132: 65-85), long PCR, nested PCR, real-time PCR, duplex PCR, multiplex PCR, quantitative PCR, or single cell PCR.

[0092] Another method for amplification involves amplification of a single stranded polynucleotide using a single oligonucleotide primer. The single stranded polynucleotide that is to be amplified contains two non-contiguous sequences that are substantially or

completely complementary to one another and, thus, are capable of hybridizing together to form a stem-loop structure. This single stranded polynucleotide already may be part of a polynucleotide analyte or may be created as the result of the presence of a polynucleotide analyte.

[0093] Another method for achieving the result of an amplification of nucleic acids is known as the ligase chain reaction (LCR). This method uses a ligase enzyme to join pairs of preformed nucleic acid probes. The probes hybridize with each complementary strand of the nucleic acid analyte, if present, and ligase is employed to bind each pair of probes together resulting in two templates that can serve in the next cycle to reiterate the particular nucleic acid sequence.

[0094] Another method for achieving nucleic acid amplification is the nucleic acid sequence based amplification (NASBA). This method is a promoter-directed, enzymatic process that induces in vitro continuous, homogeneous and isothermal amplification of a specific nucleic acid to provide RNA copies of the nucleic acid. The reagents for conducting NASBA include a first DNA primer with a 5′-tail comprising a promoter, a second DNA primer, reverse transcriptase, RNase-H, T7 RNA polymerase, NTP's and dNTP's.

[0095] Another method for amplifying a specific group of nucleic acids is the Q-beta-replicase method, which relies on the ability of Q-beta-replicase to amplify its RNA substrate exponentially. The reagents for conducting such an amplification include "midi-variant RNA" (amplifiable hybridization probe), NTP's, and Q-beta-replicase.

[0096] Another method for amplifying nucleic acids is known as 3SR and is similar to NASBA except that the RNase-H activity is present in the reverse transcriptase. Amplification by 3SR is an RNA specific target method whereby RNA is amplified in an isothermal process combining promoter directed RNA polymerase, reverse transcriptase and RNase H with target RNA. See for example Fahy et al. PCR Methods Appl. 1:25-33 (1991).

[0097] Another method for amplifying nucleic acids is the Transcription Mediated Amplification (TMA) used by Gen-Probe. The method is similar to NASBA in utilizing two enzymes in a self-sustained sequence replication. See U.S. Pat. No. 5,299,491 herein incorporated by reference.

[0098] Another method for amplification of nucleic acids is Strand Displacement Amplification (SDA) (Westin et al 2000, Nature Biotechnology, 18, 199-202; Walker et al 1992, Nucleic Acids Research, 20, 7, 1691-1696), which is an isothermal amplification technique based upon the ability of a restriction endonuclease such as HincII or BsoBI to nick the unmodified strand of a hemiphosphorothioate form of its recognition site, and

the ability of an exonuclease deficient DNA polymerase such as Klenow exo minus polymerase, or Bst polymerase, to extend the 3′-end at the nick and displace the downstream DNA strand. Exponential amplification results from coupling sense and antisense reactions in which strands displaced from a sense reaction serve as targets for an antisense reaction and vice versa.

[0099] Another method for amplification of nucleic acids is Rolling Circle Amplification (RCA) (Lizardi et al. 1998, Nature Genetics, 19:225-232). RCA can be used to amplify single stranded molecules in the form of circles of nucleic acids. In its simplest form, RCA involves the hybridization of a single primer to a circular nucleic acid. Extension of the primer by a DNA polymerase with strand displacement activity results in the production of multiple copies of the circular nucleic acid concatenated into a single DNA strand.

[00100] In some embodiments of the invention, RCA is coupled with ligation. For example, a single oligonucleotide can be used both for ligation and as the circular template for RCA. This type of polynucleotide can be referred to as a "padlock probe" or a "RCA probe." For a padlock probe, both termini of the oligonucleotide contain sequences complementary to a domain within a nucleic acid sequence of interest. The first end of the padlock probe is substantially complementary to a first domain on the nucleic acid sequence of interest, and the second end of the padlock probe is substantially complementary to a second domain, adjacent to the first domain near the first domain. Hybridization of the oligonucleotide to the target nucleic acid results in the formation of a hybridization complex. Ligation of the ends of the padlock probe results in the formation of a modified hybridization complex containing a circular polynucleotide. In some cases, prior to ligation, a polymerase can fill in the gap by extending one end of the padlock probe. The circular polynucleotide thus formed can serve as a template for RCA that, with the addition of a polymerase, results in the formation of an amplified product nucleic acid. The methods of the invention described herein can produce amplified products with defined sequences on both the 5′- and 3′-ends. Such amplified products can be used as padlock probes.

[00101] Some aspects of the invention utilize the linear amplification of nucleic acids or polynucleotides. Linear amplification generally refers to a method that involves the formation of one or more copies of the complement of only one strand of a nucleic acid or polynucleotide molecule, usually a nucleic acid or polynucleotide analyte. Thus, the primary difference between linear amplification and exponential amplification is that in the latter process, the product serves as substrate for the formation of more product,

whereas in the former process the starting sequence is the substrate for the formation of product but the product of the reaction, i.e. the replication of the starting template, is not a substrate for generation of products. In linear amplification the amount of product formed increases as a linear function of time as opposed to exponential amplification where the amount of product formed is an exponential function of time.

[00102] In some embodiments, amplification methods can be solid-phase amplification, polony amplification, colony amplification, emulsion PCR, bead RCA, surface RCA, surface SDA, etc., as will be recognized by one of skill in the art. In some embodiments, amplification methods that results in amplification of free DNA molecules in solution or tethered to a suitable matrix by only one end of the DNA molecule can be used. Methods that rely on bridge PCR, where both PCR primers are attached to a surface (see, e.g., WO 2000/018957 and Adessi et al., Nucleic Acids Research (2000): 28(20): E87) can be used. In some cases the methods of the invention can create a "polymerase colony technology," or "polony." referring to a multiplex amplification that maintains spatial clustering of identical amplicons (see Harvard Molecular Technology Group and Lipper Center for Computational Genetics website). These include, for example, in situ polonies (Mitra and Church, Nucleic Acid Research 27, e34, Dec. 15, 1999), in situ rolling circle amplification (RCA) (Lizardi et al., Nature Genetics 19, 225, July 1998), bridge PCR (U.S. Pat. No. 5,641,658), picotiter PCR (Leamon et al., Electrophoresis 24, 3769, November 2003), and emulsion PCR (Dressman et al., PNAS 100, 8817, Jul. 22, 2003). The methods of the invention provide new methods for generating and using polonies.

Other Downstream Applications

[00103] An important aspect of the invention is that the methods and compositions disclosed herein can be efficiently, cost-effectively, and with the minimal loss of biological material, utilized for downstream analyses.

[00104] In particular embodiments, the oligonucleotide-attached nucleic acid molecule is further modified, such as by cloning, including by incorporation of the modified molecule into a vector, said incorporation occurring at ends in the modified molecule generated by endonuclease cleavage within the inverted repeat.

[00105] In specific embodiments, the oligonucleotide-attached nucleic acid molecule is immobilized to a solid support, such as non-covalently or covalently.

[00106] The libraries of amplified, non-amplified, non-targeted nucleic acids, as well as enriched copies of selected sequence regions of interest are useful for massively parallel

sequencing (Next Generation Sequencing methods), multiplexed quantification of large sets of sequence regions of interest, such as by high density qPCR arrays and other highly parallel quantification platforms (selective massively parallel target pre-amplification), as well as generation of libraries with enriched population of sequence regions of interest.

Sequencing

[00107] In one embodiment, the invention provides for products ready for amplification in preparation for sequencing. In some embodiments, the target polynucleotides are pooled followed by sequencing one or more polynucleotides in the pool. Sequencing methods utilizing adaptor incorporated sequences are well known in the art and are further described, for example, in US Patent No.s 8,053,192 and 8,017,335.

[00108] Sequencing processes are generally template dependent. Nucleic acid sequence analysis that employs template dependent synthesis identifies individual bases, or groups of bases as they are added during a template mediated synthesis reaction, such as a primer extension reaction, where the identity of the base is complementary to the template sequence to which the primer sequence is hybridized during synthesis. Other such processes include ligation driven processes, where oligonucleotides or polynucleotides are complexed with an underlying template sequence, in order to identify the sequence of nucleotides in that sequence. Typically, such processes are enzymatically mediated using nucleic acid polymerases, such as DNA polymerases, RNA polymerases, reverse transcriptases, and the like, or other enzymes such as in the case of ligation driven processes, e.g., ligases.

[00109] Sequence analysis using template dependent synthesis can include a number of different processes. For example, in the ubiquitously practiced four-color Sanger sequencing methods, a population of template molecules is used to create a population of complementary fragment sequences. Primer extension is carried out in the presence of the four naturally occurring nucleotides, and with a sub-population of dye labeled terminator nucleotides, e.g., dideoxyribonucleotides, where each type of terminator (ddATP, ddGTP, ddTTP, ddCTP) includes a different detectable label. As a result, a nested set of fragments is created where the fragments terminate at each nucleotide in the sequence beyond the primer, and are labeled in a manner that permits identification of the terminating nucleotide. The nested fragment population is then subjected to size based separation, e.g., using capillary electrophoresis, and the labels associated with each different sized fragment is identified to identify the terminating nucleotide. As a result, the sequence of labels moving past a detector in the separation system provides a direct

readout of the sequence information of the synthesized fragments, and by complementarity, the underlying template (See, e.g., U.S. Pat. No. 5,171,534, incorporated herein by reference in its entirety for all purposes).

[00110] Other examples of template dependent sequencing methods include sequence by synthesis processes, where individual nucleotides are identified iteratively, as they are added to the growing primer extension product.

[00111] Pyrosequencing is an example of a sequence by synthesis process that identifies the incorporation of a nucleotide by assaying the resulting synthesis mixture for the presence of by-products of the sequencing reaction, namely pyrophosphate. In particular, a primer/template/polymerase complex is contacted with a single type of nucleotide. If that nucleotide is incorporated, the polymerization reaction cleaves the nucleoside triphosphate between the α and β phosphates of the triphosphate chain, releasing pyrophosphate. The presence of released pyrophosphate is then identified using a chemiluminescent enzyme reporter system that converts the pyrophosphate, with AMP, into ATP, then measures ATP using a luciferase enzyme to produce measurable light signals. Where light is detected, the base is incorporated, where no light is detected, the base is not incorporated. Following appropriate washing steps, the various bases are cyclically contacted with the complex to sequentially identify subsequent bases in the template sequence. See, e.g., U.S. Pat. No. 6,210,891, incorporated herein by reference in its entirety for all purposes).

[00112] In related processes, the primer/template/polymerase complex is immobilized upon a substrate and the complex is contacted with labeled nucleotides. The immobilization of the complex may be through the primer sequence, the template sequence and/or the polymerase enzyme, and may be covalent or noncovalent. For example, immobilization of the complex can be via a linkage between the polymerase or the primer and the substrate surface. A variety of types of linkages are useful for this attachment, including, e.g., provision of biotinylated surface components, using e.g., biotin-PEG-silane linkage chemistries, followed by biotinylation of the molecule to be immobilized, and subsequent linkage through, e.g., a streptavidin bridge. Other synthetic coupling chemistries, as well as non-specific protein adsorption can also be employed for immobilization. In alternate configurations, the nucleotides are provided with and without removable terminator groups. Upon incorporation, the label is coupled with the complex and is thus detectable. In the case of terminator bearing nucleotides, all four different nucleotides, bearing individually identifiable labels, are contacted with the

complex. Incorporation of the labeled nucleotide arrests extension, by virtue of the presence of the terminator, and adds the label to the complex. The label and terminator are then removed from the incorporated nucleotide, and following appropriate washing steps, the process is repeated. In the case of non-terminated nucleotides, a single type of labeled nucleotide is added to the complex to determine whether it will be incorporated, as with pyrosequencing. Following removal of the label group on the nucleotide and appropriate washing steps, the various different nucleotides are cycled through the reaction mixture in the same process. See, e.g., U.S. Pat. No. 6,833,246, incorporated herein by reference in its entirety for all purposes). For example, the Illumina Genome Analyzer System is based on technology described in WO 98/44151, hereby incorporated by reference, wherein DNA molecules are bound to a sequencing platform (flow cell) via an anchor probe binding site (otherwise referred to as a flow cell binding site) and amplified in situ on a glass slide. The DNA molecules are then annealed to a sequencing primer and sequenced in parallel base-by-base using a reversible terminator approach. Typically, the Illumina Genome Analyzer System utilizes flow-cells with 8 channels, generating sequencing reads of 18 to 36 bases in length, generating >1.3 Gbp of high quality data per run (see www.illumina.com).

[00113] In yet a further sequence by synthesis process, the incorporation of differently labeled nucleotides is observed in real time as template dependent synthesis is carried out. In particular, an individual immobilized primer/template/polymerase complex is observed as fluorescently labeled nucleotides are incorporated, permitting real time identification of each added base as it is added. In this process, label groups are attached to a portion of the nucleotide that is cleaved during incorporation. For example, by attaching the label group to a portion of the phosphate chain removed during incorporation, i.e., a β, γ, or other terminal phosphate group on a nucleoside polyphosphate, the label is not incorporated into the nascent strand, and instead, natural DNA is produced. Observation of individual molecules typically involves the optical confinement of the complex within a very small illumination volume. By optically confining the complex, one creates a monitored region in which randomly diffusing nucleotides are present for a very short period of time, while incorporated nucleotides are retained within the observation volume for longer as they are being incorporated. This results in a characteristic signal associated with the incorporation event, which is also characterized by a signal profile that is characteristic of the base being added. In related aspects, interacting label components, such as fluorescent resonant energy transfer

(FRET) dye pairs, are provided upon the polymerase or other portion of the complex and the incorporating nucleotide, such that the incorporation event puts the labeling components in interactive proximity, and a characteristic signal results, that is again, also characteristic of the base being incorporated (See, e.g., U.S. Pat. Nos. 6,056,661, 6,917,726, 7,033,764, 7,052,847, 7,056,676, 7,170,050, 7,361,466, 7,416,844 and Published U.S. Patent Application No. 2007-0134128, the full disclosures of which are hereby incorporated herein by reference in their entirety for all purposes).

[00114] In some embodiments, the nucleic acids in the sample can be sequenced by ligation. This method uses a DNA ligase enzyme to identify the target sequence, for example, as used in the polony method and in the SOLiD technology (Applied Biosystems, now Invitrogen). In general, a pool of all possible oligonucleotides of a fixed length is provided, labeled according to the sequenced position. Oligonucleotides are annealed and ligated; the preferential ligation by DNA ligase for matching sequences results in a signal corresponding to the complementary sequence at that position.

[00115] In some embodiments, sequencing comprises extension of a sequencing primer comprising a sequence hybridizable to at least a portion of the complement of the first adaptor oligonucleotide. In some embodiments, sequencing comprises extension of a sequencing primer comprising a sequence hybridizable to at least a portion of the complement of the second adaptor oligonucleotide. A sequencing primer may be of any suitable length, such as about, less than about, or more than about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, or more nucleotides, any portion or all of which may be complementary to the corresponding target sequence (e.g. about, less than about, or more than about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, or more nucleotides). In some embodiments, sequencing comprises a calibration step, wherein the calibration is based on each of the nucleotides at one or more nucleotide positions in the barcode sequences. Calibration can be useful in processing the sequencing data, for example, by facilitating or increasing the accuracy of identifying a base at a given position in the sequence.

[00116] In some embodiments, accurate identification of the sample from which a target polynucleotide is derived is based on at least a portion of the sequence obtained for the target polynucleotide and is at least 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.8%, 99.85%, 99.9%, 99.95%, 99.99%, or more accurate. In some embodiments, the sample source of a target polynucleotide is identified based on a single barcode contained in the sequence. In some embodiments, accuracy can be increased by identifying the source of

a target polynucleotide using two or more barcodes contained in the sequence. Multiple barcodes can be joined to a target polynucleotide by the incorporation of multiple barcodes into a single adaptor/primer to which a target polynucleotide is joined, and/or by joining two or more adaptors/primers having one or more barcodes to a target polynucleotide. In some embodiments, the identity of the sample source of a target polynucleotide comprising two or more barcode sequences may be accurately determined using only one of the barcode sequences that it comprises. In general, accurate identification of a sample from which a target polynucleotide is derived comprises correct identification of a sample source from among two or more samples in a pool, such as about, less than about, or more than about 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 16, 20, 24, 28, 32, 36, 40, 50, 60, 70, 80, 90, 100, 128, 192, 384, 500, 1000 or more samples in a pool.

[00117] In some embodiments, the methods are useful for preparing target polynucleotide(s) for sequencing by the sequencing by ligation methods commercialized by Applied Biosystems (e.g., SOLiD sequencing). In general, double stranded fragment polynucleotides can be prepared by the methods of the present invention, and then incorporated into a water-in-oil emulsion along with polystyrene beads and amplified, for example by PCR. In some cases, alternative amplification methods can be employed in the water-in-oil emulsion such as any of the methods provided herein. The amplified product in each water microdroplet formed by the emulsion can interact, bind, or hybridize with the one or more beads present in that microdroplet leading to beads with a plurality of amplified products of substantially one sequence. When the emulsion is broken, the beads float to the top of the sample and are placed onto an array. The methods can include a step of rendering the nucleic acid bound to the beads single-stranded or partially single stranded. Sequencing primers are then added along with a mixture of four different fluorescently labeled oligonucleotide probes. The probes bind specifically to the two bases in the polynucleotide to be sequenced immediately adjacent and 3′ of the sequencing primer to determine which of the four bases are at those positions. After washing and reading the fluorescence signal from the first incorporated probe, a ligase is added. The ligase cleaves the oligonucleotide probe between the fifth and sixth bases, removing the fluorescent dye from the polynucleotide to be sequenced. The whole process is repeated using a different sequence primer until all of the intervening positions in the sequence are imaged. The process allows the simultaneous reading of millions of DNA fragments in a 'massively parallel' manner. This 'sequence-by-ligation' technique uses probes that encode for two bases rather than just one allowing

error recognition by signal mismatching, leading to increased base determination accuracy.

[00118] In some embodiments, the methods are useful for preparing target polynucleotide(s) for sequencing by synthesis using the methods commercialized by 454/Roche Life Sciences including but not limited to the methods and apparatus described in Margulies et al., Nature (2005) 437:376-380 (2005); and U.S. Pat. Nos. 7,244,559; 7,335,762; 7,211,390; 7,244,567; 7,264,929; and 7,323,305. In general, double stranded fragment polynucleotides can be prepared by the methods of the present invention, immobilized onto beads, and compartmentalized in a water-in-oil PCR emulsion. In some cases, alternative amplification methods can be employed in the water-in-oil emulsion such as any of the methods provided herein. When the emulsion is broken, amplified fragments remain bound to the beads. The methods can include a step of rendering the nucleic acid bound to the beads single stranded or partially single stranded. The beads can be enriched and loaded into wells of a fiber optic slide so that there is approximately 1 bead in each well. Nucleotides are flowed across and into the wells in a fixed order in the presence of polymerase, sulfhydrolase, and luciferase. Addition of nucleotides complementary to the target strand can result in a chemiluminescent signal that is recorded, such as by a camera. The combination of signal intensity and positional information generated across the plate allows software to determine the DNA sequence.

[00119] In some embodiments, the methods are useful for preparing target polynucleotide(s) for sequencing by the methods commercialized by Helicos BioSciences Corporation (Cambridge, Mass.) as described in U.S. application Ser. No. 11/167,046, and U.S. Pat. Nos. 7,501,245; 7,491,498; 7,276,720; and in U.S. Patent Application Publication Nos. US20090061439; US20080087826; US20060286566; US20060024711; US20060024678; US20080213770; and US20080103058. In general, double stranded fragment polynucleotides can be prepared by the methods of the present invention, and then immobilized onto a flow-cell surface. The methods can include a step of rendering the nucleic acid bound to the flow-cell surface stranded or partially single stranded. Polymerase and labeled nucleotides are then flowed over the immobilized DNA. After fluorescently labeled nucleotides are incorporated into the DNA strands by a DNA polymerase, the surface is illuminated with a laser, and an image is captured and processed to record single molecule incorporation events to produce sequence data.

[00120] In some embodiments, the methods are useful for preparing target polynucleotide(s) for sequencing by the sequencing by ligation methods commercialized by Dover Systems. Generally, double stranded fragment polynucleotides can be prepared by the methods of the present invention. The polynucleotides can then be amplified in an emulsion in the presence of magnetic beads. Any amplification methods can be employed in the water-in-oil emulsion such as any of the methods provided herein. The resulting beads with immobilized clonal polynucleotide polonies are then purified by magnetic separation, capped, amine functionalized, and covalently immobilized in a series of flow cells. The methods can include a step of rendering the nucleic acid bound to the flow-cell surface stranded or partially single stranded. Then, a series of anchor primers are flowed through the cell, where they hybridize to the synthetic oligonucleotide sequences at the 3′ or 5′ end of proximal or distal genomic DNA tags. Once an anchor primer is hybridized, a mixture of fully degenerate nonanucleotides ('nonamers') and T4 DNA ligase is flowed into the cell; each of the nonamer mixture's four components is labeled with one of four fluorophores, which correspond to the base type at the query position. The fluorophore-tagged nonamers selectively ligate onto the anchor primer, providing a fluorescent signal that identifies the corresponding base on the genomic DNA tag. Once the probes are ligated, fluorescently labeling the beads, the array is imaged in four colors. Each bead on the array will fluoresce in only one of the four images, indicating whether there is an A, C, G, or T at the position being queried. After imaging, the array of annealed primer-fluorescent probe complex, as well as residual enzyme, are chemically striped using guanidine HCl and sodium hydroxide. After each cycle of base reads at a given position have been completed, and the primer-fluorescent probe complex has been stripped, the anchor primer is replaced, and a new mixture of fluorescently tagged nonamers is introduced, for which the query position is shifted one base further into the genomic DNA tag. Seven bases are queried in this fashion, with the sequence performed from the 5′ end of the proximal tag, followed by six base reads with a different anchor primer from the 3′ end of the proximal tag, for a total of 13 base pair reads for this tag. This sequence is then repeated for the 5′ and 3′ ends of the distal tag, resulting in another 13 base pair reads. The ultimate result is a read length of 26 bases (thirteen from each of the paired tags). However, it is understood that this method is not limited to 26 base read lengths.

[00121] In some embodiments, the methods are useful for preparing target polynucleotide(s) from selectively enriched populations of specific sequence regions of

interest in a strand-specific manner for sequencing by the methods well known in the art and further described below.

[00122] For example the methods are useful for sequencing by the method commercialized by Illumina as described U.S. Pat. Nos. 5,750,341; 6,306,597; and 5,969,119. In general, double stranded fragment polynucleotides can be prepared by the methods of the present invention to produce amplified nucleic acid sequences tagged at one (e.g., (A)/(A′) or both ends (e.g., (A)/(A′) and (C)/(C′)). In some cases, single stranded nucleic acid tagged at one or both ends is amplified by the methods of the present invention (e.g., by SPIA or linear PCR). The resulting nucleic acid is then denatured and the single stranded amplified polynucleotides are randomly attached to the inside surface of flow-cell channels. Unlabeled nucleotides are added to initiate solid-phase bridge amplification to produce dense clusters of double-stranded DNA. To initiate the first base sequencing cycle, four labeled reversible terminators, primers, and DNA polymerase are added. After laser excitation, fluorescence from each cluster on the flow cell is imaged. The identity of the first base for each cluster is then recorded. Cycles of sequencing are performed to determine the fragment sequence one base at a time. For paired-end sequencing, such as for example, when the polynucleotides are labeled at both ends by the methods of the present invention, sequencing templates can be regenerated in-situ so that the opposite end of the fragment can also be sequenced.

[00123] In some embodiments, the methods are useful for preparing target polynucleotide(s) for sequencing by the methods commercialized by Pacific Biosciences as described in U.S. Pat. Nos. 7,462,452; 7,476,504; 7,405,281; 7,170,050; 7,462,468; 7,476,503; 7,315,019; 7,302,146; 7,313,308; and US Application Publication Nos. US20090029385; US20090068655; US20090024331; and US20080206764. In general, double stranded fragment polynucleotides can be prepared by the methods of the present invention. The polynucleotides can then be immobilized in zero mode waveguide arrays. The methods may include a step of rendering the nucleic acid bound to the waveguide arrays single stranded or partially single stranded. Polymerase and labeled nucleotides are added in a reaction mixture, and nucleotide incorporations are visualized via fluorescent labels attached to the terminal phosphate groups of the nucleotides. The fluorescent labels are clipped off as part of the nucleotide incorporation. In some cases, circular templates are utilized to enable multiple reads on a single molecule.

[00124] Another example of a sequencing technique that can be used in the methods of the provided invention is nanopore sequencing (see e.g. Soni G V and Meller A. (2007) Clin

Chem 53: 1996-2001). A nanopore can be a small hole of the order of 1 nanometer in diameter. Immersion of a nanopore in a conducting fluid and application of a potential across it can result in a slight electrical current due to conduction of ions through the nanopore. The amount of current that flows is sensitive to the size of the nanopore. As a DNA molecule passes through a nanopore, each nucleotide on the DNA molecule obstructs the nanopore to a different degree. Thus, the change in the current passing through the nanopore as the DNA molecule passes through the nanopore can represent a reading of the DNA sequence.

[00125] Another example of a sequencing technique that can be used in the methods of the provided invention is semiconductor sequencing provided by Ion Torrent (e.g., using the Ion Personal Genome Machine (PGM)). Ion Torrent technology can use a semiconductor chip with multiple layers, e.g., a layer with micro-machined wells, an ion-sensitive layer, and an ion sensor layer. Nucleic acids can be introduced into the wells, e.g., a clonal population of single nucleic can be attached to a single bead, and the bead can be introduced into a well. To initiate sequencing of the nucleic acids on the beads, one type of deoxyribonucleotide (e.g., dATP, dCTP, dGTP, or dTTP) can be introduced into the wells. When one or more nucleotides are incorporated by DNA polymerase, protons (hydrogen ions) are released in the well, which can be detected by the ion sensor. The semiconductor chip can then be washed and the process can be repeated with a different deoxyribonucleotide. A plurality of nucleic acids can be sequenced in the wells of a semiconductor chip. The semiconductor chip can comprise chemical-sensitive field effect transistor (chemFET) arrays to sequence DNA (for example, as described in U.S. Patent Application Publication No. 20090026082). Incorporation of one or more triphosphates into a new nucleic acid strand at the 3′ end of the sequencing primer can be detected by a change in current by a chemFET. An array can have multiple chemFET sensors.

Genetic Analysis

[00126] The methods of the present invention can be used in the analysis of genetic information of selective genomic regions of interest as well as genomic regions which may interact with the selective region of interest. Amplification methods as disclosed herein can be used in the devices, kits, and methods known to the art for genetic analysis, such as, but not limited to those found in U.S. Pat. Nos. 6,449,562, 6,287,766, 7,361,468, 7,414,117, 6,225,109, and 6,110,709. In some cases, amplification methods of the present invention can be used to amplify target nucleic acid for DNA hybridization studies to

determine the presence or absence of polymorphisms. The polymorphisms, or alleles, can be associated with diseases or conditions such as genetic disease. In other cases the polymorphisms can be associated with susceptibility to diseases or conditions, for example, polymorphisms associated with addiction, degenerative and age related conditions, cancer, and the like. In other cases, the polymorphisms can be associated with beneficial traits such as increased coronary health, or resistance to diseases such as HIV or malaria, or resistance to degenerative diseases such as osteoporosis, Alzheimer's or dementia.

Kits

[00127] Any of the compositions described herein may be comprised in a kit. In a non-limiting example the kit, in suitable container means, comprises: one or more RNA DNA stem-loop oligonucleotides. The kit can further contain enzymes and/or reagents useful for ligation, cleavage and or amplification. The kit can contain a DNA-polymerase. The kit can contain reagents for amplification, for example reagents useful for single primer isothermal amplification methods. The kit can further optionally contain reagents for sequencing, for example, reagents useful for next-generation massively parallel sequencing methods.

[00128] The containers of the kits can generally include at least one vial, test tube, flask, bottle, syringe or other containers, into which a component may be placed, and preferably, suitably aliquoted. Where there is more than one component in the kit, the kit also can generally contain a second, third or other additional container into which the additional components can be separately placed. However, various combinations of components can be comprised in a container.

[00129] When the components of the kit are provided in one or more liquid solutions, the liquid solution can be an aqueous solution. However, the components of the kit can be provided as dried powder(s). When reagents and/or components are provided as a dry powder, the powder can be reconstituted by the addition of a suitable solvent.

[00130] A kit can include instructions for employing the kit components as well the use of any other reagent not included in the kit. Instructions can include variations that can be implemented.

[00131] In one aspect, the invention provides kits containing any one or more of the elements disclosed in the above methods and compositions. In some embodiments, a kit comprises a composition of the invention, in one or more containers. In some

embodiments, the invention provides kits comprising adaptors, primers, and/or other oligonucleotides described herein. In some embodiments, the kit further comprises one or more of: (a) a DNA ligase, (b) a DNA-dependent DNA polymerase, (c) an RNA-dependent DNA polymerase, (d) random primers, (e) primers comprising at least 4 thymidines at the 3' end, (f) a DNA endonuclease, (g) a DNA-dependent DNA polymerase having 3' to 5' exonuclease activity, (h) a plurality of primers, each primer having one of a plurality of selected sequences, (i) a DNA kinase, (j) a DNA exonuclease, (k) magnetic beads, (l) an enzyme comprising RNase H activity, (m) an RNA ligase, and (n) one or more buffers suitable for one or more of the elements contained in said kit. The adaptors, primers, other oligonucleotides, and reagents can be, without limitation, any of those described above. Elements of the kit can further be provided, without limitation, in any suitable amounts and/or using any of the combinations (such as in the same kit or same container) described above or any other suitable combination known in the art. The kits may further comprise additional agents, such as those described above, for use according to the methods of the invention. The kit elements can be provided in any suitable container, including but not limited to test tubes, vials, flasks, bottles, ampules, syringes, or the like. The agents can be provided in a form that may be directly used in the methods of the invention, or in a form that requires preparation prior to use, such as in the reconstitution of lyophilized agents. Agents may be provided in aliquots for single-use or as stocks from which multiple uses, such as in a number of reaction, may be obtained.

EXAMPLES

Example 1: *Whole transcriptome amplification using stem-loop chimeric amplification primer and the corresponding linear (non stem-loop) chimeric amplification primer*

[00132] In this example cDNA was generated from total RNA (50pg total HeLa total RNA) and control sample without added total RNA (no template control, NTC). The reagents and protocols for first and second strand cDNA were those of Ovation Pico (NuGEN Technologies). The chimeric first strand primer used for this example comprises a combination of a chimeric primer with a random sequence at the 3'-end (Primer A random, in Table 1) and a chimeric primer with an oligo dT at the 3'-portion (Primer A dT, in Table 1) and further comprises the same 5'-RNA tail. The 5'-RNA tail of the first strand cDNA primer comprises the same sequence as the single chimeric

amplification primer used in the isothermal single primer amplification step, as shown below.

[00133] The double stranded cDNA with an RNA/DNA heteroduplex at one end generated by the first and second cDNA synthesis steps above was purified as per the manufacturer instruction, and subjected to the single primer isothermal amplification (SPIA) step, using the reagents and conditions of the Ovation Pico kit (NuGEN Technologies) and either the stem-loop chimeric amplification primer (Stem-Loop amplification primer A, Table 1) or the corresponding linear chimeric amplification primer (Linear amplification primer A, Table1). The amplification primer generated from the stem-loop amplification primer as well as the linear amplification primers are chimeric primers comprising a DNA sequence at the 3'-end and an RNA sequence at the 5'-end, which are the same sequence as the corresponding tail of the first strand primers. SPIA amplification was carried out according to manufacturer instructions. SPIA amplification products were purified using MinElute columns (QIAGEN) using the manufacturer instructions and reagents.

Table 1

Primer ID	Sequence
Primer A random	*cuucuauaguuuagguaacuuuguguuuga*NNNNNN
Primer A dT	*cuucuauaguuuagguaacuuuguguuuga*TTTTTTTTTTTTTNN
Stem-loop amplification primer A	*ucaaacacuucuauaguuuagguaacuuug*TGTTTGA
Linear amplification primer A	*cuucuauaguuuagguaacuuug* TGTTTGA

[00134] Table 1 illustrates the primer sequences with ribonucleotides appearing in lower case and deoxyribonucleotides appearing in upper case. "N" represents a random nucleotide.

[00135] The yields of transcriptome amplification employing the stem-loop chimeric amplification and the corresponding linear chimeric amplification primer using the method of the invention are summarized in Table 2.

Table 2.

Sample	*Amplification primer*	*µg*	*Avg µg (Std Dev)*
50 pg RNA	Linear amplification primer A	4.87	**4.50 (0.32)**
50 pg RNA	Linear amplification primer A	4.31	
50 pg RNA	Linear amplification primer A	4.32	
NTC	Linear amplification primer A	0.73	1.36 (1.25)
NTC	Linear amplification primer A	1.99	
50 pg RNA	Stem-loop amplification primer A	5.61	**5.51 (0.11)**
50 pg RNA	Stem-loop amplification primer A	5.52	
50 pg RNA	Stem-loop amplification primer A	5.38	
NTC	Stem-loop amplification primer A	0.85	0.90 (0.11)
NTC	Stem-loop amplification primer A	0.96	

[00136] The results demonstrate the higher amplification yields obtained in reactions employing the stem-loop chimeric amplification primer of the invention as compared to the corresponding linear chimeric amplification primer. This is likely due to employing a pro-primer which is activated during the initial reaction thus leading to lower production of non-specific products as indicated by the lower yield of amplification product in the no target control.

Example 2: *Whole genome amplification (genomic DNA derived from FFPE sample) using stem-loop chimeric amplification primer and the corresponding linear (non stem-loop) chimeric amplification primer*

[00137] Whole genome amplification (WGA) was performed as described in the Ovation WGA FFPE System Users Guide (Part #6200, NuGEN Technologies Inc). Briefly, 50 or 100 ng of isolated FFPE genomic DNA was fragmented, end-repaired and adenosine-tailed to generate 5′ phosphorylated/3′ A-overhang fragments. Adaptors suitable for amplification using the method of the invention were utilized for this example (Adaptor

B, Table 3). Adaptor B comprises two oligonucleotides that when annealed form a single T-overhang to aid in ligation efficiency and a long 3' overhang that comprises the priming site for the SPIA chimeric amplification primer. After the adaptors were ligated to the fragments, the reactions were purified to remove excess adaptors using Agencourt® RNAClean® XP beads. The resulting adaptor ligated fragments were isothermally amplified using the enzyme mixture and protocol from the kit and purified using an appropriate purification method, such as Qiagen® spin columns. Either the stem-loop chimeric amplification primer (Stem-loop amplification primer B, Table 3) or the corresponding linear chimeric amplification primer (Linear amplification primer B, Table 3), were employed in the isothermal single primer amplification step.

[00138] Amplification products were purified as per the manufacturer instructions and the yields of whole genome amplification reactions were determined using Nanodrop (Thermo Scientific; Figure 6). The size-distribution of the amplification products from these reactions were determined using BioAnalyze (Agilent) (Figure 7).

Table 3

Oligonucleotide	Sequence
Adaptor B	5'-cgtagactcaTctgcctatagtgagtcgtattacc-3' |||||||||||||||| 3'-tgcatctgagtAgacgga-5'
Stem-loop chimeric amplification primer B	5'-rUrCrUrGrCrCrUrGrGrUrArArUrArCrGrArCrUrCrArCrUrArUAGGCAGA-3'
Linear chimeric amplification primer B	5'-rGrGrUrArArUrArCrGrArCrUrCrArCrUrArUAGGCAGA-3'

[00139] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of preparing a target nucleic acid molecule for further processing, comprising:

(a) attaching an RNA/DNA chimeric stem-loop oligonucleotide having a portion of its stem and at least a portion of its loop comprise of RNA to said target nucleic acid molecule thereby producing a stem-loop-target complex; and

(b) contacting said stem-loop-target complex with an RNase that cleaves RNA from a RNA/DNA heteroduplex, thereby preparing the stem-loop-target complex for further processing.

2. The method of claim 1, wherein the RNase is RNase H.

3. The method of claims 1-2, wherein further processing comprises nucleic acid sequencing.

4. The method of claim 3, wherein sequencing comprises single molecule sequencing.

5. The method of claim 3, wherein sequencing comprises sequencing by synthesis.

6. The method of claims 1-5, wherein further processing comprises nucleic acid amplification.

7. The method of claim 6, wherein a portion of the stem-loop oligonucleotide is further used to prime an amplification reaction.

8. The method of claim 6-7, wherein the amplification is isothermal amplification.

9. The method of claim 6-7, wherein the amplification is PCR.

10. The method of claims 1-9, wherein the nucleic acid is produced by fragmenting a longer nucleic acid.

11. The method of claims 1-10, wherein the target nucleic acid is blunt ended.

12. The method of claims 1-11, wherein the target nucleic acid comprises a 3' overhang.

13. The method of claims 1-12, wherein the target nucleic acid comprises a 5' overhang.

14. The method of claim 12, wherein the stem-loop oligonucleotide further comprises a 3' overhang.

15. The method of claim 13, wherein the stem-loop oligonucleotide further comprises a 5' overhang.

16. The method of claims 14-15, wherein the stem-loop oligonucleotide overhang and nucleic acid overhang are at least partially complementary to each other.

17. The method of claims 1-16, wherein the attaching step comprises attaching a first RNA/DNA chimeric stem-loop oligonucleotide to a first end of said target nucleic acid and a second RNA/DNA chimeric stem-loop oligonuclotide to a second end of said target nucleic acid.

18. The method of claim 17, wherein said first and said second RNA/DNA chimeric stem-loop oligonucleotides are the same.

19. The method of claims 1-18, wherein the attaching step comprises the use of a ligase.

20. The method of claim 19, wherein the ligase is T4 ligase.

21. The method of claims 1-20, wherein the stem-loop oligonucleotide lacks 5' phosphate.

22. The method of claims 1-21, wherein the target nucleic acid is further extended along the cleaved stem-loop oligonucleotide sequence, thereby forming a target nucleic acid extension product, wherein a portion of the stem-loop oligonucleotide acts as a template.

23. The method of claim 22, wherein the stem-loop oligonucleotide that is in duplex with the nucleic acid extension is further cleaved using an RNase that cleaves RNA from a RNA/DNA heteroduplex, thereby forming a priming ready construct.

24. The method of claim 23, wherein the priming ready construct is further amplified.

25. The method of claim 24, wherein the amplification comprises single primer isothermal amplification (SPIA).

QIAGEN Exhibit 1002
Page 861

26. The method of claims 1-25, wherein at least one loop of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

27. The method of claims 1-26, wherein at least one stem of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

28. The method of claims 1-27, wherein at least one overhang of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

29. A method for amplifying a nucleic acid molecule, comprising:

contacting in a single mixture

(a) an RNA/DNA stem-loop adaptor having a portion of its stem and at least a portion of its loop comprise of RNA;

(b) a double-stranded nucleic acid with a 3' overhang;

(c) RNase that cleaves RNA from a RNA/DNA heteroduplex;

(d) a DNA polymerase with strand displacement activity.

30. The method of claim 29, further comprising forming an extension product by priming the double-stranded nucleic acid at its 3' overhang using at least a portion of the stem-loop oligonucleotide.

31. The method of claim 30, wherein the loop-RNA in the extension product is further cleaved by the RNase, thereby forming a second double stranded nucleic acid with at least one 3' overhang.

32. The method of claims 29-31, wherein the RNase comprises RNase H.

33. The method of claims 29-32, further comprising at least a second amplification mixture with at least a second stem-loop oligonucleotide, wherein each stem-loop oligonucleotide further comprises a barcode sequence.

34. The method of claims 29-33, wherein the 3' overhang is produced by

(a) forming a product, wherein the stem-loop oligonucleotide is attached to a double stranded nucleic acid by one strand;

(b) extending the double stranded nucleic acid along the stem-loop oligonucleotide; and

(c) cleaving the stem-loop oligonucleotide using an RNase that cleaves RNA sequences in an RNA/DNA heteroduplex.

35. The method of claims 29-33, wherein the 3' overhang is produced during single primer isothermal amplification.

36. A kit for amplifying nucleic acids, comprising:

(a) a stem-loop oligonucleotide, wherein at least a portion of the stem comprises RNA and wherein at least a portion of the loop comprises RNA;

(b) an RNase that cleaves RNA sequences in an RNA/DNA duplex;

(c) a ligase; and

(d) a DNA polymerase.

37. A method of amplifying a double stranded target nucleic acid molecule, comprising:

(a) attaching a strand of a first stem-loop oligonucleotide to said double stranded target nucleic acid molecule thereby producing an oligonucleotide-attached nucleic acid molecule; and

(b) cleaving part of the stem-loop oligonucleotide using an RNase; and using a second stem-loop oligonucleotide as a primer in amplifying the double stranded target nucleic acid molecule.

38. The method of claim 37, wherein the second stem-loop oligonucleotide is an exact copy of the first stem-loop oligonucleotide.

39. The method of claims 37-38, wherein the RNase comprises RNase H.

40. The method of claims 37-39, wherein the stem-loop oligonucleotide comprises an RNA/DNA heteroduplex in a stem.

41. The method of claims 37-40, wherein a strand of the oligonucleotide-attached nucleic acid molecule is extended.

42. The method of claim 41, wherein the extension yields a second RNA/DNA heteroduplex.

43. The method of claim 42, wherein cleaving part of the stem-loop oligonucleotide comprises cleaving the second RNA/DNA heteroduplex.

44. The method of claim 43, wherein cleaving part of the stem-loop oligonucleotide generates a priming ready construct.

45. The method of claim 43, wherein the priming ready construct comprises a single stranded region that is 3' to the double stranded target nucleic acid.

46. The method of claims 37-45, wherein the amplification comprises PCR.

47. The method of claims 37-45, wherein the amplification comprises SPIA.

48. The method of claims 37-47, wherein the attaching comprises ligation.

49. The method of claim 48, wherein T4 ligase is utilized in ligation.

50. The method of claims 37-49, wherein the stem-loop oligonucleotide lacks 5' phosphate.

51. The method of claims 37-50, wherein the target nucleic acid is produced by fragmenting a longer nucleic acid.

52. The method of claims 37-51, wherein the target nucleic acid is blunt ended.

53. The method of claims 37-52, wherein the target nucleic acid comprises a 3' overhang.

54. The method of claims 37-53, wherein the target nucleic acid comprises a 5' overhang.

55. The method of claim 53, wherein the stem-loop oligonucleotide further comprises a 3' overhang.

56. The method of claim 54, wherein the stem-loop oligonucleotide further comprises a 5' overhang.

57. The method of claims 55-56, wherein the stem-loop oligonucleotide overhang and nucleic acid overhang are at least partially complementary to each other.

58. The method of claims 37-57, wherein at least one loop of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

59. The method of claims 37-58, wherein at least one stem of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

60. The method of claims 37-59, wherein at least one overhang of the stem-loop oligonucleotide further comprises one or more sequence elements selected from the group consisting of a primer binding site, a sequencing primer binding site, a barcode, a promoter, an adaptor-group sequence, a restriction enzyme recognition site, a target complementary overhang, a probe complementary overhang, a probe binding site, a random sequence, and a complementary sequence to any of the sequence elements within the group.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
14/836,936	08/26/2015	1634	1980	25115-755.202	23	1

CONFIRMATION NO. 2434

UPDATED FILING RECEIPT

12754
Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

OC000000080948296

Date Mailed: 02/26/2016

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Doug Amorese, Los Altos, CA;
Nurith Kurn, Palo Alto, CA;
Jonathan Scolnick, San Francisco, CA;
Alexandra Hui Wang, Mountain View, CA;

Applicant(s)

NuGEN Technologies, Inc., San Carlos, CA;

Power of Attorney: The patent practitioners associated with Customer Number 12754

Domestic Priority data as claimed by applicant

This appln claims benefit of 62/042,240 08/26/2014
and claims benefit of 62/188,337 07/02/2015

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.) - None.
Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

Permission to Access Application via Priority Document Exchange: Yes

Permission to Access Search Results: No

Applicant may provide or rescind an authorization for access using Form PTO/SB/39 or Form PTO/SB/69 as appropriate.

If Required, Foreign Filing License Granted: 09/10/2015

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 14/836,936**

Projected Publication Date: 06/02/2016

Non-Publication Request: No

Early Publication Request: No

Title

COMPOSITIONS AND METHODS FOR TARGETED NUCLEIC ACID SEQUENCE ENRICHMENT AND HIGH EFFICIENCY LIBRARY GENERATION

Preliminary Class

435

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications: No

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit http://www.SelectUSA.gov or call +1-202-482-6800.

Docket No. 25115-755.202

NON-PROVISIONAL PATENT APPLICATION

COMPOSITIONS AND METHODS FOR TARGETED NUCLEIC ACID SEQUENCE ENRICHMENT AND HIGH EFFICIENCY LIBRARY GENERATION

Inventors: Doug AMORESE,
Citizen of United States of America, Residing at
357 S El Monte Ave
Los Altos, CA 94022

Nurith KURN,
Citizen of Israel, Residing at
2876 Ramona Street
Palo Alto, CA 94306

Jonathan SCOLNICK,
Citizen of United States of America, Residing at
451 Kansas Street
San Francisco, CA, 94107

Alexandra Hui WANG,
Citizen of United States of America, Residing at
1390 West Dana Street
Mountain View, CA 94041

Assignee: NuGEN Technologies, Inc.
201 Industrial Road, Suite 310
San Carlos, CA 94070

Entity: Large

WSGR

Wilson Sonsini Goodrich & Rosati
PROFESSIONAL CORPORATION

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

Filed Electronically on: August 26, 2015

COMPOSITIONS AND METHODS FOR TARGETED NUCLEIC ACID SEQUENCE ENRICHMENT AND HIGH EFFICIENCY LIBRARY GENERATION

CROSS-REFERENCE

[0001] This application claims priority to U.S. Provisional Application No. 62/042,240, filed August 26, 2014 and U.S. Provisional Application No. 62/188,337, filed July 2, 2015, which applications are incorporated herein by reference. This application is related to U.S. Provisional Application No. 61/591,241, filed Jan. 26, 2012; PCT Application No. PCT/US13/23278, filed January 25, 2013, and U.S. Patent Application No. 13/750,768, which applications are incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0002] With the rapid development of next generation sequencing (NGS) technologies and platforms, whole genome sequencing is becoming increasingly feasible. Researchers are driven to generate increasing amounts of data to achieve greater understanding of variance and biological trends, and to generate data from smaller sample sizes to avoid averaging across multiple cells within a tissue.

[0003] Although the cost of whole genome sequencing is decreasing and the throughput of the NGS platforms is increasing, it is nonetheless often more practical and cost-effective to select genomic regions of interest for sequencing and analysis. Target enrichment is a commonly employed strategy in genomic DNA sequencing in which genomic regions of interest are selectively captured from a DNA sample before sequencing. Focused target enrichment is an important tool especially in the fields of study where sequencing of a large number of samples is necessary (e.g. population-based studies of disease markers or SNPs), making whole genome sequencing cost-prohibitive. Similarly, improvements have been made that enable DNA libraries to be made from nucleic acid from fewer number of cells, but these are bound by the limitations of the efficiency of ligation reactions.

[0004] Several approaches to target enrichment have been developed which vary from one another in terms of sensitivity, specificity, reproducibility, uniformity, cost and ease of use. The target enrichment methods commonly employed today can be divided into three major categories, each with its distinct advantages and disadvantages: 1) PCR-based methods; 2) capture-by-hybridization, e.g., on-array or in-solution hybrid capture; and 3) capture-by-circularization, e.g., molecular inversion probe-based methods.

[0005] The PCR-based methods employ highly parallel PCR amplification, where each target sequence in the sample has a corresponding pair of unique, sequence-specific primers. The simultaneous use of numerous primer pairs makes multiplex PCR impractical due to high level of non-specific amplification and primer-primer interactions. Recently developed microdroplet PCR technology (Tewhey *et al.*, 2009) in which each amplification reaction is physically separated into an individual droplet removes the constraints of multiplex PCR relating to non-specific amplification and primer-primer interactions. However, microdroplet PCR and other improved PCR-based methods require special instrumentation or platforms, are limited in their throughput, and,

as with conventional multiplex PCR, require a large number of individual primer pairs when enriching for a multitude of regions on interest, thus making target enrichment costly.

[0006] Hybrid capture methods are based on the selective hybridization of the target genomic regions to user-designed oligonucleotides. The hybridization can be to oligonucleotides immobilized on high or low density microarrays (on-array capture), or solution-phase hybridization to oligonucleotides modified with a ligand (e.g. biotin) which can subsequently be immobilized to a solid surface, such as a bead (in-solution capture). The hybrid capture methods require complex pools of costly long oligonucleotides and long periods (typically 48 hours) of hybridization for efficient capture. For on-array hybrid capture, expensive instrumentation and hardware is also required. Because of the relatively low efficiency of the hybridization reaction, large quantities of input DNA are needed.

[0007] The molecular inversion probe (MIP) based method relies on construction of numerous single-stranded linear oligonucleotide probes, consisting of a common linker flanked by target-specific sequences. Upon annealing to a target sequence, the probe gap region is filled via polymerization and ligation, resulting in a circularized probe. The circularized probes are then released and amplified using primers directed at the common linker region. One of the main disadvantages of the MIP-based target enrichment is its relatively low capture uniformity, meaning there is large variability in sequence coverage across the target regions. As with PCR and hybrid capture, the MIP-based method requires a large number of target-specific oligonucleotides, which can be costly.

[0008] There is a need for improved methods for selective target enrichment that allow for low-cost, high throughput capture of genomic regions of interest without specialized instrumentation. Additionally, there is also a need for high efficiency nucleic acid library generation. The methods of the invention described herein fulfills these needs.

SUMMARY OF THE INVENTION

[0009] In one aspect, a method is provided for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising: a) annealing an oligonucleotide to the nucleic acid sequence of interest in a nucleic acid fragment, wherein the oligonucleotide comprises a 3' portion complementary to the nucleic acid sequence of interest and a 5' portion comprising a first adaptor sequence; b) extending the oligonucleotide annealed to the nucleic acid sequence of interest in the nucleic acid fragment with a polymerase, thereby generating an oligonucleotide extension product comprising sequence complementary to the nucleic acid sequence of interest and the first adaptor sequence at a first end; c) ligating a sequence complementary to a second adaptor sequence to the oligonucleotide extension product, thereby generating an oligonucleotide extension product comprising the first adaptor sequence at the first end, the sequence complementary to the nucleic acid sequence of interest, and a sequence complementary to the second adaptor sequence at a second end; and d) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest, thereby generating an enriched nucleic acid sequence of interest. In

some cases, the method further comprises, before step a), fragmenting a nucleic acid in the sample, thereby generating the nucleic acid fragment of step a) comprising the sequence of interest. In some cases, the nucleic acid fragment is double-stranded prior to step a). In some cases, the nucleic acid fragment comprises deoxyribonucleic acid (DNA). In some cases, the DNA is genomic DNA. In some cases, the DNA is cDNA. In some cases, the method further comprises denaturing the double-stranded nucleic acid fragment prior to step b), wherein the nucleic acid fragment comprising the nucleic acid sequence of interest in step b) is single-stranded prior to annealing to the oligonucleotide. In some cases, the 3' portion of the oligonucleotide comprises a random sequence. In some cases, the 3' portion of the oligonucleotide comprises a sequence designed to anneal to the nucleic acid sequence of interest. In some cases, ligating the sequence complementary to the second adaptor sequence comprises ligating the sequence complementary to the second adaptor sequence to a 3' end of the oligonucleotide extension product. In some cases, the second adaptor sequence is within a strand of a double-stranded adaptor. In some cases, the first adaptor sequence and the second adaptor sequence are distinct. In some cases, the first adaptor sequence and/or the second adaptor sequence comprise a barcode sequence. In some cases, the first adaptor sequence and/or the second adaptor sequence comprise sequence that can anneal to a sequencing primer, or complement of a sequence that can anneal to a sequencing primer. In some cases, the polymerase is a DNA polymerase. In some cases, the method further comprises annealing a plurality of oligonucleotides to different nucleic acid sequences of interest in nucleic acid fragments, wherein each of the plurality of oligonucleotides comprises a 3' portion complementary to a nucleic acid sequence of interest among the nucleic acid sequences of interest and a 5' portion comprising a first adaptor sequence; extending the plurality of oligonucleotides annealed to the different nucleic acid sequences of interest in the nucleic acid fragments with a polymerase, thereby generating a plurality of oligonucleotide extension products comprising sequence complementary to the nucleic acid sequences of interest and the first adaptor sequence at a first end; ligating a sequence complementary to a second adaptor sequence to the plurality of oligonucleotide extension products, thereby generating a plurality of oligonucleotide extension products comprising the first adaptor sequence at the first end, the sequence complementary to the nucleic acid sequence of interest, and a sequence complementary to the second adaptor sequence at a second end; and amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequences of interest, thereby generating enriched nucleic acid sequences of interest.

[0010] In some cases, the amplifying of step d) comprises a polymerase chain reaction. In some cases, the method further comprises enriching a plurality of nucleic acid sequences of interest using the steps of the method. In some cases, the method further comprises sequencing the plurality of enriched nucleic acid sequences of interest. In some cases, the nucleic acid sequence of interest comprises a gene. In some cases, the gene is a cancer gene. In some cases, the method further comprises sequencing the nucleic acid sequence of interest. In some cases, the sequencing comprises use of a reversible-dye terminator.

[0011] In another aspect, a method is provided for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising: a) combining a transposome comprising a transposase and a transposon sequence with a nucleic acid in the sample, thereby generating a nucleic acid fragment comprising the nucleic acid sequence of interest and a transposon sequence appended to a first end of the nucleic acid fragment; b) annealing an oligonucleotide to the nucleic acid fragment comprising the transposon sequence at the first end, wherein the oligonucleotide comprises a 3' portion that is complementary to the transposon sequence and a 5' portion comprising a first adaptor sequence; c) extending the first oligonucleotide annealed to the transposon sequence with a polymerase, thereby generating a first oligonucleotide extension product comprising sequence complementary to the nucleic acid fragment and a first adaptor sequence at the first end; d) annealing a second oligonucleotide to a sequence complementary to the nucleic acid sequence of interest in the nucleic acid fragment comprising the first adaptor sequence at the first end, wherein the second oligonucleotide comprises a 3' portion that anneals to the complement of the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence; e) extending the second oligonucleotide annealed to the complement of the nucleic acid sequence of interest in the nucleic acid fragment comprising the first adaptor sequence at the first end with a polymerase, thereby generating a second oligonucleotide extension product comprising sequence complementary to the first adaptor sequence at the first end, the nucleic acid sequence of interest, and the second adaptor sequence at the second end; and f) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest. In some cases, the transposase is a Tn5 transposase. In some cases, the nucleic acid fragment of step a) is double stranded. In some cases, the method further comprises denaturing the double stranded nucleic acid fragment prior to step b), wherein the nucleic acid fragment in step b) is single-stranded prior to annealing to the oligonucleotide. In some cases, the first oligonucleotide extension product is annealed to the nucleic acid fragment after the extension step c). In some cases, the method further comprises denaturing the first oligonucleotide extension product from the nucleic acid fragment prior to step d), whereby the first oligonucleotide extension product in step d) is single stranded. In some cases, the 3' portion of the second oligonucleotide comprises a random sequence. In some cases, the 3' portion of the second oligonucleotide comprises a sequence designed to anneal to the nucleic acid sequence of interest. In some cases, the first adaptor sequence and the second adaptor sequence are distinct. In some cases, the first adaptor sequence and/or the second adaptor sequence comprise a barcode sequence. In some cases, the first adaptor sequence and/or the second adaptor sequence comprise sequence designed to anneal to a sequencing primer. In some cases, the polymerase is a DNA polymerase. In some cases, the amplification of step f) comprises polymerase chain reaction. In some cases, the annealing step d) comprises annealing a plurality of second oligonucleotides to a plurality of nucleic acid fragments, wherein the plurality of second oligonucleotides comprise a 3' portion that anneals to the complement of a different nucleic acid sequence of interest in the nucleic acid fragments and a 5' portion comprising the second adaptor sequence. In some cases, the method further comprises using steps a) –c) to produce a plurality of first oligonucleotide extension products comprising

sequence complementary to nucleic acid fragments and a first adaptor sequence at the first end; annealing each of a plurality of second oligonucleotides to a sequence complementary to a nucleic acid sequence of interest in the nucleic acid fragments comprising the first adaptor sequence at the first end, wherein the plurality of second oligonucleotides comprise a 3' portion that anneals to a complement of a nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence; extending the plurality of second oligonucleotides annealed to a complement of a nucleic acid sequence of interest in a nucleic acid fragment comprising the first adaptor sequence at the first end with a polymerase, thereby generating a plurality of second oligonucleotide extension products comprising sequence complementary to the first adaptor sequence at the first end, a nucleic acid sequence of interest, and the second adaptor sequence at the second end; and amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest.

[0012] In some cases, the method further comprises enriching a plurality of nucleic acid sequences of interest using the steps of the method. In some cases, the method further comprises sequencing a plurality of enriched nucleic acid sequences of interest. In some cases, the sequencing comprises use of a reversible-dye terminator.

[0013] In another aspect, a method is provided for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising: a) combining a transposome comprising a transposase and a transposon sequence with a nucleic acid in the sample, thereby generating a nucleic acid fragment comprising the nucleic acid sequence of interest and a transposon sequence appended to a first end of the nucleic acid fragment; b) annealing an oligonucleotide to the nucleic acid fragment comprising the transposon sequence at the first end, wherein the oligonucleotide comprises a 3' portion that is complementary to the nucleic acid sequence of interest in the nucleic acid fragment and a 5' portion comprising a first adaptor sequence; c) extending the oligonucleotide annealed to the nucleic acid sequence of interest with a polymerase, thereby generating an oligonucleotide extension product comprising sequence complementary to the transposon sequence at a first end, sequence complementary to the nucleic acid sequence of interest, and a first adaptor sequence at a second end; and d) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the transposon sequence and a second primer that anneals to the complement of the first adaptor sequence to enrich for the nucleic acid sequence of interest. In some cases, the transposase is a Tn5 transposase. In some cases, the nucleic acid fragment of step a) is double stranded. In some cases, the method further comprises denaturing the double stranded nucleic acid fragment prior to step b), wherein the nucleic acid fragment in step b) is single-stranded prior to annealing to the oligonucleotide. In some cases, the oligonucleotide extension product is annealed to the nucleic acid fragment after the extension of step c). In some cases, the method further comprises denaturing the oligonucleotide extension product from the nucleic acid fragment prior to step d), whereby the oligonucleotide extension product in step d) is single stranded. In some cases, the first adaptor sequence comprises a barcode sequence. In some cases, the first adaptor sequence comprises sequence designed to anneal to a sequencing primer. In some cases, the polymerase is a DNA polymerase. In some cases, the method further

comprises generating a plurality of nucleic acid fragments comprising a nucleic acid sequence of interest and a transposon sequence appended to a first end of the nucleic acid fragment; annealing a plurality of oligonucleotides to the nucleic acid fragments comprising the transposon sequence at the first end, wherein each of the plurality of oligonucleotides comprises a 3' portion that is complementary to the nucleic acid sequence of interest in a nucleic acid fragment and a 5' portion comprising a first adaptor sequence; extending the plurality of oligonucleotides annealed to the nucleic acid sequences of interest with a polymerase, thereby generating a plurality of oligonucleotide extension products comprising sequence complementary to the transposon sequence at a first end, sequence complementary to the nucleic acid sequence of interest, and a first adaptor sequence at a second end; and amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the transposon sequence and a second primer that anneals to the complement of the first adaptor sequence to enrich for the nucleic acid sequences of interest.

[0014] In some cases, the amplification of step f) comprises polymerase chain reaction. In some cases, the method further comprises enriching a plurality of nucleic acid sequences of interest using the steps of the method. In some cases, the method further comprises sequencing a plurality of enriched nucleic acid sequences of interest. In some cases, the sequencing comprises use of a reversible-dye terminator.

[0015] In some aspects, a method is provided for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising: a) annealing a first oligonucleotide to a nucleic acid fragment comprising the nucleic acid sequence of interest, wherein the first oligonucleotide comprises a 3' portion complementary to a sequence in the nucleic acid fragment and a 5' portion comprising a first adaptor sequence; b) extending the first oligonucleotide annealed to the nucleic acid fragment comprising the nucleic acid sequence of interest with a polymerase, thereby generating a first oligonucleotide extension product comprising sequence complementary to the nucleic acid sequence of interest and the first adaptor sequence at a first end; c) annealing a second oligonucleotide to the complement of the nucleic acid sequence of interest in the first oligonucleotide extension product, wherein the second oligonucleotide comprises a 3' portion complementary to the sequence complementary to the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence; d) extending the second oligonucleotide annealed to the complement of the nucleic acid sequence of interest with a polymerase, thereby generating a second oligonucleotide extension product comprising a sequence complementary to the first adaptor sequence at a first end, the nucleic acid sequence of interest, and the second adaptor sequence at a second end; and e) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest, thereby generating an enriched nucleic acid sequence of interest. In some cases, the method further comprises, before step a), fragmenting a nucleic acid in the sample, thereby generating the nucleic acid fragment of step a) comprising the nucleic acid sequence of interest. In some cases, the nucleic acid fragment is double-stranded prior to step a). In some cases, the nucleic acid fragment comprises deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). In some cases, the DNA is genomic DNA. In some cases, the DNA is cDNA. In some cases, the 3' portion of the first

oligonucleotide comprises DNA and the 5' portion of the first oligonucleotide comprises RNA. In some cases, the method further comprises denaturing the double-stranded nucleic acid fragment prior to step a), wherein the nucleic acid fragment comprising the nucleic acid sequence of interest in step a) is single-stranded prior to annealing to the first oligonucleotide. In some cases, the 3' portion of the first oligonucleotide comprises a random sequence. In some cases, the 3' portion of the second oligonucleotide comprises a sequence designed to anneal to the nucleic acid sequence of interest. In some cases, the first adaptor sequence and the second adaptor sequence are distinct. In some cases, the first adaptor sequence and/or the second adaptor sequence comprise barcode sequence. In some cases, the first adaptor sequence and/or the second adaptor sequence comprise sequence that can anneal to a sequencing primer, or complement of a sequence that can anneal to a sequencing primer. In some cases, the polymerase is a DNA polymerase.

[0016] In some cases, the method further comprises: annealing a plurality of first oligonucleotides to different nucleic acid fragments comprising different nucleic acid sequences of interest, wherein each of the plurality of first oligonucleotides comprises a 3' portion complementary to a sequence in the nucleic acid fragment among the nucleic acid fragments and a 5' portion comprising a first adaptor sequence; extending the plurality of first oligonucleotides annealed to the different nucleic acid fragments comprising different nucleic acid sequences of interest with a polymerase, thereby generating a plurality of first oligonucleotide extension products each comprising sequence complementary to the nucleic acid sequences of interest and the first adaptor sequence at a first end; annealing a plurality of second oligonucleotides to the complement of the nucleic acid sequences of interest in the plurality of first oligonucleotide extension products, wherein each of the plurality of second oligonucleotides comprises a 3' portion complementary to the sequence complementary to the nucleic acid sequence of interest among the nucleic acid sequences of interest and a 5' portion comprising a second adaptor sequence; extending the plurality of second oligonucleotides annealed to the complement of the nucleic acid sequences of interest in the first oligonucleotide extension products, thereby generating a plurality of second oligonucleotide extension products each comprising a sequence complementary to the first adaptor sequence at a first end, the nucleic acid sequence of interest among the nucleic acid sequences of interest, and the second adaptor sequence at a second end; and amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequences of interest, thereby generating enriched nucleic acid sequences of interest. In some cases, the method further comprises prior to step e), digesting the second oligonucleotide extension product with RNase H. In some cases, amplifying in step e) comprises a polymerase chain reaction. In some cases, the method further comprises enriching a plurality of nucleic acid sequences of interest using steps a)-e). In some cases, the method further comprises sequencing the plurality of enriched nucleic acid sequences of interest. In some cases, the nucleic acid sequence of interest comprises a gene. In some cases, the gene is a cancer gene. In some cases, the method further comprises sequencing the enriched nucleic acid sequence of interest. In some cases, the sequencing comprises use of a reversible-dye terminator.

[0017] In another aspect, a method is provided for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising: a) fragmenting the nucleic acids thereby generating nucleic acid fragments, wherein the nucleic acid fragments comprise the nucleic acid sequence of interest; b) ligating a first adaptor sequence to a 5' end of each of the nucleic acid fragments thereby generating nucleic acid fragments comprising the first adaptor sequence at the 5' end; c) annealing one or more oligonucleotides to the nucleic acid fragments comprising the first adaptor sequence at the 5' end, wherein each of the one or more oligonucleotides comprise a 3' portion that is complementary to a nucleic acid sequence of interest present in one or more of the nucleic acid fragments, and a 5' portion comprising a second adapter sequence; d) extending the one or more oligonucleotides with a polymerase thereby generating one or more oligonucleotide extension products comprising a sequence complementary to the first adaptor sequence at a first end, a sequence complementary to the nucleic acid sequence of interest and the second adaptor sequence at a second end; and e) amplifying the one or more oligonucleotide extension products using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest, wherein the nucleic acid sequence of interest comprises a gene fusion or a gene rearrangement. In some cases, the nucleic acids in the sample comprise RNA. In some cases, the RNA is converted to double-stranded cDNA prior to step b). In some cases, the gene fusion or gene rearrangement is in a cancer gene. In some cases, the gene rearrangement is a translocation or an inversion. In some cases, the one or more oligonucleotides anneals to a portion of at least one gene in the gene fusion. In some cases, the at least one gene in the gene fusion is BCR, ABL, ALK, RET, or ROS1. In some cases, the method further comprises sequencing the nucleic acid sequence of interest after step e). In some cases, the method further comprises denaturing the nucleic acid fragments prior to step c), thereby generating single-stranded nucleic acid fragments comprising the first adaptor sequence at the 5' end. In some cases, the first adaptor sequence, the second adaptor sequence or a combination thereof comprises barcode sequence.

INCORPORATION BY REFERENCE

[0018] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0020] **FIG. 1** depicts selective target enrichment using ligation of a single forward adaptor at the ends of the DNA fragments in the DNA library. The sequence-specific oligonucleotide that anneals to the target region of interest contains a common reverse adaptor sequence at its 5'end, and following sequence-specific oligonucleotide extension, PCR is performed using a set of primers specific to the forward and reverse adaptors.

[0021] **FIG. 2** depicts an alternative ligation protocol wherein the DNA-fragment-adaptor complex is denatured following ligation without nick repair and adaptor fill-in, generating ligation products where non-complementary ends exist on each insert.

[0022] **FIG. 3** depicts selective target enrichment using ligation of partial duplex adaptors. Cleavage of the 5' end of the long strand of the partial duplex adaptor (and the corresponding complementary sequence of the extended sequence-specific oligonucleotide) by a nucleic acid modifying enzyme specific for double-stranded DNA allows for ligation of a new adaptor pair , and consequently, amplification with primers corresponding to the new adaptors.

[0023] **FIG. 4** depicts high efficiency NGS library generation using random priming. The oligonucleotide that anneals to the DNA fragment contains a common reverse adaptor sequence at its 5' end, and following primer extension, PCR is performed using a set of primers specific to the forward and reverse adaptors.

[0024] **FIG. 5** depicts the Ovation Target Enrichment System workflow.

[0025] **FIG. 6** depicts Agencourt RNAClean XP Bead purification process overview.

[0026] **FIG. 7** depicts fragment distribution when 1μL of a target enriched library is loaded into a Bioanalyzer High Sensitivity DNA Chip.

[0027] **FIG. 8** depicts stylized qPCR amplification plot.

[0028] **FIG. 9** depicts enrichment of cDNA from 100ng universal human reference (UHR) RNA using the Ovation Target Enrichment System.

[0029] **FIG. 10** depicts selective target enrichment using random priming of a forward adaptor and target-specific priming of a reverse adaptor.

[0030] **FIG. 11** depicts selective target enrichment using target-specific priming to append a reverse adaptor to nucleic acid fragments followed by ligation to append a forward adaptor.

[0031] **FIG. 12** depicts selective target enrichment using a transposase-mediated fragmenting and tagging reaction to append a forward adaptor to the ends of nucleic acid fragments followed by target-specific priming to append a reverse adaptor.

[0032] **FIG. 13** depicts selective target enrichment using a transposase-mediated fragmenting and tagging reaction to append a first adaptor to the end of nucleic acid fragments followed by target-specific priming to append a second adaptor.

[0033] **FIG. 14** depicts selective target enrichment using random priming to append a first adaptor nucleic acid fragments followed by target-specific priming to append a second adaptor.

DETAILED DESCRIPTION OF THE INVENTION

General

[0034] The methods provided herein can be used for the selective enrichment of a plurality of defined target sequences from complex DNA with a set of common primers and adaptors, thus circumventing the need for multiplex PCR and multiple primer pairs. A multiplicity of target regions of interest are envisioned: for example, the regions of interest can represent all known coding regions, the entire exome, selected regions of coding genomic regions representing selected pathways, selected genomic regions known to comprise genomic variation related to altered phenotype, entire or selected regions of a specific chromosome, and the like.The methods described herein can be used for high efficiency nucleic acid library production as well. A target region of interest can include sequence used for primer annealing (e.g., a probe "landing zone").

[0035] Altogether, the methods described herein can create a simple, low cost, high throughput system for target enrichment and library preparation.

[0036] Reference will now be made in detail to exemplary embodiments described herein. While the disclosed methods and compositions will be described in conjunction with the exemplary embodiments, it will be understood that these exemplary embodiments are not intended to limit the invention. Methods and compositions described herein encompass alternatives, modifications and equivalents.

[0037] Provided herein are methods and compositions for the enrichment of specific target sequences of interest from a sample comprising nucleic acids. The methods described herein can enrich target sequences using conventional duplex adaptors and/or partial duplex adaptors, sequence specific oligonucleotides, restriction enzymes and ligation. The methods can further enable enrichment of target sequences from specific strands of template nucleic acids which can be further amplified using a variety of amplification methods. Provided herein are methods for high efficiency generation of libraries comprising specific nucleic acid sequences of interest.

[0038] Provided herein are methods and compositions for the enrichment of target nucleic acid sequences from a sample comprising nucleic acids. The method can comprise fragmenting nucleic acids in an input sample to generate nucleic acid fragments. The nucleic acids can be DNA, or RNA. The nucleic acids can be single or double stranded. The DNA can be genomic DNA or cDNA or any combination thereof. The nucleic acids in an input sample can be double stranded DNA. Fragmentation of the nucleic acids can be achieved through methods known in the art. Fragmentation can be through physical fragmentation methods and/or enzymatic fragmentation methods. Physical fragmentation methods can include nebulization, sonication, and/or hydrodynamic shearing. The fragmentation can be accomplished mechanically comprising subjecting the nucleic acids in the input sample to acoustic sonication. The fragmentation can comprise treating the nucleic acids in the input sample with one or more enzymes under conditions suitable for the one or more enzymes to generate double-stranded nucleic acid breaks. Examples of enzymes useful in the generation of nucleic acid or polynucleotide fragments include sequence specific and non-sequence specific nucleases. Non-limiting examples of nucleases include DNase I, Fragmentase, restriction endonucleases, variants thereof, and combinations thereof. Reagents for carrying out enzymatic fragmentation reactions can be commercially available (e.g, from New England Biolabs). For

example, digestion with DNase I can induce random double-stranded breaks in DNA in the absence of Mg^{++} and in the presence of Mn^{++}. Fragmentation can comprise treating the nucleic acids in the input sample with one or more restriction endonucleases. Fragmentation can produce fragments having 5' overhangs, 3' overhangs, blunt ends, or a combination thereof. In some embodiments, such as when fragmentation comprises the use of one or more restriction endonucleases, cleavage of sample polynucleotides leaves overhangs having a predictable sequence. The method can include the step of size selecting the fragments via standard methods known in the art such as column purification or isolation from an agarose gel.

[0039] The nucleic acids in the input sample can be fragmented into a population of fragmented nucleic acid molecules or polynucleotides of one or more specific size range(s). The fragments can have an average length from about 10 to about 10,000 nucleotides. The fragments can have an average length from about 50 to about 2,000 nucleotides. The fragments can have an average length from about 100 to about 2,500, about 10 to about 1,000, about 10 to about 800, about 10 to about 500, about 50 to about 500, about 50 to about 250, or about 50 to about 150 nucleotides. The fragments can have an average length less than 10,000 nucleotides, such as less than 5,000 nucleotides, less than 2,500 nucleotides, less than 2,500 nucleotides, less than 1,000 nucleotides, less than 500 nucleotides, such as less than 400 nucleotides, less than 300 nucleotides, less than 200 nucleotides, or less than 150 nucleotides.

[0040] Fragmentation of the nucleic acids can be followed by end repair of the nucleic acid fragments. End repair can include the generation of blunt ends, non-blunt ends (i.e sticky or cohesive ends), or single base overhangs such as the addition of a single dA nucleotide to the 3′-end of the nucleic acid fragments, by a polymerase lacking 3′-exonuclease activity. End repair can be performed using any number of enzymes and/or methods known in the art including, but not limited to, commercially available kits such as the Encore ™ Ultra Low Input NGS Library System I. End repair can be performed on double stranded DNA fragments to produce blunt ends wherein the double stranded DNA fragments contain 5' phosphates and 3' hydroxyls.The double-stranded DNA fragments can be blunt-end polished (or "end repaired") to produce DNA fragments having blunt ends, prior to being joined to adapters. Generation of the blunt ends on the double stranded fragments can be generated by the use of a single strand specific DNA exonuclease such as for example exonuclease 1, exonuclease 7 or a combination thereof to degrade overhanging single stranded ends of the double stranded products. The double stranded DNA fragments can be blunt ended by the use of a single stranded specific DNA endonuclease, for example, but not limited to, mung bean endonuclease or S1 endonuclease. The double stranded products can be blunt ended by the use of a polymerase that comprises single stranded exonuclease activity such as for example T4 DNA polymerase, or any other polymerase comprising single stranded exonuclease activity or a combination thereof to degrade the overhanging single stranded ends of the double stranded products. The polymerase comprising single stranded exonuclease activity can be incubated in a reaction mixture that does or does not comprise one or more dNTPs. A combination of single stranded nucleic acid specific exonucleases and one or more polymerases can be used to blunt end the double stranded fragments generated by fragmenting the sample comprising nucleic acids. The nucleic acid fragments can be made blunt ended by filling in the overhanging

single stranded ends of the double stranded fragments. For example, the fragments can be incubated with a polymerase such as T4 DNA polymerase or Klenow polymerase or a combination thereof in the presence of one or more dNTPs to fill in the single stranded portions of the double stranded fragments. The double stranded DNA fragments can be made blunt by a combination of a single stranded overhang degradation reaction using exonucleases and/or polymerases, and a fill-in reaction using one or more polymerases in the presence of one or more dNTPs.

[0041] In some embodiments, the 5' and/or 3' end nucleotide sequences of fragmented nucleic acids are not modified or end-repaired prior to appending with the adapter oligonucleotides of the present invention. For example, fragmentation by a restriction endonuclease can be used to leave a predictable overhang, followed by ligation with one or more adapter oligonucleotides comprising an overhang complementary to the predictable overhang on a nucleic acid fragment. In another example, cleavage by an enzyme that leaves a predictable blunt end can be followed by ligation of blunt-ended nucleic acid fragments to adapter oligonucleotides comprising a blunt end. In some embodiments, end repair can be followed by an addition of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more nucleotides, such as one or more adenine, one or more thymine, one or more guanine, or one or more cytosine, to produce an overhang. Nucleic acid fragments having an overhang can be joined to one or more adapter oligonucleotides having a complementary overhang, such as in a ligation reaction. For example, a single adenine can be added to the 3' ends of end repaired DNA fragments using a template independent polymerase, followed by ligation to one or more adapters each having a thymine at a 3' end. In some embodiments, adapter oligonucleotides can be joined to blunt end double-stranded nucleic acid fragments which have been modified by extension of the 3' end with one or more nucleotides followed by 5' phosphorylation. In some cases, extension of the 3' end can be performed with a polymerase such as for example Klenow polymerase or any of the suitable polymerases provided herein, or by use of a terminal deoxynucleotide transferase, in the presence of one or more dNTPs in a suitable buffer containing magnesium. In some embodiments, nucleic acid fragments having blunt ends can be joined to one or more adapters comprising a blunt end. Phosphorylation of 5' ends of nucleic acid fragments can be performed for example with T4 polynucleotide kinase in a suitable buffer containing ATP and magnesium. The fragmented nucleic acid molecules may optionally be treated to dephosphorylate 5' ends or 3' ends, for example, by using enzymes known in the art, such as phosphatases.

[0042] The methods described herein for enriching for target nucleic acid sequences further comprise appending a first adaptor to the nucleic acid fragments generated by the methods described herein. Appending the first adaptor to the nucleic acid fragments generated by methods described herein can be achieved using a ligation reaction or a priming reaction. In one embodiment, appendage of a first adaptor to the nucleic acid fragments comprises ligation. In one embodiment, ligation of the first adaptor to the nucleic acid fragments can be following end repair of the nucleic acid fragments. In another embodiment, the ligation of the first adaptor to the nucleic acid fragments can be following generation of the nucleic acid fragments without end repair of the nucleic acid fragments. The first adaptor can be any type of adaptor known in the art including, but not limited to, conventional duplex or double stranded adaptors in which the adaptor comprises two complementary strands. In

one embodiment, the first adaptor can be a forward adaptor. In a preferred embodiment, the first adaptor can be a double stranded DNA adaptor. In one embodiment, the first adaptor can be an oligonucleotide of known sequence and, thus, allow generation and/or use of sequence specific primers for amplification and/or sequencing of any polynucleotides to which the first adaptor(s) is appended or attached. In one embodiment, the first adaptor can be a conventional duplex adaptor, wherein the first adaptor comprises sequence well known in the art. In a preferred embodiment, the first adaptor can be appended to the nucleic acid fragments generated by the methods described herein in multiple orientations. In a preferred embodiment, the methods described herein can involve the use of a first duplex adaptor comprising double stranded DNA of known sequence that is blunt ended and can bind to the double stranded nucleic acid fragments generated by the methods described herein in one of two orientations. In one embodiment, the first adaptor can be ligated to each of the nucleic acid fragments such that each of the nucleic acid fragments comprises the same first adaptor. In other words, each of the nucleic acid fragments comprises a common first adaptor. In another embodiment, a first adaptor can be appended or ligated to a library of nucleic acid fragments generated by the methods described herein such that each nucleic acid fragment in the library of nucleic acid fragments comprises the first adaptor ligated to one or both ends.

[0043] In one embodiment, the first adaptor can be ligated or appended to the 5' and/or 3' ends of the nucleic acid fragments generated by the methods described herein. The first adaptor can comprise two strands wherein each strand comprises a free 3' hydroxyl group but neither strand comprises a free 5' phosphate. In one embodiment, the free 3' hydroxyl group on each strand of the first adaptor can be ligated to a free 5' phosphate present on either end of the nucleic acid fragments of the present invention. In this embodiment, the first adaptor comprises a ligation strand and a non-ligation strand whereby the ligation strand can be ligated to the 5'phosphate on either end of the nucleic acid fragment while a nick or gap can be present between the non-ligation strand of the first adaptor and the 3' hydroxyl on either end of the nucleic acid fragment. In one embodiment, the nick or gap can be filled in by performing a gap repair reaction. In one embodiment, the gap repair can be performed with a DNA dependent DNA polymerase with strand displacement activity. In one embodiment, the gap repair can be performed using a DNA dependent DNA polymerase with weak or no strand displacement activity. In one embodiment, the ligation strand of the first adaptor can serve as the template for the gap repair or fill-in reaction. In this embodiment, the gap repair or fill-in reaction comprises an extension reaction wherein the ligation strand of the first adaptor serves as a template and leads to the generation of nucleic acid fragments with complementary termini or ends as depicted, for example, in FIG. 1. In one embodiment, the gap repair can be performed using *Taq* DNA polymerase. In one embodiment, the ligation of the first adaptor to the nucleic acid fragments generated by the methods described herein may not be followed by gap repair as depicted, for example, in FIG. 2. In this embodiment, the nucleic acid fragments comprise first adaptor sequence ligated only at the 5' end of each strand.

[0044] Appending a first adaptor to a nucleic acid fragment can comprise a priming reaction, for example, in FIG. 11. An oligonucleotide comprising, at a first end, sequence complementary to a target sequence of interest present in the nucleic acid fragment and, at a second end, sequence comprising the first adaptor sequence, wherein

the first adaptor sequence is not complementary to the target nucleic acid can be annealed to the nucleic acid fragment. In some examples, the first adaptor can be appended by annealing a random primer containing random sequence to the nucleic acid fragment, as shown in FIG. 14. In this example, the random primer can be a composite primer (e.g., a 3' portion containing DNA and a 5' portion containing RNA). The nucleic acid fragment can be single stranded. A double-stranded nucleic acid fragment can be denatured to generate a single-stranded nucleic acid fragment. Denaturation can be achieved using any of the methods known in the art including, but not limited to, physical, thermal, and/ or chemical denaturation. Denaturation can be achieved using thermal or heat denaturation. The single stranded target nucleic acid sequence of interest can be present in one or more single stranded nucleic acid fragments. Different or distinct target nucleic acid sequences of interest can be present in one or more of the single stranded nucleic acid fragments. One or more oligonucleotides can comprise sequence complementary to the same sequence of interest present in one or more single stranded nucleic acid fragments. The one or more oligonucleotides can comprise sequence that is complementary to different parts or regions of the same sequence of interest. The different regions can be adjacent to each other. The different regions can be non-adjacent to each other. The one or more oligonucleotides that comprise sequence complementary to the same target nucleic acid sequence of interest can further comprise the same first adaptor sequence. One or more oligonucleotides can comprise sequence complementary to different or distinct sequences of interest which can be present in one or more single stranded nucleic acid fragments. The one or more oligonucleotides that comprise sequence complementary to different or distinct target nucleic acid sequences of interest can further comprise the same first adaptor sequence. The sequence complementary to the target sequence of interest can be at the 3'end of the oligonucleotide and the first adaptor sequence can be at the 5'end of the oligonucleotide. The first adaptor sequence can be non-complementary to the target nucleic acid sequence of interest. The first adaptor sequence can serve as a tail. The first adaptor sequence can be a conventional adaptor sequence. The first adaptor can comprise reverse adaptor sequence. The first adaptor sequence can be of known sequence and, thus, can allow generation and/or use of sequence specific primers for amplification and/or sequencing of any polynucleotides to which the first adaptor sequence is appended or attached. The oligonucleotide can be annealed to the nucleic acid fragments without prior denaturation. Annealing of the oligonucleotide can be via formation of a triple helix or triplex between the oligonucleotide and a double stranded nucleic acid fragment. The double stranded nucleic acid fragment can comprise a sequence of interest and can be present amongst a plurality of double stranded nucleic acid fragments. The oligonucleotide can comprise sequence complementary to the sequence of interest in the double stranded nucleic acid fragment. The use of the oligonucleotide comprising sequence complementary to a target sequence of interest present in a nucleic acid fragment amongst one or more of a plurality of nucleic acid fragments can allow for selective binding and subsequent enrichment of the nucleic acid fragment using the methods described herein. In some cases, the annealing can occur in solution.

[0045] Following annealing of the oligonucleotide as described above, a polymerase can be used to extend the oligonucleotide. The polymerase can be a DNA dependent DNA polymerase. The DNA dependent DNA

polymerase can be any of the DNA dependent DNA polymerases as described herein and extension of the oligonucleotide can be by any of the methods known in the art. An oligonucleotide comprising the first adaptor sequence, wherein the first adaptor sequence is not complementary to the target nucleic acid, and sequence complementary to a target sequence of interest present in a nucleic acid fragment can be annealed to the nucleic acid fragment and extended with a polymerase to generate an oligonucleotide extension product comprising the first adaptor sequence at a first end. The nucleic acid fragment can be present amongst a plurality of nucleic acid fragments. In this embodiment, the oligonucleotide extension product can only be generated for a nucleic acid fragment that contains the target sequence of interest.

[0046] After generating a nucleic acid molecule comprising a first adaptor sequence at a first end by any of the methods described above (e.g., ligation or priming reaction), a second adaptor sequence may be appended to the nucleic acid molecule at a second end. The second adaptor sequence can be appended to the nucleic acid molecule at a second end by ligation, for example, as shown in FIG. 11. The second adaptor sequence can be appended to the nucleic acid molecule at a second end by a priming reaction.

[0047] In some embodiments, the second adaptor sequence is appended to a nucleic acid fragment comprising a first adaptor at a first end by a priming reaction. In one embodiment, the nucleic acid fragment is double stranded. In one embodiment, the double stranded nucleic acid fragment is denatured prior to the priming reaction to generate a single stranded nucleic acid fragment comprising a first adaptor sequence at a first end. Denaturation can be achieved using any of the methods as described above. In one embodiment, denaturation of the first adaptor-nucleic acid fragment complex generates single stranded nucleic acid fragments comprising first adaptor sequence at only the 5'end of the nucleic acid fragments as depicted, for example, in FIG. 2. In another embodiment, denaturation of the first adaptor-nucleic acid fragment complex generates single stranded nucleic acid fragments comprising first adaptor sequence at both the 5'end and 3'end of the nucleic acid fragments as depicted, for example, in FIG.1.

[0048] In one embodiment, the nucleic acid fragments comprising first adaptor sequence appended to either the 5' end or both the 5' and 3' end can be denatured to generate single stranded nucleic acid fragments comprising first adaptor sequence appended to either the 5' end or both the 5' and 3' end. In one embodiment, the methods of the present invention described herein can be used to generate a plurality of single stranded nucleic acid fragments comprising first adaptor sequence appended to either the 5' end or both the 5' and 3' end.

[0049] In some embodiments, a second adaptor sequence may be appended to a single stranded nucleic acid fragment comprising first adaptor sequence appended to either the 5' end or both the 5' and 3' ends. In one embodiment, appending a second adaptor sequence to a single stranded nucleic acid fragment comprising first adaptor sequence appended to either the 5' end or both the 5' and 3' ends can comprise a priming reaction. In this embodiment, an oligonucleotide comprising, at a first end, sequence complementary to a target sequence of interest present in a single stranded nucleic acid fragment, and at a second end, sequence from a second adaptor, wherein the second adaptor sequence is not complementary to the target nucleic acid can be annealed to the single stranded nucleic acid fragments. In one embodiment, the second adaptor sequence can be sequence from a

reverse adaptor. In one embodiment, the target nucleic acid sequence of interest can be present in one or more of the single stranded nucleic acid fragments. In one embodiment, different or distinct target nucleic acid sequences of interest can be present in one or more of the single stranded nucleic acid fragments. In one embodiment, one or more oligonucleotides can comprise sequence complementary to the same sequence of interest present in one or more single stranded nucleic acid fragments. In this embodiment, the one or more oligonucleotides can comprise sequence that is complementary to different parts or regions of the same sequence of interest. In one embodiment, the different regions can be adjacent to each other. In one embodiment, the different regions can be non-adjacent to each other. In a preferred embodiment, the one or more oligonucleotides that comprise sequence complementary to the same target nucleic acid sequence of interest further comprise the same second adaptor sequence. In another embodiment, one or more oligonucleotides can comprise sequence complementary to different or distinct sequences of interest which can be present in one or more single stranded nucleic acid fragments. In a preferred embodiment, the one or more oligonucleotides that comprise sequence complementary to different or distinct target nucleic acid sequences of interest further comprise the same second adaptor sequence. In one embodiment, the sequence complementary to the target sequence of interest can be at the 3'end of the oligonucleotide and the second adaptor sequence can be at the 5'end of the oligonucleotide. In a preferred embodiment, the second adaptor sequence is non-complementary to the target nucleic acid sequence of interest. In this manner, the second adaptor sequence serves as a tail. The second adaptor sequence can be a conventional adaptor sequence. In a preferred embodiment, the second adaptor sequence can be a conventional adaptor sequence that is different than or distinct from the sequence of the first adaptor appended to the single stranded nucleic acid fragment as described above. In one embodiment, the second adaptor sequence can be of known sequence and, thus, allow generation and/or use of sequence specific primers for amplification and/or sequencing of any polynucleotides to which the second adaptor sequence is appended or attached. In a separate embodiment, the oligonucleotide can be annealed to the nucleic acid fragments comprising the first adaptor sequence appended to either the 5' end or both the 5' and 3' end without prior denaturation. In this embodiment, annealing of the oligonucleotide can be via formation of a triple helix or triplex between the oligonucleotide and a double stranded nucleic acid fragment comprising the first adaptor sequence appended to either the 5' end or both the 5' and 3' ends of the double stranded nucleic acid fragment. In this embodiment, the double stranded nucleic acid fragment comprises a sequence of interest and can be present amongst a plurality of double stranded nucleic acid fragments comprising first adaptor sequence appended to either the 5' end or both the 5' and 3' end. Further to this embodiment, the oligonucleotide comprises sequence complementary to the sequence of interest in the double stranded nucleic acid fragment. Overall, the use of the oligonucleotide comprising sequence complementary to a target sequence of interest present in a nucleic acid fragment amongst one or more or a plurality of nucleic acid fragments allows for selective binding and subsequent enrichment of said nucleic acid fragment using the methods described herein.

[0050] Following annealing of the oligonucleotide as described above, a polymerase can be used to extend the oligonucleotide. In one embodiment, the polymerase can be a DNA dependent DNA polymerase. In one

embodiment, the DNA dependent DNA polymerase can be any of the DNA dependent DNA polymerases as described herein and extension of the oligonucleotide can be by any of the methods known in the art. In one embodiment, an oligonucleotide comprising the second adaptor sequence, wherein the second adaptor sequence is not complementary to the target nucleic acid, and sequence complementary to a target sequence of interest present in a nucleic acid fragment comprising a first adaptor appended to one and/or both ends can be annealed to the nucleic acid fragment and extended with a polymerase to generate an oligonucleotide extension product comprising the first adaptor sequence at a first end and the second adaptor sequence at a second end. In some examples, the oligonucleotide extension product may form an RNA/DNA heteroduplex, for example, as depicted in FIG. 14. In one embodiment, the nucleic acid fragment can be present amongst a plurality of nucleic acid fragments comprising first adaptor appended to one and/or both ends. In this embodiment, the oligonucleotide extension product can only be generated for a nucleic acid fragment that contains the target sequence of interest.

[0051] A second adaptor sequence can be appended to a second end of a nucleic acid molecule comprising a first adaptor sequence at a first end by ligation as shown in, for example, FIG. 11. The nucleic acid molecule can be an oligonucleotide extension product comprising a first adaptor sequence at a first end as disclosed herein. The second adaptor sequence can be sequence from a forward adaptor. Ligation of the second adaptor to the nucleic acid fragment can follow end repair of the oligonucleotide extension product. Ligation of the second adaptor to the oligonucleotide extension product can follow generation of the oligonucleotide extension product without end repair of the oligonucleotide extension product. The second adaptor can be any type of adaptor known in the art including, but not limited to, conventional duplex or double stranded adaptors in which the adaptor comprises two complementary strands. The second adaptor can be a double stranded DNA adaptor. The second adaptor can be a single stranded DNA adaptor. The second adaptor can be an oligonucleotide of known sequence and, thus, allow generation and/or use of sequence specific primers for amplification and/or sequencing of any polynucleotides to which the second adaptor(s) is appended or attached. The second adaptor can be a conventional duplex adaptor, wherein the second adaptor comprises sequence well known in the art. The second adaptor can be appended to the oligonucleotide extension product generated by the methods described herein in multiple orientations. The methods described herein can involve the use of a second duplex adaptor comprising double stranded DNA of known sequence that is blunt ended and can bind to the oligonucleotide extension product comprising a first adaptor at a first end generated by the methods described herein in one of two orientations. The second adaptor can be ligated to each of the oligonucleotide extension products such that each of the oligonucleotide extension products comprises the same second adaptor. Each of the oligonucleotide extension products can comprise a common second adaptor. A second adaptor can be appended or ligated to a library of oligonucleotide extension products generated by the methods described herein such that each oligonucleotide extension products in the library of oligonucleotide extension products comprises the second adaptor ligated to one or both ends.

[0052] The second adaptor can be ligated or appended to a 5' and/or 3' end of the oligonucleotide extension product generated by the methods described herein. The second adaptor can comprise two strands wherein each strand comprises a free 3' hydroxyl group but neither strand comprises a free 5' phosphate. The free 3' hydroxyl

group on each strand of the second adaptor can be ligated to a free 5' phosphate present on either end of the oligonucleotide extension products. The second adaptor can comprise a ligation strand and a non-ligation strand whereby the ligation strand can be ligated to the 5'phosphate on either end of the oligonucleotide extension product while a nick or gap can be present between the non-ligation strand of the first adaptor and the 3' hydroxyl on either end of the oligonucleotide extension product. The nick or gap can be filled in by performing a gap repair reaction. The gap repair can be performed with a DNA dependent DNA polymerase with strand displacement activity. The gap repair can be performed using a DNA dependent DNA polymerase with weak or no strand displacement activity. The ligation strand of the second adaptor can serve as the template for the gap repair or fill-in reaction. The gap repair or fill-in reaction can comprise an extension reaction wherein the ligation strand of the second adaptor serves as a template and leads to the generation of oligonucleotide extension products with complementary termini or ends. The gap repair can be performed using *Taq* DNA polymerase. In some cases, the ligation of the second adaptor to the nucleic acid fragments generated by the methods described herein may not be followed gap repair. In this embodiment, the oligonucleotide extension products comprise second adaptor sequence ligated only at the 5' end of each strand.

[0053] Methods provided herein can be used to generate nucleic acid molecules (i.e., oligonucleotide extension products) comprising a first adaptor sequence at a first end and a second adaptor sequence at a second end. Any method or any combinations of methods as disclosed herein can be used to append one or more adaptors to one or more ends of a nucleic acid fragment. The first adaptor and the second adaptor can be appended by ligation. The first adaptor and the second adaptor can be appended by a priming reaction as shown, for example, in FIG. 10. The first adaptor can be appended by ligation and the second adaptor can be appended by a priming reaction. The first adaptor can be appended by a priming reaction and the second adaptor can be appended by ligation. In one example, the first adaptor can be appended by annealing a primer comprising a target specific sequence at the 3'-end and the first adaptor sequence at a second end, to the nucleic acid of interest followed by extension of the primer with DNA polymerase to generate a partial duplex comprising a blunt end. In this example, the second adaptor can be appended to the partial duplex by ligating the second adaptor to the blunt end without the need for an end repair process, as shown in FIG. 11.

[0054] In one embodiment, the oligonucleotide extension products comprising a first adaptor at a first end and a second adaptor at a second end as generated by the methods described herein can be subjected to an amplification reaction. In one embodiment, the amplification reaction can be exponential, and can be carried out at various temperature cycles or can be isothermal. In one embodiment, the amplification can be polymerase chain reaction. In one embodiment, the amplification reaction can be isothermal. In one example, the oligonucleotide extension product can be separated from the template nucleic acid fragment in order to generate a single stranded oligonucleotide extension product with first adaptor sequence on the 5' end and second adaptor sequence on the 3' end. The single stranded oligonucleotide extension product can then be amplified using a first primer comprising sequence complementary to the first adaptor and a second primer comprising sequence complementary to the second adaptor sequence. In an alternative example, the oligonucleotide extension product

can be separated from the template nucleic acid fragment in order to generate a single stranded oligonucleotide extension product with a first adaptor sequence on the 3' end and a second adaptor sequence on the 5' end. The single stranded oligonucleotide extension product can then be amplified using a first primer that anneals to the complement of the second adaptor sequence and a second primer that anneals to the complement of the first adaptor sequence. In this manner only oligonucleotide extension products comprising both the first and the second adaptor sequence will be amplified and thus enriched. In one embodiment, the first adaptor and/or the second adaptor sequence can comprise an identifier sequence. In one embodiment, the identifier sequence can be barcode sequence. In one embodiment, the barcode sequence can be the same or different for the first adaptor and the second adaptor sequence. In one embodiment, the first adaptor and/or the second adaptor sequence can comprise sequence that can be used for downstream applications such as, for example, but not limited to, sequencing. In one embodiment, the first adaptor and/or the second adaptor sequence can comprise flow cell sequences which can be used for sequencing with the sequencing method developed by Illumina and described herein.

[0055] In an aspect of the disclosure, a method for enriching a nucleic acid sequence of interest in a sample comprising nucleic acids is provided comprising combining a transposome comprising a transposase and a transposon sequence with nucleic acids in the sample. The nucleic acids can be fragmented generating a nucleic acid fragment comprising the nucleic acid sequence of interest. The transposon sequence can be appended to a first end of the nucleic acid fragment generating a nucleic acid fragment comprising a transposon sequence at a first end. A "transposase" can be any enzyme that catalyzes the transfer of a transposon (or transposable element) from one region of the genome to another. A transposase enzyme (as part of a transposome complex) can be used to tag a nucleic acid molecule with the transposon by fragmenting the nucleic acid to generate nucleic acid fragments and by appending a transposon sequence to the nucleic acid fragments at a first end thereby generating nucleic acid fragments comprising the transposon sequence at a first end. The transposase can form a functional complex with a transposon sequence that is capable of catalyzing a transposition reaction. The transposase can bind to the transposon sequence and append the transposon sequence to the nucleic acid molecules in a process sometimes termed "tagmentation." Tagmentation can comprise a fragmentation step and an appending step. The fragmentation step and the appending step can be carried out by the transposase enzyme. The fragmenting step and the appending step can be simultaneous or near simultaneous. Fragmentation can generate nucleic acid fragments of various lengths as provided herein. The fragmenting step and appending step can be random. The fragmenting step and the appending step can be site-selective. When fragmentation of a nucleic acid is carried out by the transposase enzyme, it may not be necessary to perform other fragmentation steps as disclosed herein (e.g., mechanical shearing, sonicating, etc).

[0056] Any transposase that can tag a nucleic acid molecule with a transposon sequence may be used. Methods of tagging nucleic acid molecules with transposases are known in the art and kits for performing such methods are commercially available (e.g., Nextera DNA-sample prep kit from Epicentre). In some examples, the transposase

enzyme is the Tn5 transposase. In other examples, the transposase enzyme may be Tn10 transposase, Mu transposase, or P-element transposase.

[0057] The nucleic acid fragment may be double stranded. The double stranded nucleic acid fragment can be denatured, thereby generating single stranded nucleic acid fragments. The transposon sequence can be appended to a first end of the nucleic acid fragment. The transposon sequence can have a known sequence. Tailed primers can be used to append an adaptor to the nucleic acid fragment. One or more first oligonucleotides can be annealed to the nucleic acid fragments comprising the transposon sequence at a first end. The one or more first oligonucleotides can comprise a 3' portion that is complementary to the transposon sequence and a 5' portion comprising a first adaptor sequence. The one or more first oligonucleotides annealed to the transposon sequence can be extended with a polymerase, thereby generating one or more first oligonucleotide extension products comprising sequence complementary to the nucleic acid fragments and the first adaptor sequence at the first end. The polymerase may be a DNA polymerase. Any method of annealing and extending (e.g., a priming reaction) as disclosed herein may be used. After generating a nucleic acid fragment comprising a first adaptor at a first end by the methods described herein, any method of appending a second adaptor to the nucleic acid fragment can be used (e.g., ligation or priming reaction). One or more second oligonucleotides can be annealed to the nucleic acid sequence of interest in the nucleic acid fragments comprising the first adaptor sequence at the first end. The one or more second oligonucleotides can comprise a 3' portion that is complementary to the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence. Each of the one or more second oligonucleotides can comprise a 3' portion that is complementary to a portion of a different nucleic acid sequence of interest in the nucleic acid fragments and a 5' portion comprising the second adaptor sequence. The one or more second oligonucleotides can be extended, thereby generating one or more second oligonucleotide extension products comprising sequence complementary to the first adaptor sequence at the first end, the sequence complementary to the nucleic acid sequence of interest, and the second adaptor sequence at the second end. The one or more second oligonucleotide extension products can be amplified using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest. A library of nucleic acid sequences may be created. FIG. 12 illustrates an example of selective target enrichment using a transposome-mediated tagmentation protocol.

[0058] In an alternate embodiment, the methods of the present invention can be used to generate a library of nucleic acid fragments or inserts wherein each nucleic acid fragment comprises an adaptor at one or both ends. In one embodiment, the adaptors can be present at both ends and can be distinct from each other. In one embodiment, the adaptors can be present at both ends and can comprise the same adaptor sequence. The generation of the library comprising nucleic acid inserts with distinct adaptors at both ends can involve the methods for generating oligonucleotide extension products comprising first adaptor sequence on one end and second adaptor sequence on the other end as described herein with the exception that the oligonucleotide that binds to the nucleic acid fragments and can be extended comprises random sequence. The first adaptor may be appended to the nucleic acid fragment by ligation and the second adaptor may subsequently be appended by a

priming reaction, as described herein. The oligonucleotide can comprise random sequence at the 3' portion that is hybridizable to one or more nucleic acid fragments and can further comprises second adaptor sequence at the 5'-portion. Extension of the oligonucleotide along the nucleic acid fragment and the corresponding first adaptor can generates a product, or products, comprising the second adaptor at one end and a sequence complementary to the first adaptor at the other end, as illustrated in Figure 4. The first adaptor can be appended to the nucleic acid fragment by a priming reaction and the second adaptor can subsequently be appended by ligation, as described herein. The oligonucleotide can comprise random sequence at the 3' portion that is hybridizable to one or more nucleic acid fragments and can further comprise first adaptor sequence at the 5'-portion. Extension of the oligonucleotide along the nucleic acid fragment can generate an oligonucleotide extension product, or products, comprising the first adaptor at a first end. The second adaptor can then be ligated to a second end of the oligonucleotide extension product generating an oligonucleotide extension product comprising a first adaptor sequence at a first end and a second adaptor sequence at a second end. Random sequence present in the oligonucleotide can bind to and be extended on one or more nucleic acid inserts. One or more oligonucleotides comprising a 3' portion comprising random sequence and a 5' portion comprising second adaptor sequence can be annealed to a library comprising nucleic acid inserts comprising a first adaptor sequence on one or both ends of each of the nucleic acid inserts. One or more second adaptors can be ligated to a library comprising nucleic acid inserts comprising a first adaptor sequence on one end of each of the nucleic acid inserts. The first adaptor can be the same or common to each of the nucleic acid inserts. The first adaptor sequence can be the same or common to each of the one or more oligonucleotides. The second adaptor can be the same or common to each of the nucleic acid inserts. The second adaptor sequence can be the same or common to each of the one or more oligonucleotides. The methods described above can be used to generate a library of nucleic acid inserts wherein each of the nucleic acid inserts comprises a common first adaptor on one end and a common second adaptor sequence on a second end. The first adaptor and the second adaptor sequence can be distinct from each other. The first adaptor and the second adaptor sequence can comprise the same adaptor sequence. The methods described herein can be used for the high efficiency generation of a library of nucleic acid sequences.

[0059] In yet another alternate embodiment to the methods of the invention as described above, the first adaptor can be a double stranded DNA adaptor comprising a partial duplex, wherein the two strands of the adaptor can be different lengths with a complementary region and an overhanging region at the 5' end. In this embodiment, the 5' end of the long strand of the partial duplex adaptor can comprise a unique site for a nucleic acid modifying enzyme, such as a restriction enzyme, that is absent from the short strand of the duplex adaptor. In a further embodiment, the 3' end of the short strand adaptor can be modified by a replacement of the 3' OH-group by a blocking group, for example, a dideoxynucleotide (ddCMP, ddAMP, ddTMP, or ddGMP) to prevent polymerase extension. In this embodiment, the first adaptor comprising the partial duplex can be ligated to nucleic acid fragments generated by the methods described herein. In one embodiment, ligation of the partial duplex first adaptor can be followed by a gap repair reaction as described above. In this embodiment, ligation of the partial duplex first adaptor is not followed by a gap repair reaction. In a preferred embodiment, the partial duplex first

adaptor comprises a free 5' phosphate on the short strand and a free 3'hydroxyl on the long strand. In this embodiment, ligation of the partial duplex adaptor generates double stranded nucleic acid fragments wherein both ends of the double stranded nucleic acid fragment comprise the long strand and short strand of the partial duplex first adaptor. A double stranded partial duplex first adaptor-nucleic acid fragment complex can be generated by ligation. In one embodiment, the double stranded partial duplex first adaptor-nucleic acid fragment complex can be denatured to generate a single stranded nucleic acid fragment comprising the long strand of the first adaptor on a first end and the short strand of the first adaptor on a second end. In this embodiment, the first end is the 5' end and the second end is the 3' end. In one embodiment, the first adaptor can be appended to one or more nucleic acid fragments as generated by the methods described herein such that each of the nucleic acid fragments comprises the same first adaptor or, in other words, the first adaptor can be common to each of the nucleic acid inserts. An oligonucleotide or primer comprising sequence complementary to a sequence of interest in the single stranded nucleic acid fragment can be annealed to the single stranded nucleic acid fragment and extended using a polymerase. In one embodiment, the polymerase can be a DNA dependent DNA polymerase. In one embodiment, the DNA dependent DNA polymerase can be any of the DNA dependent DNA polymerases as described herein and extension of the oligonucleotide can be by any of the methods known in the art. Extension of the primer annealed to the single stranded nucleic acid fragment generates an oligonucleotide extension product comprising sequence complementary to the long strand of the first adaptor on one end. In one embodiment, the oligonucleotide extension product remains hybridized to the single stranded nucleic acid fragment such that the restriction and/or cleavage site specific for a nucleic acid modifying enzyme is made double stranded. The double stranded site can then be cleaved by the nucleic acid modifying enzyme specific for the double stranded restriction site. In one embodiment, the nucleic acid modifying enzyme can be a restriction enzyme. In one embodiment, the restriction enzyme can be specific for a double stranded restriction site. In one embodiment, cleavage of the restriction site can generate a blunt end or non-blunt end. In one embodiment, end repair by any of the methods described herein can be performed on the end of the nucleic acid fragment following cleavage. Cleavage of the restriction and/or cleavage site generates a site to which a second adaptor can be ligated. Ligation of the second adaptor can be through any of the methods for ligation as described herein. In one embodiment, ligation generates a double stranded nucleic acid fragment comprising the second adaptor on a first end and a partial duplex on a second end, wherein the partial duplex comprises a 3' overhang comprising the sequence of the short strand of the first adaptor. The double stranded nucleic acid fragment can then be denatured using any of the methods for denaturation disclosed herein to generate a single stranded nucleic acid fragment comprising the second adaptor sequence on the first end and the sequence of the short strand of the first adaptor on the second end. In one embodiment, the first end and second end comprise the 5' end and 3' end, respectively. In one embodiment, the second adaptor can be appended to one or more nucleic acid fragments following cleavage of the double stranded restriction site such that each of the nucleic acid fragments comprises the same second adaptor or, in other words, the second adaptor can be common to each of the nucleic acid inserts. The single stranded nucleic acid fragment can then be amplified using a first primer specific for the second adaptor and a second primer

specific for sequence present in the short strand of the first adaptor. In one embodiment, the amplification reaction can be exponential, and may be carried out at various temperature cycles or isothermal. In one embodiment, the amplification can be polymerase chain reaction. In one embodiment, the amplification reaction can be isothermal. Overall, only a fragment comprising the second adaptor and the short strand of the first adaptor will be amplified or enriched. In so far as the method provides for enrichment of targeted fragments of the library, and not enrichment of oligonucleotide extension products generated by the extension of the oligonucleotide comprising sequence complementary to a target sequence of interest, there is no distortion of the original DNA library, and the enrichment is independent of the insert length. Because the 3' end of the short strand of the partial duplex adaptor is 3' blocked, the method enables directional or asymmetric ligation. In one embodiment, the oligonucleotide that comprises sequence complementary to a sequence of interest in a nucleic acid fragment further comprises reverse adaptor sequence. In this embodiment, the sequence complementary to a sequence of interest in the nucleic acid fragment can be present in a 3' portion of the oligonucleotide and the reverse adaptor sequence can be present at a 5' portion. Further to this embodiment, the reverse adaptor sequence can be a common or conventional adaptor sequence and can be different or distinct from the first and/or second adaptors. Further still to this embodiment, the methods described above can lead to the generation of a single stranded nucleic acid fragment comprising the second adaptor at one end and the reverse adaptor sequence at the other end. Subsequent to this embodiment, the single stranded nucleic acid fragment can be enriched through amplification using a first primer specific to the second adaptor and a second primer specific to the third adaptor sequence.

[0060] The methods of the inventions are further applicable to any enrichment of target nucleic acid sequences of interest from libraries comprising fragments of nucleic acid of a sample appended with adaptor sequence at one or both ends, wherein the libraries are generated using ligation of the adaptor or adaptor sequences to one or both ends as described herein or by ligation independent methods, such as for example Nextera, a transposome driven method. In one embodiment, the nucleic acid can be DNA such as genomic DNA or cDNA. In one embodiment, the nucleic acid can be double stranded. Enrichment of nucleic acid sequences of interest can be achieved using the methods described herein for target enrichment. In one embodiment, the method for enriching for target nucleic acid sequences of interest from a library comprising nucleic inserts with adaptors appended to one or both ends comprises denaturing the nucleic acid inserts to generate a library of single stranded nucleic acid inserts. In one embodiment, each of the nucleic acid inserts can comprise a first adaptor sequence on one end and a second adaptor sequence on an opposite end. In one embodiment, the first adaptor and the second adaptor can be distinct from each other. In one embodiment, the first adaptor and the second adaptor can comprise the same adaptor sequence. In one embodiment, each of the nucleic acid inserts can comprise a first adaptor sequence on one end and a second adaptor sequence on an opposite end such that denaturation generates a library of single stranded nucleic acid inserts comprising the first adaptor sequence on one end and the second adaptor sequence on an opposite end. Denaturation can be achieved using any of the methods described herein. Further to the embodiments described above, one or more oligonucleotides can be annealed to the single stranded nucleic acid

inserts. In one embodiment, each of the one or more oligonucleotides comprises a 3' portion that is complementary to a target nucleic acid sequence of interest present in one or more of the nucleic acid inserts, and a 5' portion comprising a third adaptor sequence. In one embodiment, the third adaptor sequence is distinct from either or both of the first adaptor and the second adaptor. The one or more oligonucleotides can be extended with a polymerase (e.g., a DNA polymerase) thereby generating one or more oligonucleotide extension products with the first or second adaptor at a first end and the third adaptor sequence at a second end. In one embodiment, the first end comprises the 5' end and the second end comprises the 3' end. The one or more oligonucleotide extension products can be amplified using a first primer that can be complementary to the first or second adaptor and a second primer that can be complementary to the third adaptor sequence to enrich for nucleic acid fragments comprising the first or second adaptor and the third adaptor sequence at each end. In one embodiment, the first and second adaptors can be common to each of the nucleic acid inserts in the library. In one embodiment, the third adaptor sequence can be common to each of the one or more oligonucleotides. Overall, the target enrichment methods as described above can be used to generate a composition comprising a library of nucleic acid inserts enriched for any target sequence of interest from a non-enriched library comprised of nucleic acid inserts with an adaptor ligated to one or both ends.

[0061] A schematic of a preferred embodiment of the methods described herein for enriching for target sequences of interest is illustrated in FIG. 1. Overall, FIG. 1 depicts a method for isolating or enriching for a nucleic acid fragment or insert comprising a target nucleic acid sequence from a library or plurality of nucleic acid fragments. The method in FIG. 1 involves generation of a ligated library of nucleic acid fragments or inserts wherein each fragment or insert of the ligated library comprises a common forward adaptor and a fragment or insert specific reverse adaptor distinct from the forward adaptor such that subtractive PCR using a primer directed against the common forward adaptor and a primer directed against the reverse adaptor enriches for a nucleic acid fragment or insert comprising a target nucleic acid sequence. The input for the method depicted in FIG.1 is fragmented DNA. The fragmented DNA is double stranded and comprises a plurality or library of DNA fragments. In one embodiment, the DNA fragments can be derived from complex DNA, such as double-stranded DNA, genomic DNA or mixed DNA from more than one organism. In one embodiment, the DNA fragments can be derived from RNA that has been converted to cDNA through a first strand synthesis reaction using any of the methods well known in the art for generating cDNA from an RNA template which can include, but is not limited to, combining the RNA with a primer (e.g., random primer), and reverse transcribing the RNA template with an RNA-dependent DNA polymerase. In one embodiment, the DNA fragments can be derived from RNA that has been converted to double stranded cDNA through a first and second strand synthesis reaction using any of the methods well known in the art. Fragmentation of the DNA to produce the DNA fragments can be achieved through any of the methods described herein for fragmenting nucleic acids which can include, but are not limited to, physical (e.g., sonication), and/or chemical (e.g., restriction enzyme treatment) fragmentation reactions.

[0062] As depicted in FIG. 1, a single forward adaptor is ligated to the DNA fragments. In one embodiment, the single forward adaptor can comprise known sequence. In one embodiment, the single forward adaptor can be a

common adaptor. In one embodiment, the DNA fragments can be subjected to an end repair reaction as described herein to produce blunt ends. In this embodiment, the single forward adaptor can also comprise blunt ends and ligation between the single forward adaptor and the DNA fragments can be through blunt end ligation as described herein. Ligation can be facilitated through the use of enzymes (e.g., T4 DNA ligase) and methods known in the art, including, but not limited to, commercially available kits such as the Encore ™ Ultra Low Input NGS Library System. In FIG. 1, the forward adaptor can contain a strand (the ligation strand) that ligates with the free 5'phosphate on a 5' end of the DNA fragments and a strand that does not ligate (non-ligation strand) to a 3' end of the DNA fragments. In one embodiment, the ligation reaction can lead to the generation of a nick or gap between the non-ligation strand of the single forward adaptor and the 3' end of the DNA fragments. In this embodiment, the nick or gap can be repaired or filled in through a gap repair or fill-in reaction wherein the 3' end of the DNA fragments can be extended with a polymerase (preferably with a DNA dependent DNA polymerase such as *Taq* DNA polymerase) wherein the ligation strand of the forward adaptor can serve as template. In this embodiment, the gap repair generates DNA fragments with complementary ends. As depicted in FIG. 1, the DNA fragments with complementary ends are denatured to generate a denatured library comprising single stranded DNA fragments with complementary ends. Denaturation can be achieved using any of the methods known in the art which can include, but are not limited to, heat denaturation, and/or chemical denaturation.

[0063] As depicted in FIG. 1, a custom oligonucleotide with a reverse adaptor tail is annealed to the single stranded DNA fragments with complementary ends. In one embodiment, the custom oligonucleotide with a reverse adaptor tail can comprise a 3' portion comprising sequence complementary to a target sequence of interest in one of the single-stranded DNA fragments and a 5' portion comprising reverse adaptor sequence that is not complementary to the single-stranded DNA fragments in the denatured library. In one embodiment, the reverse adaptor sequence can be known sequence. In a preferred embodiment, the reverse adaptor sequence can be distinct from the single forward adaptor as described herein. In one embodiment, a plurality of custom oligonucleotides with a reverse adaptor tail can be added to the denatured library wherein the plurality of custom oligonucleotides with a reverse adaptor tail comprise a 3' portion comprising sequence complementary to a target sequence of interest in one or more of the single-stranded DNA fragments of the denatured library and a 5' portion comprising a reverse adaptor sequence that is not complementary to the single-stranded DNA fragments in the denatured library. In one embodiment, the reverse adaptor tail comprises the same reverse adaptor sequence in each of the plurality of custom oligonucleotides with a reverse adaptor tail, and wherein the reverse adaptor sequence is distinct from the forward adaptor sequence. In another embodiment, the reverse adaptor tail comprises a different reverse adaptor sequence for each of the plurality of custom oligonucleotides with a reverse adaptor tail, and wherein each of the different reverse adaptor sequences is distinct from the forward adaptor sequence. In one embodiment, the 3' portion of the custom oligonucleotide with a reverse adaptor tail can be a specific sequence, wherein the custom oligonucleotide comprises a sequence complementary to the target sequence of interest and provides a means for targeted enrichment of sequence or sequences of interest using the methods of the invention. In another embodiment the 3' portion of the custom oligonucleotide with reverse

adaptor tail can be a randomly generated sequence hybridizable to random sequences of the library of fragments with adaptor sequences on one or both ends, providing means for efficient, non-enriched library generation employing the methods of the invention.

[0064] Following annealing of the custom oligonucleotide with a reverse adaptor tail to a sequence of interest in a single-stranded DNA fragment of the denatured library, the custom oligonucleotide with a reverse adaptor tail is extended using any method known in the art, which can include but is not limited to, extension using a DNA dependent DNA polymerase using the single stranded DNA fragment of the denatured library as a template. Extension of the custom oligonucleotide with a reverse adaptor tail generates an oligonucleotide extension product with forward adaptor sequence at one end and reverse adaptor sequence at the other end. In this embodiment, the custom oligonucleotide with a reverse adaptor tail can only anneal to and be extended on DNA fragments in the denatured library comprising the target sequence of interest for which the custom oligonucleotide with a reverse adaptor tail is directed. As illustrated in FIG. 1, a subtractive polymerase chain reaction (PCR) procedure is subsequently performed using a first primer directed against the forward adaptor sequence and a second primer directed toward the reverse adaptor sequence such that only the oligonucleotide extension product with the forward adaptor sequence at one end and the reverse adaptor sequence at the other end can be amplified and thus enriched.

[0065] FIG. 2 depicts another embodiment of subtractive PCR enrichment method as described for FIG. 1, wherein ligation of a duplex forward adaptor (P1) to a double stranded nucleic acid fragment is not subjected to gap repair. The duplex forward adaptor (P1) comprises a strand (the ligation strand; P1) that ligates with the free 5'phosphate on a 5' end of the nucleic acid fragment and a strand that does not ligate (non-ligation strand; P1rc) to a 3' end of the nucleic acid fragment. In this embodiment, the ligation strand is ligated to the 5' end of both strands of a double stranded nucleic fragment whereas a gap or nick is generated between the non-ligation strand and the 3'end of both strands of the double stranded nucleic acid fragment. As depicted in FIG. 2, the ligation of the P1 adaptor to the nucleic acid fragment is followed by denaturation without a gap repair or fill-in reaction thereby generating a single stranded nucleic acid fragment with non-complementary ends. In this embodiment, the single stranded nucleic acid fragment with non-complementary ends comprises a P1 forward adaptor sequence at a 5' end and fragment specific sequence at a 3' end. In a further embodiment, the single stranded nucleic fragment with non-complementary ends can be further processed as described above and illustrated in FIG. 1 to generate a single stranded nucleic fragment with the P1 forward adaptor sequence at the 5' end and a distinct reverse adaptor sequence on the 3' end.

[0066] FIG. 4 illustrates another embodiment of the present invention for the high efficiency generation of libraries comprising nucleic acid fragments with distinct adaptors on each end. In this embodiment, the methods for generating a denatured library comprising single stranded DNA fragments with a single forward adaptor sequence on a 5' end and a reverse adaptor sequence on a 3' end is similar to the method described for FIG. 1. However, in FIG. 4, an oligonucleotide with a reverse adaptor tail is used whereby the oligonucleotide comprises a 3' portion comprising random sequence and a reverse adaptor tail wherein the reverse adaptor tail comprises

reverse adaptor sequence that is distinct from the single forward adaptor sequence. As illustrated in FIG. 4, PCR can be carried out with a first primer directed against the single forward adaptor sequence and a second primer directed toward the reverse adaptor sequence wherein both the first and second primers further comprise flow cell sequences. In this manner, single stranded DNA fragments with a single forward adaptor sequence on a 5' end and a reverse adaptor sequence on a 3' end comprise flow cell sequences that can be used to adhere the amplified single stranded DNA fragments to flow cells for subsequent sequencing by the method commercialized by Illumina, as described U.S. Pat. Nos. 5,750,341; 6,306,597; and 5,969,119.

[0067] FIG. 3 illustrates a method for enrichment of a target nucleic acid sequence, or sequences of interest, contained in a double stranded nucleic acid insert from a complex library. In one embodiment, the complex library comprises nucleic acid inserts from a genomic DNA sample, In FIG. 3, the single forward adaptor comprises a partial duplex forward adaptor comprising a long strand, A, that forms a partial duplex with a short strand, B. Strand A of the partial duplex adaptor further comprises a restriction enzyme site while strand B does not contain the restriction enzyme site. Strand B further comprises a blocking group wherein the 3' end of strand B is modified by replacement of the 3' OH group with a blocking group that can prevent polymerase extension. In one embodiment, the partial duplex forward adaptor is ligated to the double stranded nucleic acid fragments such that a double stranded insert with the partial duplex forward adaptor appended to both ends is generated. In this embodiment, the 5' end of strand B of the partial duplex adaptor can contain a free 5' phosphate which can be ligated to a free 3' OH present on one or both strands of the double stranded insert. Subsequent denaturation generates a single stranded insert comprising sequence A on a 5'end and strand B on a 3' end. A primer, C, directed against a specific sequence of interest within the single stranded insert can be annealed to the specific sequence and extended with a DNA polymerase using the single stranded insert as a template. In one embodiment, primer C can be a sequence specific primer and is employed for enrichment of target, or targets, of interest according to the methods of the invention. In one embodiment, primer C can be a random primer. Extension of primer C with DNA polymerase generates an extended primer C product that comprises sequence complementary to sequence A at its' 3' end in a double stranded complex with the template insert strand such that a double stranded restriction site has been generated between sequence A and its' complement. In one embodiment, the double stranded restriction enzyme recognition site can be cleaved by a restriction enzyme specific for the double stranded restriction site thereby generating a truncated, or cleaved, adaptor sequence at one end. A second forward adaptor comprising a common, or conventional, duplex adaptor D is then ligated to the cleavage site using any of the ligation methods described herein, thereby generating a double stranded complex comprising the second forward adaptor D at one end and a 3' overhang comprising strand B on the opposite end. The double stranded complex comprising the second forward adaptor D at one end and a 3' overhang comprising strand B on the opposite end is denatured and amplified using a first primer directed against the second forward adaptor D and a second primer directed against strand B. In this manner, the methods depicted in FIG. 3 can be used to enrich for specific sequences of interest from a complex library since the methods are designed such that the second forward adaptor D can only bind to the double stranded cleavage site generated by restriction enzyme

digestion of the double stranded restriction site created between sequence A and its complement following extension of primer C. As described, primer C can be directed against a target sequence of interest present in a single, or multiple insert, or inserts, amongst a plurality of inserts. Moreover, the method can be made strand specific by designing primer C to bind to target sequence of interest present on one strand or the other of an insert in an amongst a plurality of inserts.

[0068] Unless otherwise specified, terms and symbols of genetics, molecular biology, biochemistry and nucleic acid 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.

Input Nucleic Acid

[0069] The input can be a nucleic acid. In one embodiment, the input can be DNA. In one embodiment, the input nucleic acid can be complex DNA, such as double-stranded DNA, genomic DNA or mixed DNA from more than one organism. In one embodiment, the input can be RNA. In one embodiment, the RNA can be obtained and purified using standard techniques in the art and include RNAs in purified or unpurified form, which include, but are not limited to, mRNAs, tRNAs, snRNAs, rRNAs, retroviruses, small non-coding RNAs, microRNAs, polysomal RNAs, pre-mRNAs, intronic RNA, viral RNA, cell free RNA and fragments thereof. The non-coding RNA, or ncRNA may include snoRNAs, microRNAs, siRNAs, piRNAs and long nc RNAs. In one embodiment, the input nucleic acid can be cDNA. The cDNA can be generated from RNA, e.g., mRNA. The cDNA can be single or double stranded. The input DNA can be of a specific species, for example, human, rat, mouse, other animals, specific plants, bacteria, algae, viruses, and the like. The input complex also can be from a mixture of genomes of different species such as host-pathogen, bacterial populations and the like. The input DNA can be cDNA made from a mixture of genomes of different species. Alternatively, the input nucleic acid can be from a synthetic source. The input DNA can be mitochondrial DNA. The input DNA can be cell-free DNA. The cell-free DNA can be obtained from, e.g., a serum or plasma sample. The input DNA can comprise one or more chromosomes. For example, if the input DNA is from a human, the DNA can comprise one or more of chromosome 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, X, or Y. The DNA can be from a linear or circular genome. The DNA can be plasmid DNA, cosmid DNA, bacterial artificial chromosome (BAC), or yeast artificial chromosome (YAC). The input DNA can be from more than one individual or organism. The input DNA can be double stranded or single stranded. The input DNA can be part of chromatin. The input DNA can be associated with histones.

[0070] In some embodiments, the oligonucleotides targeting the selected sequence regions of interest are designed to hybridize to single-stranded nucleic acid targets. In one embodiment, the oligonucleotides targeting

the selected sequence regions of interest are designed to hybridize to single-stranded DNA targets. In the case where the input nucleic acid sample comprises genomic DNA or other double-stranded DNA, the input nucleic acid sample can be first denatured to render the target single stranded and enable hybridization of the oligonucleotides to the desired sequence regions of interest. In these embodiments, the methods and compositions described herein can allow for region-specific enrichment and amplification of sequence regions of interest. In some embodiments, the other double-stranded DNA can be double-stranded cDNA generated by first and second strand synthesis of one or more target RNAs.

[0071] In other embodiments, the oligonucleotides targeting the selected sequence regions of interest are designed to hybridize to double-stranded nucleic acid targets, without denaturation of the double stranded nucleic acids. In other embodiments, the oligonucleotides targeting the selected sequence regions of interest are designed to hybridize to a double-stranded DNA target, without denaturation of the dsDNA. In these embodiments, the oligonucleotides targeting the selected sequence regions of interest are designed to form a triple helix (triplex) at the selected sequence regions of interest. The hybridization of the oligonucleotides to the double-stranded DNA sequence regions of interest can be carried out without prior denaturation of the double stranded nucleic acid sample. In such embodiments, the methods and compositions described herein can allow for region-specific enrichment as well as strand-specific enrichment and amplification of sequence regions of interest. This method can be useful for generation of copies of strand specific sequence regions of interest from complex nucleic acid without the need to denature the dsDNA input DNA, thus enabling enrichment and analysis of multiplicity of sequence regions of interest in the native complex nucleic acid sample. The method can find use for studies and analyses carried out in situ, enable studies and analysis of complex genomic DNA in single cells or collection of very small well defined cell population, as well as permit the analysis of complex genomic DNA without disruption of chromatin structures.

[0072] A "target nucleic acid sequence" or "target sequence" as used herein, is a polynucleotide sequence of interest, for which enrichment is desired. The target sequence may be known or not known, in terms of its actual sequence. Generally, a "template", as used herein, is a polynucleotide that contains the target nucleic acid sequence. The terms "target sequence," "target nucleic acid sequence," "target nucleotide sequence," "regions of interest," or "sequence of interest" and, variations thereof, are used interchangeably.

Oligonucleotides of the invention

[0073] As used within the invention, the term "oligonucleotide" refers to a polynucleotide chain, typically less than 200 residues long, most typically between 15 and 100 nucleotides long, but also intended to encompass longer polynucleotide chains. Oligonucleotides may be single-or double-stranded. As used in this invention, the term "oligonucleotide" may be used interchangeably with the terms "primer" and "adaptor".

[0074] As used herein, the terms "hybridization"/ "hybridizing" and "annealing" are used interchangeably and refer to the pairing of complementary nucleic acids.

[0075] The term "primer", as used herein, can refer to a nucleotide sequence, generally with a free 3' hydroxyl group, that is capable of hybridizing with a template (such as one or more target polynucleotides, one or more target DNAs, one or more target RNAs or a primer extension product) and is also capable of promoting polymerization of a polynucleotide complementary to the template. A primer can be, for example, an oligonucleotide. It can also be, for example, a sequence of the template (such as a primer extension product or a fragment of the template created following RNase (e.g., RNase H) cleavage of a template-DNA complex) that is hybridized to a sequence in the template itself (for example, as a hairpin loop), and that is capable of promoting nucleotide polymerization. Thus, a primer can be an exogenous (e.g., added) primer or an endogenous (e.g., template fragment) primer. A primer may contain a non-hybridizing sequence that constitutes a tail of the primer. A primer may still be hybridizing to a target even though its sequences are not fully complementary to the target.

[0076] The primers of the invention are generally oligonucleotides that are employed in an extension reaction by a polymerase along a polynucleotide template, such as in PCR, SPIA or cDNA synthesis, for example. The oligonucleotide primer can be a synthetic polynucleotide that is single stranded, containing a sequence at its 3′-end that is capable of hybridizing with a sequence of the target polynucleotide. Normally, the 3′ region of the primer that hybridizes with the target nucleic acid has at least 80%, preferably 90%, more preferably 95%, most preferably 100%, complementarity to a sequence or primer binding site.

[0077] "Complementary", as used herein, can refer to complementarity to all or only to a portion of a sequence. The number of nucleotides in the hybridizable sequence of a specific oligonucleotide primer should be such that stringency conditions used to hybridize the oligonucleotide primer will prevent excessive random non-specific hybridization. Usually, the number of nucleotides in the hybridizing portion of the oligonucleotide primer will be at least as great as the defined sequence on the target polynucleotide that the oligonucleotide primer hybridizes to, namely, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least about 20, and generally from about 6 to about 10 or 6 to about 12 or 12 to about 200 nucleotides, usually about 20 to about 50 nucleotides. In general, the target polynucleotide is larger than the oligonucleotide primer or primers as described previously.

[0078] In some cases, the identity of the investigated target polynucleotide sequence is known, and hybridizable sequence specific oligonucleotides or primers can be synthesized precisely according to the antisense sequence of the aforesaid target polynucleotide sequence. In some embodiments, multiple sequence-specific oligonucleotides or primers are employed to hybridize to a multiplicity of genomic regions of interest, allowing for selective enrichment of the regions of interest. In so far as the genomic regions may be very long, multiple oligonucleotides can be designed to hybridize to different sequence regions within the genomic regions of interest. In other embodiments, when the target polynucleotide sequence is unknown, the hybridizable sequence of an oligonucleotide or primer is a random sequence. Oligonucleotides or primers comprising random sequences may be referred to as "random primers", or "random oligonucleotides," as described herein. In one embodiment, an oligonucleotide or primer of the present invention hybridizable to a target sequence may comprise a mixture of primers or oligonucleotides designed to hybridize to a plurality (e.g. 2, 3, 4, about 6, 8, 10, 20, 40, 80, 100, 125,

150, 200, 250, 300, 400, 500, 600, 800, 1000, 1500, 2000, 2500, 3000, 4000, 5000, 6000, 7000, 8000, 10,000, 20,000, 25,000 or more) of target sequences. In some cases, the plurality of target sequences may comprise a group of related sequences, random sequences, a whole transcriptome or fraction (e.g. substantial fraction) thereof, or any group of sequences such as mRNA. In some embodiments, the primers can be directed to known sequences present in the adaptors used in the invention as described herein. In this embodiment, the primers can comprise groups of primers comprising one or more primers in each group, wherein each group of primers can be directed against distinct adaptors.

[0079] Tailed primers or oligonucleotides can be employed in certain embodiments of the invention. In general, a tailed primer comprises a 3' portion that is hybridizable to one or more target polynucleotides, and a 5' portion that is not hybridizable to the one or more target polynucleotides. In general, the non-hybridizable 5' portion does not hybridize to the one or more target polynucleotides under conditions in which the hybridizable 3' portion of the tailed primer hybridizes to the one or more target polynucleotides. In some embodiments, the non-hybridizable 5' portion comprises an adaptor sequence. In some embodiments, the non-hybridizable 5' portion comprises a common or conventional adaptor sequence. In some embodiments, the non-hybridizable 5' portion comprises a common or conventional adaptor sequence that is distinct or different from the sequence of other adaptors used in the present invention. In some embodiments, the non-hybridizable 5' portion comprises a promoter-specific sequence. Generally, a promoter-specific sequence comprises a single-stranded DNA sequence region which, in double-stranded form is capable of mediating RNA transcription. Examples of promoter-specific sequences are known in the art, and include, without limitation, T7, T3, or SP6 RNA polymerase promoter sequences. When the tailed primer is extended with a DNA polymerase, a primer extension product with a 5' portion comprising a defined sequence can be created. This primer extension product can then have a second primer anneal to it, which can be extended with a DNA polymerase to create a double stranded product comprising a defined sequence at one end. In some embodiments, where the non-hybridizable 5' portion of one or more tailed primers comprises a promoter-specific sequence, creation of a double-stranded product comprising a defined sequence at one end generates a double-stranded promoter sequence that is capable of mediating RNA transcription. In some embodiments, a double-stranded promoter sequence can be generated by hybridizing to the promoter-specific sequence an oligonucleotide comprising a sequence complementary to the promoter-specific sequence. In some embodiments, formation of a double-stranded promoter can be followed by the generation of single-stranded RNA by RNA transcription of sequence downstream of the double-stranded promoter, generally in a reaction mixture comprising all necessary components, including but not limited to ribonucleoside triphosphates (rNTPs) and a DNA-dependent RNA polymerase. Tailed primers can comprise DNA, RNA, or both DNA and RNA. In some embodiments, the tailed primer consists of DNA.

[0080] Composite primers can be employed in certain embodiments of the invention. Composite primers are primers that are composed of RNA and DNA portions. In some aspects, the composite primer can be a tailed composite primer comprising, for example, a 3'-DNA portion and a 5'-RNA portion. In the tailed composite primer, a 3'-portion, all or a portion of which comprises DNA, is complementary to a polynucleotide; and a 5'-

portion, all or a portion of which comprises RNA, is not complementary to the polynucleotide and does not hybridize to the polynucleotide under conditions in which the 3'-portion of the tailed composite primer hybridizes to the polynucleotide target. When the tailed composite primer is extended with a DNA polymerase, a primer extension product with a 5'-RNA portion comprising a defined sequence can be created. This primer extension product can then have a second primer anneal to it, which can be extended with a DNA polymerase to create a double stranded product with an RNA/DNA heteroduplex comprising a defined sequence at one end. The RNA portion can be selectively cleaved from the partial heteroduplex to create a double-stranded DNA with a 3'-single-stranded overhang which can be useful for various aspects of the present invention including allowing for isothermal amplification using a composite amplification primer.

[0081] A "random primer," as used herein, can be a primer that generally comprises a sequence that is designed not necessarily based on a particular or specific sequence in a sample, but rather is based on a statistical expectation (or an empirical observation) that the sequence of the random primer is hybridizable (under a given set of conditions) to one or more sequences in the sample. A random primer will generally be an oligonucleotide or a population of oligonucleotides comprising a random sequence(s) in which the nucleotides at a given position on the oligonucleotide can be any of the four nucleotides, or any of a selected group of the four nucleotides (for example only three of the four nucleotides, or only two of the four nucleotides). In some cases all of the positions of the oligonucleotide or population of oligonucleotides can be any of two or more nucleotides. In other cases, only a portion of the oligonucleotide, for instance a particular region, will comprise positions which can be any of two or more bases. In some cases, the portion of the oligonucleotide which comprises positions which can be any of two or more bases is about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or about 15-20 nucleotides in length. In some cases, a random primer may comprise a tailed primer having a 3'-region that comprises a random sequence and a 5'-region that is a non-hybridizing sequence that comprises a specific, non-random sequence. The 3'-region may also comprise a random sequence in combination with a region that comprises poly-T sequences. The sequence of a random primer (or its complement) may or may not be naturally-occurring, or may or may not be present in a pool of sequences in a sample of interest. As is well understood in the art, a "random primer" can also refer to a primer that is a member of a population of primers (a plurality of random primers) which collectively are designed to hybridize to a desired and/or a significant number of target sequences. A random primer may hybridize at a plurality of sites on a nucleic acid sequence. The use of random primers provides a method for generating primer extension products complementary to a target polynucleotide or target nucleic sequence which does not require prior knowledge of the exact sequence of the target. In some embodiments one portion of a primer is random, and another portion of the primer comprises a defined sequence. For example, in some embodiments, a 3'-portion of the primer will comprise a random sequence, while the 5'-portion of the primer comprises a defined sequence. In some embodiments a 3'-random portion of the primer will comprise DNA, and a 5'- defined portion of the primer will comprise RNA, in other embodiments, both the 3' and 5'-portions will comprise DNA. In some embodiments, the 5'-portion will contain a defined sequence and the 3'-portion will comprise a poly-dT sequence that is hybridizable to a multiplicity of RNAs in a sample (such as all mRNA). In

some embodiments, a “random primer,” or primer comprising a randomly generated sequence, comprises a collection of primers comprising one or more nucleotides selected at random from two or more different nucleotides, such that all possible sequence combinations of the nucleotides selected at random may be represented in the collection. In some embodiments, generation of one or more random primers does not include a step of excluding or selecting certain sequences or nucleotide combinations from the possible sequence combinations in the random portion of the one or more random primers.

[0082] In one embodiment, the oligonucleotides of the invention can be tailed oligonucleotides. In one embodiment, the 5’-tail can comprise RNA and is non hybridizable to the RNA in the sample. In one embodiment, the 5’-tail can comprise DNA and is non hybridizable to the DNA in the sample. In one embodiment, the 5’-tail can comprise an adaptor that is not hydridizable to the DNA and/or nucleic acid fragments derived from the sample comprising nucleic acid. In one embodiment, the 5’-tail can comprise an adaptor sequence that is not hydridizable to the DNA and/or nucleic acid fragments derived from the sample comprising nucleic acid. In some embodiments, the 5’-tail can comprise a common adaptor sequence that is not hydridizable to the DNA and is distinct from any other adaptor or adaptor sequence used in the methods of the invention described herein. In some embodiments, the 5’-tail can comprise an identifier sequence. In some embodiments, the identifier sequence can comprise a barcode sequence. In some embodiments, the 5’-tail can comprise a common adaptor sequence that is not hydridizable to the DNA and a barcode sequence.

[0083] The term “adaptor”, as used herein, refers to an oligonucleotide of known sequence, the ligation of which to a target polynucleotide or a target polynucleotide strand of interest enables the generation of amplification-ready products of the target polynucleotide or the target polynucleotide strand of interest. The target polynucleotide molecules may be fragmented or not prior to the addition of adaptors.

[0084] Various adaptor designs are envisioned which are suitable for generation of amplification-ready products of target sequence regions/strands of interest. For example, the two strands of the adaptor may be self-complementary, non-complementary or partially complementary. A common feature of the adaptors depicted in Figure 3 of the invention is the partial duplex design, wherein the two strands of the adaptor are different lengths with a complementary region and an overhanging region at the 5’ end. The 5’ end of the long strand of the partial duplex adaptor contains a unique site for a nucleic acid modifying enzyme, such as a restriction enzyme, that is absent from the short strand of the duplex adaptor. The 3’ end of the short strand adaptor is modified by a replacement of the 3’ OH-group is by a blocking group, for example, a dideoxynucleotide (ddCMP, ddAMP, ddTMP, or ddGMP) to prevent polymerase extension.

[0085] In some embodiments of the invention, the adaptors comprise an additional identifier sequence, e.g., a barcode sequence. As used herein, the term "barcode" refers to a known nucleic acid sequence that allows some feature of a polynucleotide with which the barcode is associated to be identified. In some embodiments, the feature of the polynucleotide to be identified is the sample from which the polynucleotide is derived. In some embodiments, barcodes are at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or more nucleotides in length. In some embodiments, barcodes are shorter than 10, 9, 8, 7, 6, 5, or 4 nucleotides in length. In some embodiments, each

barcode in a plurality of barcodes differ from every other barcode in the plurality at at least three nucleotide positions, such as at least 3, 4, 5, 6, 7, 8, 9, 10, or more positions. In some embodiments, barcodes associated with some polynucleotides are of different length than barcodes associated with other polynucleotides. In general, barcodes are of sufficient length and comprise sequences that are sufficiently different to allow the identification of samples based on barcodes with which they are associated. In some embodiments, both the forward and reverse adapter comprise at least one of a plurality of barcode sequences. In some embodiments, the first, second, and/or third adaptor comprises at least one of a plurality of barcode sequences. In some embodiments, each reverse adapter comprises at least one of a plurality of barcode sequences, wherein each barcode sequence of the plurality of barcode sequences differs from every other barcode sequence in the plurality of barcode sequences. In some embodiments, both the first adapter and the second adapter comprise at least one of a plurality of barcode sequences. In some embodiments, barcodes for second adapter oligonucleotides are selected independently from barcodes for first adapter oligonucleotides. In some embodiments, first adapter oligonucleotides and second adapter oligonucleotides having barcodes are paired, such that adapters of the pair comprise the same or different one or more barcodes. In some embodiments, the methods of the invention further comprise identifying the sample from which a target polynucleotide is derived based on the barcode sequence to which the target polynucleotide is joined. In general, a barcode comprises a nucleic acid sequence that when joined to a target polynucleotide serves as an identifier of the sample from which the target polynucleotide was derived.

[0086] Recently, many improvements have been made in adaptor design that have reduced the occurrence of adapter dimer. These improvements include the use of nucleotide analogs and structured oligonucleotides, and have allowed for use of higher concentrations of oligonucleotides in ligation reactions. The higher concentrations of adapters in ligation reactions have enabled researchers to produce high quality libraries from as few as 150 copies of genome. Ligation of adaptors to the ends of DNA fragments, in particular those fragments containing the regions of interest is suitable for carrying out the methods of the invention. Various ligation modalities are envisioned, dependent on the choice of nucleic acid modifying enzymes and the resulting double-stranded DNA cleavage. For example, when a blunt end product comprising the target region/sequence of interest is generated, blunt end ligation can be suitable. Alternatively, where the cleavage is carried out using a restriction enzyme of known sequence specificity, leading to the generation of cleavage sites with known sequence overhangs, suitable ends of the adaptors can be designed to enable hybridization of the adaptor to the cleavage site of the sequence region of interest and subsequent ligation. Reagents and methods for efficient and rapid ligation of adaptors are commercially available and are known in the art.

Nucleic acid modifying enzymes

[0087] The nucleic acid (NA) -modifying enzyme can be DNA-specific modifying enzyme. The NA-modifying enzyme can be selected for specificity for double-stranded DNA. The enzyme can be a duplex-specific endonuclease, a blunt-end frequent cutter restriction enzyme, or other restriction enzyme. Examples of blunt-end cutters include DraI or SmaI. The NA-modifying enzyme can be an enzyme provided by New England Biolabs .

The NA-modifying enzyme can be a homing endonuclease (a homing endonuclease can be an endonuclease that does not have a stringently-defined recognition sequence). The NA-modifying enzyme can be a nicking endonuclease (a nicking endonuclease can be an endonuclease that can cleave only one strand of DNA in a double-stranded DNA substrate). The NA-modifying enzyme can be a high fidelity endonuclease (a high fidelity endonuclease can be an engineered endonuclease that has less "star activity" than the wild-type version of the endonuclease).

[0088] In a preferred embodiment, the NA-modifying enzyme is a sequence and duplex-specific, DNA modifying enzyme.

DNA-Dependent DNA Polymerases

[0089] DNA-dependent DNA polymerases for use in the methods and compositions of the invention are capable of effecting extension of a primer or oligonucleotide according to the methods of the invention. In one embodiment, a preferred DNA-dependent DNA polymerase can be one that is capable of extending a nucleic acid primer in the presence of the DNA and/or cDNA template. Exemplary DNA dependent DNA polymerases suitable for the methods of the present invention include but are not limited to Klenow polymerase, with or without 3'-exonuclease, Bst DNA polymerase, Bca polymerase, .phi.29 DNA polymerase, Vent polymerase, Deep Vent polymerase, Taq polymerase, T4 polymerase, and E. coli DNA polymerase 1, derivatives thereof, or mixture of polymerases. In some cases, the polymerase does not comprise a 5'-exonuclease activity. In other cases, the polymerase comprises 5' exonuclease activity. In some cases, the primer or oligonucleotide extension of the present invention may be performed using a polymerase comprising strong strand displacement activity such as for example Bst polymerase. In other cases, the primer extension of the present invention may be performed using a polymerase comprising weak or no strand displacement activity. One skilled in the art may recognize the advantages and disadvantages of the use of strand displacement activity during the primer extension step, and which polymerases may be expected to provide strand displacement activity (see e.g., New England Biolabs Polymerases).

Methods of Amplification

[0090] The methods, compositions and kits described herein can be useful to generate amplification-ready products for downstream applications such as massively parallel sequencing (e.g., next generation sequencing methods), generation of libraries with enriched population of sequence regions of interest, or hybridization platforms. Methods of amplification are well known in the art. Suitable amplification reactions can be exponential or isothermal and can include any DNA amplification reaction, including but not limited to polymerase chain reaction (PCR), strand displacement amplification (SDA), linear amplification, multiple displacement amplification (MDA), rolling circle amplification (RCA), single primer isothermal amplification (SPIA, see e.g. U.S. Pat. No. 6,251,639), Ribo-SPIA, or a combination thereof. In some cases, the amplification methods for providing the template nucleic acid may be performed under limiting conditions such that only a few

rounds of amplification (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 etc.), such as for example as is commonly done for cDNA generation. The number of rounds of amplification can be about 1-30, 1-20, 1-15, 1-10, 5-30, 10-30, 15-30, 20-30, 10-30, 15-30, 20-30, or 25-30.

[0091] PCR is an in vitro amplification procedure based on repeated cycles of denaturation, oligonucleotide primer annealing, and primer extension by thermophilic template dependent polynucleotide polymerase, resulting in the exponential increase in copies of the desired sequence of the polynucleotide analyte flanked by the primers. The two different PCR primers, which anneal to opposite strands of the DNA, are positioned so that the polymerase catalyzed extension product of one primer can serve as a template strand for the other, leading to the accumulation of a discrete double stranded fragment whose length is defined by the distance between the 5' ends of the oligonucleotide primers.

[0092] LCR uses a ligase enzyme to join pairs of preformed nucleic acid probes. The probes hybridize with each complementary strand of the nucleic acid analyte, if present, and ligase is employed to bind each pair of probes together resulting in two templates that can serve in the next cycle to reiterate the particular nucleic acid sequence.

[0093] SDA (Westin et al 2000, Nature Biotechnology, 18, 199-202; Walker et al 1992, Nucleic Acids Research, 20, 7, 1691-1696), is an isothermal amplification technique based upon the ability of a restriction endonuclease such as HincII or BsoBI to nick the unmodified strand of a hemiphosphorothioate form of its recognition site, and the ability of an exonuclease deficient DNA polymerase such as Klenow exo minus polymerase, or Bst polymerase, to extend the 3'-end at the nick and displace the downstream DNA strand. Exponential amplification results from coupling sense and antisense reactions in which strands displaced from a sense reaction serve as targets for an antisense reaction and vice versa.

[0094] Some aspects of the invention utilize linear amplification of nucleic acids or polynucleotides. Linear amplification generally refers to a method that involves the formation of one or more copies of the complement of only one strand of a nucleic acid or polynucleotide molecule, usually a nucleic acid or polynucleotide analyte. Thus, the primary difference between linear amplification and exponential amplification is that in the latter process, the product serves as substrate for the formation of more product, whereas in the former process the starting sequence is the substrate for the formation of product but the product of the reaction, i.e., the replication of the starting template, is not a substrate for generation of products. In linear amplification the amount of product formed increases as a linear function of time as opposed to exponential amplification where the amount of product formed is an exponential function of time.

[0095] In some embodiments, the amplification is exponential, e.g. in the enzymatic amplification of specific double stranded sequences of DNA by a polymerase chain reaction (PCR). In other embodiments the amplification method is linear. In other embodiments the amplification method is isothermal.

Downstream Applications

[0096] An important aspect of the invention is that the methods and compositions disclosed herein can be efficiently and cost-effectively utilized for downstream analyses, such as next generation sequencing or

hybridization platforms, with minimal loss of biological material of interest. The methods of the present invention can also be used in the analysis of genetic information of selective genomic regions of interest (e.g., analysis of SNPs or other disease markers) as well as genomic regions which may interact with the selective region of interest.

Sequencing

[0097] For example, the methods of the invention are useful for sequencing by the method commercialized by Illumina, as described U.S. Pat. Nos. 5,750,341; 6,306,597; and 5,969,119. In general, double stranded fragment polynucleotides can be prepared by the methods of the present invention to produce amplified nucleic acid sequences tagged at one (e.g., (A)/(A′) or both ends (e.g., (A)/(A′) and (C)/(C′)). In some cases, single stranded nucleic acid tagged at one or both ends is amplified by the methods of the present invention (e.g., by SPIA or linear PCR).The resulting nucleic acid is then denatured and the single-stranded amplified polynucleotides are randomly attached to the inside surface of flow-cell channels. Unlabeled nucleotides are added to initiate solid-phase bridge amplification to produce dense clusters of double-stranded DNA. To initiate the first base sequencing cycle, four labeled reversible terminators, primers, and DNA polymerase are added. After laser excitation, fluorescence from each cluster on the flow cell is imaged. The identity of the first base for each cluster is then recorded. Cycles of sequencing are performed to determine the fragment sequence one base at a time.

[0098] In some embodiments, the methods of the invention are useful for preparing target polynucleotides for sequencing by the sequencing by ligation methods commercialized by Applied Biosystems (e.g., SOLiD sequencing). In other embodiments, the methods are useful for preparing target polynucleotides for sequencing by synthesis using the methods commercialized by 454/Roche Life Sciences, including but not limited to the methods and apparatus described in Margulies et al., *Nature*(2005) 437:376-380 (2005); and U.S. Pat. Nos. 7,244,559; 7,335,762; 7,211,390; 7,244,567; 7,264,929; and 7,323,305. In other embodiments, the methods are useful for preparing target polynucleotide(s) for sequencing by the methods commercialized by Helicos BioSciences Corporation (Cambridge, Mass.) as described in U.S. application Ser. No. 11/167,046, and U.S. Pat. Nos. 7,501,245; 7,491,498; 7,276,720; and in U.S. Patent Application Publication Nos. US20090061439; US20080087826; US20060286566; US20060024711; US20060024678; US20080213770; and US20080103058. In other embodiments, the methods are useful for preparing target polynucleotide(s) for sequencing by the methods commercialized by Pacific Biosciences as described in U.S. Pat. Nos. 7,462,452; 7,476,504; 7,405,281; 7,170,050; 7,462,468; 7,476,503; 7,315,019; 7,302,146; 7,313,308; and US Application Publication Nos. US20090029385; US20090068655; US20090024331; and US20080206764.

[0099] Another example of a sequencing technique that can be used in the methods of the provided invention is nanopore sequencing (see e.g. Soni G V and Meller A. (2007) *Clin Chem* 53: 1996-2001). A nanopore can be a small hole of the order of 1 nanometer in diameter. Immersion of a nanopore in a conducting fluid and application of a potential across it can result in a slight electrical current due to conduction of ions through the nanopore. The amount of current that flows is sensitive to the size of the nanopore. As a DNA molecule passes through a

nanopore, each nucleotide on the DNA molecule obstructs the nanopore to a different degree. Thus, the change in the current passing through the nanopore as the DNA molecule passes through the nanopore can represent a reading of the DNA sequence.

[00100] Another example of a sequencing technique that can be used in the methods of the provided invention is semiconductor sequencing provided by Ion Torrent (e.g., using the Ion Personal Genome Machine (PGM)). Ion Torrent technology can use a semiconductor chip with multiple layers, e.g., a layer with micro-machined wells, an ion-sensitive layer, and an ion sensor layer. Nucleic acids can be introduced into the wells, e.g., a clonal population of single nucleic can be attached to a single bead, and the bead can be introduced into a well. To initiate sequencing of the nucleic acids on the beads, one type of deoxyribonucleotide (e.g., dATP, dCTP, dGTP, or dTTP) can be introduced into the wells. When one or more nucleotides are incorporated by DNA polymerase, protons (hydrogen ions) are released in the well, which can be detected by the ion sensor. The semiconductor chip can then be washed and the process can be repeated with a different deoxyribonucleotide. A plurality of nucleic acids can be sequenced in the wells of a semiconductor chip. The semiconductor chip can comprise chemical-sensitive field effect transistor (chemFET) arrays to sequence DNA (for example, as described in U.S. Patent Application Publication No. 20090026082). Incorporation of one or more triphosphates into a new nucleic acid strand at the 3′ end of the sequencing primer can be detected by a change in current by a chemFET. An array can have multiple chemFET sensors.

Genetic Analysis

[00101] The methods of the present invention can be used in the analysis of genetic information of selective genomic regions of interest as well as genomic regions which may interact with the selective region of interest. Amplification methods as disclosed herein can be used in the devices, kits, and methods known to the art for genetic analysis, such as, but not limited to those found in U.S. Pat. Nos. 6,449,562, 6,287,766, 7,361,468, 7,414,117, 6,225,109, and 6,110,709. In some cases, amplification methods of the present invention can be used to amplify target nucleic acid of interest for DNA hybridization studies to determine the presence or absence of polymorphisms. The polymorphisms, or alleles, can be associated with diseases or conditions such as genetic disease. In other cases the polymorphisms can be associated with susceptibility to diseases or conditions, for example, polymorphisms associated with addiction, degenerative and age related conditions, cancer, and the like. In other cases, the polymorphisms can be associated with beneficial traits such as increased coronary health, or resistance to diseases such as HIV or malaria, or resistance to degenerative diseases such as osteoporosis, Alzheimer's or dementia.

Kits

[00102] Any of the compositions described herein may be comprised in a kit. In a non-limiting example, the kit, in a suitable container, comprises: an adaptor or several adaptors, one or more of oligonucleotide primers and

reagents for ligation, primer extension and amplification. The kit may also comprise means for purification, such as a bead suspension.

[00103] The containers of the kits will generally include at least one vial, test tube, flask, bottle, syringe or other containers, into which a component may be placed, and preferably, suitably aliquotted. Where there is more than one component in the kit, the kit also will generally contain a second, third or other additional container into which the additional components may be separately placed. However, various combinations of components may be comprised in a container.

[00104] When the components of the kit are provided in one or more liquid solutions, the liquid solution can be an aqueous solution. However, the components of the kit may be provided as dried powder(s). When reagents and/or components are provided as a dry powder, the powder can be reconstituted by the addition of a suitable solvent.

[00105] A kit will preferably include instructions for employing, the kit components as well the use of any other reagent not included in the kit. Instructions may include variations that can be implemented.

[00106] In one aspect, the invention provides kits containing any one or more of the elements disclosed in the above methods and compositions. In some embodiments, a kit comprises a composition of the invention, in one or more containers. In some embodiments, the invention provides kits comprising adapters, primers, and/or other oligonucleotides described herein. In some embodiments, the kit further comprises one or more of: (a) a DNA ligase, (b) a DNA-dependent DNA polymerase, (c) an RNA-dependent DNA polymerase, (d) a forward adapter (e) one or more oligonucleotides comprising reverse adaptor sequence and (f) one or more buffers suitable for one or more of the elements contained in said kit. The adapters, primers, other oligonucleotides, and reagents can be, without limitation, any of those described above. Elements of the kit can further be provided, without limitation, in any of the amounts and/or combinations (such as in the same kit or same container) described above. The kits may further comprise additional agents, such as those described above, for use according to the methods of the invention. For example, the kit can comprise a first forward adaptor that is a partial duplex adaptor as described herein, a second forward adapter, and a nucleic acid modifying enzyme specific for a restriction and/or cleavage site present in the first forward adaptor. The kit elements can be provided in any suitable container, including but not limited to test tubes, vials, flasks, bottles, ampules, syringes, or the like. The agents can be provided in a form that may be directly used in the methods of the invention, or in a form that requires preparation prior to use, such as in the reconstitution of lyophilized agents. Agents may be provided in aliquots for single-use or as stocks from which multiple uses, such as in a number of reaction, may be obtained.

[00107] In one embodiment, the kit comprises a plurality of forward adaptor oligonucleotides, wherein each of said forward adaptor oligonucleotides comprises at least one of a plurality of barcode sequences, wherein each barcode sequence of the plurality of barcode sequences differs from every other barcode sequence in said plurality of barcode sequences at at least three nucleotide positions, and instructions for using the same. Forward adapters comprising different barcode sequences can be supplied individually or in combination with one or more additional forward adapters having a different barcode sequence. In some embodiments, the kit can comprises a plurality of first and second forward adapter oligonucleotides. Second forward adapter oligonucleotides can be

supplied separately from or in combination with one or more first forward adapters, and/or one or more different second adapters. Combinations of first and second forward adapters can be supplied in accordance with combinations described above. In some embodiments, the kit can comprises a plurality of oligonucleotides comprising reverse adaptor sequence. In one embodiment, the kit can comprises a plurality of oligonucleotides comprising reverse adaptor sequence, wherein each of the plurality of oligonucleotides comprising reverse adaptor sequence further comprises sequence complementary to a specific target sequence of interest present in a nucleic acid. In one embodiment, the kit can comprises a plurality of oligonucleotides comprising reverse adaptor sequence, wherein each of the plurality of oligonucleotides comprising reverse adaptor sequence further comprises random sequence. In one embodiment, the kit comprises a plurality of oligonucleotides with reverse adaptor sequence, wherein each of said oligonucleotides with reverse adaptor sequence comprises at least one of a plurality of barcode sequences, wherein each barcode sequence of the plurality of barcode sequences differs from every other barcode sequence in said plurality of barcode sequences at at least three nucleotide positions, and instructions for using the same. Oligonucleotides with reverse adaptor sequence comprising different barcode sequences can be supplied individually or in combination with one or more additional oligonucleotides with reverse adaptor sequence having a different barcode sequence.

[00108] In some cases, methods, compositions, and/or kits described herein can be used to detect a fusion event (e.g., any fusion event). In some cases, the fusions can be a fusion found in cancer. In some cases, the cancer fusion is a fusion as curated by the Wellcome Trust Sanger Institute's Catalog of Somatic Mutations in Cancer (COSMIC). (//www.sanger.ac.uk/resources/databases/cosmic.html). In some cases, the number of unique fusions that can be detected in a single assay is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 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, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, or 500. In some cases, the number of unique cancer fusions that can be detected in a single assay is about 1 to about 10, about 5 to about 25, about 20 to about 50, about 25 to about 50, about 50 to about 100, about 100 to about 200, about 100 to about 300, about 100 to about 400, about 100 to about 450, about 100 to about 500, about 200 to about 300, about 200 to about 400, about 200 to about 500, about 300 to about 400, or about 300 to about 500. In some cases, the number of unique cancer fusions that be detected in a single assay is about 10, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475 or 500. In some cases, fusions are detected using genomic nucleic acid (e.g., genomic DNA) as a starting material. In some cases, fusions are detected using RNA (e.g., messenger RNA) as a starting material. In some cases, the methods and compositions described herein can be used to enrich for and detect gene fusion events in a nucleic acid sample. In these examples, a target specific primer or oligonucleotide may anneal to a portion of at least one gene in the gene fusion. In other examples, a target specific primer or oligonucleotide may anneal to a portion of both genes in the gene fusion (i.e., at the boundary of the gene fusion event).

[00109] In some cases, a target specific primer anneals to an exon of a gene. In some cases, a target specific primer anneals to an intron of a gene. In some cases, each of a plurality of target specific primers anneals an exon in a unique gene. In some cases, a primer anneals to a sequence spanning an exon/intron junction. In some cases, a single assay can employ a plurality of primers, wherein the primers anneal to at least 10, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10,000, 20,000, 30,000, 40,000, or 50,000 unique exons. In some cases, a single assay employs a plurality of primers that anneal to about 5769 exons. In some cases, a single assay employs a plurality of primers that anneal to about 100 to about 500, about 500 to about 1000, about 1000 to about 2000, about 2000 to about 3000, about 3000 to about 4000, about 4000 to about 10,000, about 5000 to about 10,000, about 4000 to about 6000, or about 5000 to about 7500 unique exons.

[00110] A kit can comprise about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 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, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 600, 700, 800, 900, or 1000 different barcodes. A kit can comprise at least 8, 32, or 96 barcodes. A kit can comprise 8, 32, or 96 barcodes.

[00111] In some cases, the methods, compositions, and/or kits described herein can be used to discover novel gene fusions. In some cases, the methods, compositions, and/or kits described herein can be used to detect alternative splice forms. The methods, compositions, and/or kits described herein can be used to detect a low abundance of gene fusion transcripts. The amount of starting material can be, e.g., about, or at least 0.1 ng, 0.5 ng, 1 ng, 5 ng, 10 ng, 20 ng, 30 ng, 40 ng, 50 ng, 75 ng, 100 ng, 200 ng, 300 ng, 400 ng, 500 ng, 750 ng, 1000 ng, 5000 ng, or 10,000 ng of nucleic acid, e.g., RNA. The amount of starting material can be, e.g., about 0.1 ng to about 1 ng, about 1 ng to about 10 ng, about 10 ng, about 10 ng to about 50 ng, about 10 ng to about 100 ng, about 10 ng to about 500 ng, about 10 ng to about 1000 ng, about 100 ng to about 500 ng, or about 100 ng to about 1000 ng. In some cases, the amount of starting material is about or less than 10 ng of total RNA. In some cases, the amount of starting material is about or greater than 10 ng of total RNA. In some cases, nucleic acid (e.g., RNA) is obtained from a fresh sample. In some cases, nucleic acid is obtained from a formalin-fixed tissue. In some cases, the sample is a formalin-fixed paraffin-embedded tissue.

[00112] In some cases, the methods, compositions, and/or kits described herein can be used for basic research. In some cases, the methods, compositions, and/or kits described herein can be used for in RNA-Seq based cancer diagnostic and/or prognostic tests.

[00113] In some cases, a nucleic acid sample comprises a fusion of genetic material (e.g., a translocation). In some cases a fusion can be in cancer. In some cases, the cancer is acute myeloid leukemia; bladder cancer, including upper tract tumors and urothelial carcinoma of the prostate; bone cancer, including chondrosarcoma, Ewing's sarcoma, and osteosarcoma; breast cancer, including noninvasive, invasive, phyllodes tumor, Paget's disease, and breast cancer during pregnancy; central nervous system cancers, adult low-grade infiltrative

supratentorial astrocytoma/oligodendroglioma, adult intracranial ependymoma, anaplastic astrocytoma/anaplastic oligodendroglioma/glioblastoma multiforme, limited (1-3) metastatic lesions, multiple (>3) metastatic lesions, carcinomatous lymphomatous meningitis, nonimmunosuppressed primary CNS lymphoma, and metastatic spine tumors; cervical cancer; chronic myelogenous leukemia (CML); colon cancer, rectal cancer, anal carcinoma; esophageal cancer; gastric (stomach) cancer; head and neck cancers, including ethmoid sinus tumors, maxillary sinus tumors, salivary gland tumors, cancer of the lip, cancer of the oral cavity, cancer of the oropharynx, cancer of the hypopharynx, occult primary, cancer of the glottic larynx, cancer of the supraglottic larynx, cancer of the nasopharynx, and advanced head and neck cancer; hepatobiliary cancers, including hepatocellular carcinoma, gallbladder cancer, intrahepatic cholangiocarcinoma, and extrahepatic cholangiocarcinoma; Hodgkin disease/lymphoma; kidney cancer; melanoma; multiple myeloma, systemic light chain amyloidosis, Waldenstrom's macroglobulinemia; myelodysplastic syndromes; neuroendocrine tumors, including multiple endocrine neoplasia, type 1, multiple endocrine neoplasia, type 2, carcinoid tumors, islet cell tumors, pheochromocytoma, poorly differentiated/small cell/atypical lung carcinoids; Non-Hodgkin's Lymphomas, including chronic lymphocytic leukemia/small lymphocytic lymphoma, follicular lymphoma, marginal zone lymphoma, mantle cell lymphoma, diffuse large B-Cell lymphoma, Burkitt's lymphoma, lymphoblastic lymphoma, AIDS-Related B-Cell lymphoma, peripheral T-Cell lymphoma, and mycosis fungoides/Sezary Syndrome; non-melanoma skin cancers, including basal and squamous cell skin cancers, dermatofibrosarcoma protuberans, Merkel cell carcinoma; non-small cell lung cancer (NSCLC), including thymic malignancies; occult primary; ovarian cancer, including epithelial ovarian cancer, borderline epithelial ovarian cancer (Low Malignant Potential), and less common ovarian histologies; pancreatic adenocarcinoma; prostate cancer; small cell lung cancer and lung neuroendocrine tumors; soft tissue sarcoma, including soft-tissue extremity, retroperitoneal, intra-abdominal sarcoma, and desmoid; testicular cancer; thymic malignancies, including thyroid carcinoma, nodule evaluation, papillary carcinoma, follicular carcinoma, Hiirthle cell neoplasm, medullary carcinoma, and anaplastic carcinoma; uterine neoplasms, including endometrial cancer or uterine sarcoma.

[00114] In some cases, the methods described herein can be used to assess the expression level of a gene. In some cases, the expression level can be, for example, higher than normal, normal, or below normal. In some cases, a nucleic acid analyzed using a method, compositions, or kit described herein lack a mutation (e.g., wild-type) or contains one or more mutations (e.g., de novo mutation, nonsense mutation, missense mutation, silent mutation, frameshift mutation, insertion, substitution, point mutation, single nucleotide polymorphism (SNP), deletion, rearrangement, amplification, chromosomal translocation, interstitial deletion, chromosomal inversion, loss of heterozygosity, loss of function, gain of function, dominant negative, or lethal). In some cases, methods, compositions or kits described herein can be used to determine copy number alteration or a copy number variation.

[00115] **Fusion-caused cancers**

[00116] In some cases, fusion genes can be found in cancer. Cancers in which a gene fusion can be found include acute lymphoblastic leukemia (ALL), acute nonlymphocytic leukemia (ANLL), acute myelogenous, acute

myelogenous leukemia (AML), congenital fibrosarcoma (CFS), congenital mesoblastic nephroma (CMN), secretory ductal carcinoma of breast, congenital mesoblastic nephroma (CMN), angiomatoid fibrous histiocytoma (AFH), myxoid liposarcoma (MLS), acute myelomonocytic, alveolar rhabdomyosarcoma (ARMS), acute promyelocytic leukemia (APL), anaplastic large cell lymphoma (ALCL), B-cell non-Hodgkin lymphoma (NHL), and marginal zone B-cell lymphoma (MZBCL) or mucosa-associated lymphoid tissue (MALT).

[00117] **Cancer gene fusions**

[00118] In some cases, a cancer fusion can be TEL-AML1 fusion, E2A-PBX (PBX1) fusion, BCR-ABL fusion, MLL-AF4 fusion, IGH-MYC fusion, and TCR-RBTN2 fusion in ALL. In other cases, a cancer fusion can be RET and NTRK1 fusion in papillary thyroid carcinoma, PAX8-PPARG in follicular thyroid carcinoma, MECT1-MAML2 in mucoepidermoid carcinoma, TFE3 and TFEB fusion in kidney carcinomas, BRD4-NUT in midline carcinomas, ETV6-NTRK3 in secretory breast carcinomas, TMPRSS2-ETS fusion in prostate carcinomas, and TMPRSS2-ERG in prostate cancer. In certain embodiments, a cancer fusion can be a MLL-AF4 fusion in ALL, MLL-MLLT3 fusion in ANLL, MLL-MLLT1 fusion in ALL and ANLL, MLL-CREBBP fusion in AML, ABL-BCR fusion in CML, ALL, and ANLL, ETV6-NTRK3 fusion in CFS, and CMN, ETV6 – RUNX1 fusion in B cell ALL, FKHRL1-PAX3 fusion in ARMS, RARA-PML fusion in APL, RARA-NPM1 fusion in ALCL, or DDX10-NUP98 fusion in MDS and ANLL. In some cases, the cancer fusion is a RET fusion. In other cases, the cancer fusion is a ROS1 fusion. In other cases, the cancer fusion is an ALK fusion.

[00119] Non-limiting examples of genes that can be involved in cancer fusions include: ABI1, ABL1, ABL2, ACBD6, ACCN1, ACLY, ACSL3, ACTB, AF15Q14, AF1Q, AF3p21, AFF4, AGPAT5, AGTRAP, AHRR, AKAP9, ALDH2, ALK, ALO17, AMD1, ARFGEF2, ARFIP1, ARHGAP6, ARHGEF12, ARHH, ARID1A, ARNT, ASPSCR1, ASXL1, ATF1, ATG4C, ATIC, ATP8B2, BAIAP2L1, BBS9, BCAP29, BCAS3, BCAS4, BCL10, BCL11A, BCL11B, BCL2, BCL3/5, BCL6, BCL7A, BCL9, BCOR, BCR, BIRC3, BRAF, BRD3, BRD4, BTG1, C15orf21, C16orf75, C2orf44, CACNA2D4, CAMTA1, CANT1, CARS, CBFA2T3, CBFB, CBL, CCDC6, CCNB1IP1, CCND1, CCND2, CCND3, CD273, CD274, CD74, CDH11, CDK6, CDR1, CDX1, CDX2, CEP1, CEP85L, CHCHD7, CHD7, CHEK2, CHIC2, CHN1, CIC, CIITA, CLCN6, CLDN18, CLK2, CLTC, CLTCL1, CMKOR1, COL1A1, COL1A2, COL6A3, COX6C, CPB2, CREB1, CREB3L1, CREB3L2, CREBBP, CRLF2, CRTC1, CRTC3, CSF1, CTAGE5, CTNNB1, CUTA, CXorf67, CYP39A1, DAZL, DDIT3, DDX10, DDX5, DDX6, DEK, DUS4L, EBF1, EFNA5, EHF, EIF3E, EIF3K, EIF4A2, EIF4E2, ELF4, ELK4, ELL, ELN, EML1, EML4, EP300, EPC1, EPS15, ERC1, ERCC2, ERG, ERO1L, ESRP1, ETS1, ETV1, ETV4, ETV5, ETV6, EWSR1, EZR, FACL6, FAM131B, FAM22A, FBXL18, FBXO38, FCGR2B, FCHSD1, FERMT2, FEV, FGFR1, FGFR1OP, FGFR3, FHDC1, FHIT, FIP1L1, FLI1, FLII, FLJ27352, FLT3, FN1, FNBP1, FOXO1A, FOXO3A, FOXP1, FSTL3, FUS, FVT1, GAB2, GABBR2, GAS5, GAS6, GAS7, GLI1, GMDS, GMPS, GNAI1, GOLGA5, GOPC, GPBP1L1, GPC3, GPHN, GRAF, GSTP1, HAS2, HCMOGT-1, HEAB, HERPUD1, HEY1, HIP1, HIST1H4I, HJURP, HLF, HLXB9, HMGA1, HMGA2, HN1, HNRNPA2B1, HOOK3, HOXA11, HOXA13, HOXA9, HOXC11, HOXC13, HOXD11, HOXD13, HSPCA, HSPCB, IKZF1, IL2, IL21R, IL6R, INTS4, IQCG, IRF2BP2, IRF4, IRTA1, ITK, ITPR2, JAK2, JAZF1, KDELR2, KDM5A, KIAA0284, KIAA1549, KIF3B,

KIF5B, KLC1, KLK2, KRAS, KTN1, LAF4, LASP1, LCK, LCP1, LCX, LGR5, LHFP, LIFR, LMO1, LMO2, LPP, LRIG3, LYL1, MACROD1, MAF, MAFB, MALAT1, MALT1, MAML2, MARK4, MAST1, MAST2, MBOAT2, MBTD1, MCPH1, MDS2, MEAF6, MECOM, MET, MIPOL1, MKL1, MKRN1, MLF1, MLL, MLLT1, MLLT10, MLLT2, MLLT3, MLLT4, MLLT6, MLLT7, MN1, MSF, MSI2, MSN, MTCP1, MUC1, MYB, MYC, MYH11, MYH9, MYO1C, MYO1F, MYST4, NAB2, NACA, NACC2, NCOA1, NCOA2, NCOA4, NDRG1, NF1, NFATC1, NFATC2, NFIA, NFIB, NFIX, NFKB2, NIN, NONO, NOP2, NOTCH1, NPM1, NR4A3, NRG1, NSD1, NTN1, NTRK1, NTRK2, NTRK3, NUMA1, NUP107, NUP214, NUP98, NUTM1, ODZ4, OLIG2, OMD, OMG, OSCP1, P2RY8, PACS1, PAFAH1B2, PAFAH1B3, PAX3, PAX5, PAX7, PAX8, PBX1, PCM1, PCSK7, PDE4DIP, PDE8B, PDGFB, PDGFRA, PDGFRB, PER1, PHF1, PICALM, PIM1, PKD1L1, PLA2R1, PLAG1, PLXND1, PML, PMX1, PNUTL1, POU2AF1, POU5F1, PPARG, PPFIBP1, PPP1R1B, PPP2R2A, PRCC, PRDM16, PRDX4, PRKAR1A, PRKAR2A, PRKCE, PRKG2, PSIP2, PTPRK, PTPRZ1, QKI, RAB5EP, RAD51L1, RAF1, RALGDS, RANBP17, RAP1GDS1, RARA, RASA3, RBM14, RBM15, RBMS1, RET, RGS17, RGS22, RLF, RNF130, RNF216, ROD1, ROS1, RPL22, RPN1, RRP15, RSPO2, RSPO3, RUNDC2A, RUNX1, RUNX1T1, RUNXBP2, SDC4, SEC16A, SEC31A, SEPT6, SET, SFPQ, SFRS3, SH3GL1, SIL, SIP1, SLC22A1, SLC26A11, SLC26A6, SLC34A2, SLC45A3, SLCO1B3, SMARCA5, SNAP91, SP3, SPTBN1, SQSTM1, SRGAP3, SS18, SS18L1, SSBP2, SSH2, SSX1, SSX2, SSX4, STARD3, STAT5B, STAT6, STL, STRN, SULF2, SUSD1, SUZ12, SYCP1, SYK, SYT1, TACC1, TACC3, TADA2A, TAF15, TAL1, TAL2, TAS2R38, TBCEL, TBL1XR1, TCEA1, TCF12, TCF3, TCF7L1, TCF7L2, TCL1A, TCL6, TECTA, TFE3, TFEB, TFG, TFPT, TFRC, THRAP3, TLX1, TLX3, TMCC1, TMPRSS2, TNFRSF17, TNFRSF1B, TOP1, TOP2B, TOX4, TPM3, TPM4, TPR, TRIM24, TRIM27, TRIM33, TRIP11, TTL, UBE2L3, USH1G, USP6, VCL, VTI1A, WHSC1, WHSC1L1, WIF1, WT1, WWTR1, XKR3, YWHAE, YY1, ZC3H7B, ZNF145, ZNF198, ZNF278, ZNF331, ZNF384, ZNF444, ZNF521, ZNF700, ZNF703, and ZNF9.

Products based on the Methods of the Invention

[00120] Products based on the methods of the invention may be commercialized by the Applicants under the trade name Ovation ®. Ovation ® is a trademark of NuGEN Technologies, Inc.

EXAMPLES

Example 1 – Characterization of the Human Oral Microbiome by Selective Enrichment of Bacterial 16S Ribosomal DNA Sequences.

[00121] Sample nucleic acid

Microbial genomic DNA is isolated from human saliva using the OMNIgene-DISCOVER sample collection kit (DNA Genotek) according to the manufacturer's instructions. Extracted DNA is then fragmented via sonication to an average length of 400 bp and purified using Agencourt AMPure XP beads (Beckman Coulter Genomics).

[00122] Generation of control and test libraries with ligated forward adaptors

The NuGEN Ovation Ultralow Library System (NuGEN Technologies) is used to generate two next generation sequencing libraries from 100 ng of the purified sample. The first library, an unenriched control, is made as recommended by the manufacturer. A second 'test' library, the input for downstream enrichment steps, is generated using the same library construction kit modified as follows. Briefly, DNA is blunted and prepared for ligation under the standard end-repair reaction conditions described in the kit. Fragments are then ligated to the forward adaptor only. As depicted in Figure 2, ligation attaches the forward adaptor to each end of each DNA fragment, leaving a single-strand nick on the opposite strand. Adaptor fill-in will be performed, thus generating ligation products where complementary ends exist on each insert.

[00123] Ligation products of at least 100 bp in length are purified by selective binding to Agencourt AMPure XP beads and taken forward into the enrichment process.

[00124] Amplification

Ribosomal DNA fragments from the test library are selectively amplified with two distinct steps: 1) gene-specific primer extension; and 2) PCR with universal adaptor sequences. The primer extension step is performed with oligonucleotides containing a 3' gene-specific region and a 5'common region that contains a portion of the Illumina reverse adaptor sequence. Consensus 16S sequences making up the gene-specific segment are selected by comparing the ribosomal operons from 40 diverse bacterial species using the ClustalW multiple sequence alignment program (European Bioinformatics Institute). Oligonucleotides representing each of the 18 highly conserved sequence blocks identified across the 16S genomic loci are synthesized and mixed in equimolar proportions.

[00125] The pool of primer extension probes is combined with the test DNA library (above) containing the forward adaptor and the HotStarTaq PCR mastermix (QIAGEN, USA) containing buffer, dNTPs, and a thermally-activated *Taq* DNA polymerase. This solution is placed in a thermal cycler, heated to 95°C for 15 minutes to activate the polymerase and cooled to 70°C for 5 minutes to allow the 16S primers to anneal to DNA inserts and extend into the forward adaptor site. Amplification primers that bind to the forward and reverse adaptor sites are added. Selection for fragments that contain both the forward (test library) and reverse (5' common region on 16S primers) adaptor, and the respective universal priming sites, is accomplished with PCR using a 3-step temperature routine (94°C for 30 seconds, 60°C for 30 seconds, 72° C for 1 minute) for 25 cycles. PCR products are purified using AMPure XP beads and analyzed with a 2100 Bioanalyzer (Agilent Technologies).

[00126] Sequencing and data analysis

[00127] Single end sequencing reads of 100 nt length are obtained for both the control and enriched test libraries using a MiSeq System (Illumina). Raw sequencing data is processed using Illumina base calling software and mapped to a ribosomal RNA database. Sequences that do not align to bacterial rRNA are mapped to human and bacterial full genome reference sequences. Fold enrichment is determined by calculating the number of rRNA reads as a percentage of total mapped reads in the control and test samples.

Example 2 – Characterization of changes over time to the Human Oral Microbiome by Selective Enrichment of Bacterial 16S Ribosomal DNA Sequences.

[00128] Sample nucleic acid

Microbial genomic DNA is isolated from human saliva using the OMNIgene-DISCOVER sample collection kit (DNA Genotek) according to the manufacturer's instructions at 1 hour intervals for 16 hours following use of dental rinse. Extracted DNA is then fragmented via sonication to an average length of 400 bp and purified using Agencourt AMPure XP beads (Beckman Coulter Genomics).

[00129] Generation of DNA fragments with ligated forward adapters

Components from the NuGEN Ovation Ultralow Library System (NuGEN Technologies) are used to generate 16 independent next generation sequencing libraries from 100 ng of the purified sample. Briefly, DNA is blunted and prepared for ligation under the standard end-repair reaction conditions described in the kit. Fragments are then ligated to the forward adapter only. As depicted in Figure 2, ligation attaches the forward adapter to each end of each DNA fragment, leaving a single-strand nick on the opposite strand. Adapter fill-in will be performed, thus generating ligation products where complementary ends exist on each insert.

[00130] Ligation products of at least 100 bp in length are purified by selective binding to Agencourt AMPure XP beads and taken forward into the enrichment process.

[00131] Primer extension

Libraries containing ribosomal genes are generated by introducing the reverse adapter attached to the 5' end of oligonucleotides specific to conserved regions within these genes. There are two distinct steps: 1) annealing of the gene-specific primer; and 2) extension of that primer through the action of a DNA polymerase. The resulting product is a functional library containing the forward adapter on one end and the reverse adapter on the other end. The gene-specific primer extension step is performed with oligonucleotides containing a 3' gene-specific region and a 5' region that contains a portion of the Illumina reverse adapter sequence. Embedded in the reverse adapter sequence is a variable region of 8 bases that differentiates this adapter from the 16 other adapters used with the other samples. Thus, 16 gene-specific libraries have been generated; one from each sample. Each library has a common forward adapter. Each library also contains a common sequence on the opposite end but within that common sequence there is a unique 8 nucleotide region. Consensus 16S sequences making up the gene-specific segment are selected by comparing the ribosomal operons from 40 diverse bacterial species using the ClustalW multiple sequence alignment program (European Bioinformatics Institute). Oligonucleotides representing each of the 18 highly conserved sequence blocks identified across the 16S genomic loci are synthesized and mixed in equimolar proportions.

[00132] Individual samples with forward adapters ligated onto each strand are combined with the primer extension probes (described above) in 16 independent reactions. These are mixed with HotStarTaq PCR mastermix (QIAGEN, USA) containing buffer, dNTPs, and a thermally-activatable *Taq* DNA polymerase. This solution is placed in a thermal cycler, heated to 95°C for 15 minutes to activate the polymerase and cooled to 70°C for 5 minutes to allow the 16S primers to anneal to DNA inserts and extend into the forward adapter site.

[00133] Amplification

The 16 individual primer extension products (above) are pooled, amplification primers that are complementary to the 5' ends of the forward and reverse adapter sites but also contain portions complementary to flow cell oligonucleotide sequences are added. Selection for fragments that contain both the forward and reverse (5' common region on 16S primers) adapter, and the respective universal priming sites, is accomplished with PCR using a 3-step temperature routine (94°C for 30 seconds, 60°C for 30 seconds, 72° C for 1 minute) for 25 cycles. PCR products are purified using AMPure XP beads and analyzed with a 2100 Bioanalyzer (Agilent Technologies).

[00134] Sequencing and data analysis

[00135] Single end sequencing reads of 100 nt length are obtained for both the control and enriched test libraries using a MiSeq System (Illumina). Raw sequencing data is processed using Illumina base calling software. Samples from the various time points are binned based on their unique 8 base code and mapped to a ribosomal RNA database. Sequences that do not align to bacterial rRNA are mapped to human and bacterial full genome reference sequences. Changes in microbial populations are assessed by comparing 16S read counts from the different organisms in the samples over time.

Example 3 – Characterization of transcriptional activity of individual cells within a population.

[00136] Sample nucleic acid

Individual cells are isolated from whole blood using a FACS cell sorter. The cells are suspended in 10 µl of Prelude Lysis solution (a component of NuGEN Technologies, One Direct system), resulting in lysis of the cell membrane while the nuclear membrane remains intact. Sixteen of the single cell suspensions are selected for expression profiling. Briefly, kit reagents are used as described by the manufacturer to generate first and second strand cDNA from the total RNA present in the lysate. Double stranded cDNA products are purified using Agencourt AMPure XP beads (Beckman Coulter Genomics).

[00137] Generation of fragments with ligated forward adapters

Components from the NuGEN Ovation Ultralow Library System (NuGEN Technologies) are used to generate next generation sequencing libraries from each of the purified sample. Briefly, DNA is blunted and prepared for ligation under the standard end-repair reaction conditions described in the kit. Fragments are then ligated to the forward adapter only. As depicted in Figure 2, ligation attaches the forward adapter to each end of each DNA fragment, leaving a single-strand nick on the opposite strand. Adapter fill-in will be performed, thus generating ligation products where complementary ends exist on each insert.

[00138] Ligation products of at least 100 bp in length are purified by selective binding to Agencourt AMPure XP beads and taken forward into library generation.

[00139] Primer extension

Libraries are generated by introducing the reverse adapter attached to the 5' end of a random hexamer. There are two distinct steps: 1) annealing of the primer; and 2) extension of that primer through the action of a DNA polymerase. The resulting product is a functional library containing the forward adapter on one end and reverse

adapter on the other end. The primer extension step is performed with oligonucleotides containing a 3' random region and a 5' region that contains a portion of the Illumina reverse adapter sequence. Embedded in the reverse adapter sequence is a variable region of 8 bases that differentiates this adapter from the 16 other adapters used with the other samples. Thus, 16 libraries have been generated; one from each sample. Each library has a common forward adapter. Each library also contains a common sequence on the opposite end but within that common sequence there is a unique 8 nucleotide region.

[00140] Individual samples with forward adapters ligated onto each strand are combined with the primer extension probes (described above) in 16 independent reactions. These are mixed with HotStarTaq PCR mastermix (QIAGEN, USA) containing buffer, dNTPs, and a thermally-activatable *Taq* DNA polymerase. This solution is placed in a thermal cycler, heated to 95°C for 15 minutes to activate the polymerase and cooled to 70°C for 5 minutes to allow the primers to anneal to DNA inserts and extend into the forward adapter site.

[00141] Amplification

Amplification primers that are complementary to the 5' ends of the forward and reverse adapter sites but also contain portions complementary to flow cell oligonucleotide sequences are added to the 16 individual primer extension products (above). Selection for fragments that contain both the forward and reverse adapter, and the respective universal priming sites, is accomplished with PCR using a 3-step temperature routine (94°C for 30 seconds, 60°C for 30 seconds, 72° C for 1 minute) for 25 cycles. PCR products are purified using AMPure XP beads and analyzed with a 2100 Bioanalyzer (Agilent Technologies).

[00142] Sequencing and data analysis

Equal masses of each of the amplified libraries (above) are pooled and diluted to working concentrations according to manufacturer's recommendations. Single end sequencing reads of 100 nt length are obtained for libraries using a MiSeq System (Illumina). Raw sequencing data is processed using Illumina base calling software. Samples from the various time points are binned based on their unique 8 base code and mapped to a reference database. Based on the mapping characteristics, individual samples or a new pool of samples can be rerun on the sequencer to obtain greater read depth. Samples with poor gene coverage will be eliminated from the pool.

Example 4 – cDNA enrichment from target RNA

[00143] First Strand cDNA Synthesis

[00144] In some cases the following protocol can be used: First Strand Primer Mix (A1), First Strand Buffer Mix (A2), First Strand Enhancer (A4), First Strand Enzyme Mix (A3) and the Nuclease-free Water (D1) from -20°C storage. Also remove the Agencourt RNAClean XP Beads from 4°C storage and place at room temperature. Spin down the contents of A3 and place on ice. Thaw the other reagents at room temperature, mix by vortexing, spin and place on ice. Leave the nuclease-free water at room temperature. Add 2 µL of A1 to a 0.2 mL PCR tube. Add 10-200 ng of total RNA to the primer, bringing the total volume to 6 µL with nuclease-free water if

necessary. Mix by pipetting 5 times, spin and place on ice. Place the tubes in a pre-warmed thermal cycler programmed to run Program 1: 65°C for 1 min.

[00145] Immediately remove the tubes from the thermal cycler and snap chill on ice. Prepare First Strand master mix by combining and mixing according to the volumes shown below (volumes listed are for a single reaction):

First Strand Buffer Mix (A2)	First Strand Enhancer Mix (A4)	First Strand Enzyme Mix (A3)
2.5 µl	1 µl	0.5 µl

[00146] Add 4 µL of the First Strand Master Mix to each sample tube. Mix by pipetting, cap and spin the tubes and place on ice. Place the tubes in a pre-warmed thermal cycler programmed to run Program 2: 25°C for 5 min, 40°C for 5 min, 70°C for 15 min, and hold at 4°C. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice. Continue immediately to the Second Strand cDNA Synthesis.

[00147] Second Strand cDNA Synthesis

In some cases, the following protocol can be used: Remove Second Strand Buffer (B1), Second Strand Enzyme Mix (B2), and Second Strand Stop Mix (B3) from -20°C storage. Spin down the contents of B2 and place on ice. Thaw reagents B1 and B3 at room temperature, mix by vortexing, spin and place on ice. Prepare a master mix by combining B1 and B2 with water (D1) in a 0.5 mL capped tube, according to the volumes shown below (volumes listed are for a single reaction):

Water (D1)	Second Strand Buffer Mix (B1)	Second Strand Enzyme Mix (B2)
55.5 µl	7.5 µl	2.0 µl

[00148] Add 65 µL of the Second Strand Master Mix to each First Strand reaction tube. Mix by pipetting, spin and place on ice. Place the tubes in a pre-cooled thermal cycler programmed to run Program 3: 16°C for 60 min, and hold at 4°C. Remove the tubes from the thermal cycler and spin to collect condensation. Add 2 µL of Second Strand Stop Buffer (B3). Mix by pipetting and spin. Continue with cDNA purification.

[00149] cDNA Purification

[00150] In some cases, the following protocol can be used: Ensure RNAClean XP beads are fully resuspended and at room temperature. Aliquot 120 µl of RNAClean XP beads into a fresh tube for each sample to be purified. In some cases, the beads can be added first to the tube, before the sample is added. Add 78 µl of second strand reaction to the beads and mix by pipetting the full volume at least 7 times. Incubate for 5 minutes at room temperature. Place tubes on a magnet for 5 minutes to collect the beads. Remove and discard 185 µl of supernatant; attempt to avoid discarding any beads. Wash beads by gently pipetting 200 µl of fresh 85% ethanol into each tube while still on the magnet. Allow to sit for 30 seconds; in some cases, allow to sit for no more than 1 minute. Remove all ethanol and allow beads to dry on magnet for 10 minutes. Remove tubes from magnet and

fully resuspend beads in 11 µl of DNA Resuspension Buffer. Incubate at room temperature for 2 minutes. Place tubes on magnet for 5 minutes to collect beads. Transfer purified cDNA supernatant into a fresh tube. Proceed to End Repair Step of Ovation Target Enrichment protocol using 10 µl of cDNA as input. In some case, Covaris fragmentation of cDNA can be performed. In other cases, Covaris fragmentation of cDNA can be omitted.

Example 5 – cDNA enrichment from target RNA using, e.g., Ovation Target Enrichment System

[00151] 8-, 32-, and 96-sample cancer panel and FFPE cancer panel, 32- and 96-sample custom kits

[00152] I - Introduction

[00153] Methods, compositions and kits provided herein, e.g., the Ovation Target Enrichment System can provide custom target enrichment and multiplex library preparation solution for targeted resequencing. A fully customizable probe design can enable the interrogation in some cases of up to 10 Mb of custom content defined by a user. In some cases, the system can work with as little as 10 ng of fresh/frozen tissue gDNA as well as DNA isolated from FFPE tissue. A single-tube workflow can be completed in 24 hours with an optional fast protocol that can be completed in as little as 8 hours from sample to target enriched library. A barcoding design can allow both a high degree of sample multiplexing and the unambiguous identification of true PCR duplicates. In some cases, paired-end sequencing is not used. Each kit includes reagents for multiplexed target enrichment and Illumina sequencing library creation for 8, 32 or 96 samples, respectively.

[00154] *B. Workflow*

[00155] In the workflow shown in Figure 5, fragmented samples are end-repaired and individually ligated with barcoded adaptors. Following a bead purification step, samples are combined for multiplex target enrichment during which time the targeting probes are annealed and extended, creating amplifiable library molecules. Following a second bead purification, an optional qPCR step can determine the optimal number of PCR cycles to use in the final library amplification step. After library amplification, final bead purification prepares the multiplexed, target enriched library for sequencing. Using this workflow, a single user can easily process 32 samples or more at a time.

[00156] C. *Performance Specifications*

[00157] Methods, compositions and kits described herein, e.g., the Ovation Target Enrichment System can produce custom target-enriched, multiplex libraries suitable for sequencing on the Illumina Genome Analyzer IIx/IIe (GAII), MiSeq, NextSeq, HiScan SQ or HiSeq 2000/2500 NGS platforms. The methods, compositions and kits can produce target-enriched libraries from about 10 to about 500 ng fresh or FFPE-derived genomic DNA. Sequencing metrics such as on-target rates and uniformity of coverage can vary from one custom probe design to another.

[00158] D. *Quality Control*

[00159] Reagents provided herein, e.g., the Ovation Target Enrichment System can undergo functional testing to assess performance specifications.

[00160] E. *Storage and Stability*

[00161] Compositions and kits provided herein, e.g., the Ovation Target Enrichment System can be shipped on dry ice and can be unpacked immediately upon receipt. A product can contain components with multiple storage temperatures. Vials labeled Agencourt® RNAClean® XP Beads (clear cap) can be removed from the top of the shipping carton upon delivery and be stored at 4°C. All other components can be stored at -20°C on internal shelves of a freezer without a defrost cycle. In some cases, a kit is tested to perform to specifications after as many as six freeze/thaw cycles. Kits handled and stored according to the above guidelines can perform to specifications for at least six months.

[00162] II. Components

[00163] A. *Reagents Provided*

Table 1. Ovation Target Enrichment Core Library Reagents

COMPONENTS	0400-08 PART NUMBER	0400-32 PART NUMBER	0400-96 PART NUMBER	VIAL CAP	VIAL NUMBER
End Repair Buffer Mix	S01860	S01899	S01907	Blue	ER1 VER 9
End Repair Enzyme Mix	S01510	S01533	S01909	Blue	ER2 VER 4
End Repair Enhancer	S01862	S01709	S01908	Blue	ER3 VER 2
End Repair FFPE Enhancer	S01938	S01935	S01939	Blue	ER5
Ligation Buffer Mix	S01936	S01534	S01910	Yellow	L1 VER 4
Ligation Enzyme Mix	S01467	S01535	S01911	Yellow	L3 VER 4
DNA Resuspension Buffer Mix	S01901	S01901 (2)	S01901 (3)	Green	DR1
Target Extension Buffer Mix	S01866	S01900	S01912	Clear	TX1
Target Extension Enzyme Mix	S01870	S01902	S01913	Clear	TX2
Amplification Buffer Mix	S01871	S01903	S01914	Red	P1 VER 4
Amplification Primer Mix	S01934	S01906	S01917	Red	P2 VER 14
Amplification Enzyme Mix	S01873	S01764	S01895	Red	P3 VER 2
Amplification Enhancer Mix	S01872	S01904	S01915	Red	P6

Ovation Cancer Panel Target Enrichment System Reagents (Part Nos. 0400-08, 0400-32, and 0400-96), *continued*					
COMPONENTS	**0400-08 PART NUMBER**	**0400-32 PART NUMBER**	**0400-96 PART NUMBER**	**VIAL CAP**	**VIAL NUMBER**
Nuclease Free Water	S01001	S01001, S01113	S01001 (2), S01113 (3)	Green	D1
RNACleanXP Beads	S01502	S01502	S01502	Clear	—
Ligation Adaptor Mix	S01941 S01942 S01943 S01944 S01945 S01946 S01947 S01948	N/A	N/A	Yellow	L2V15DR-BC1 L2V15DR-BC2 L2V15DR-BC3 L2V15DR-BC4 L2V15DR-BC5 L2V15DR-BC6 L2V15DR-BC7 L2V15DR-BC8
Target Enrichment 32-Plex Adaptor Plate	N/A	S01937	N/A	N/A	N/A
Target Enrichment 96-Plex Adaptor Plate	N/A	N/A	S01940	N/A	N/A

Table 2. Ovation Target Enrichment Probes

TARGET ENRICHMENT SYSTEM	08 RXN KIT PROBE PART NUMBER	32 RXN KIT PROBE PART NUMBER	96 RXN KIT PROBE PART NUMBER	VIAL CAP	VIAL NUMBER
Ovation Cancer Panel Target Enrichment System (Part Nos. 9079-08, 9079-32, 9079-96)	S02007	S02008	S02008 (3)	Clear	N/A
Ovation Cancer Panel FFPE Target Enrichment System (Part Nos. 9099-08, 9099-32, 9099-96)	S02009	S02010	S02010 (3)	Clear	N/A
Ovation Custom Target Enrichment System (Part Nos. 9###-32, 9###-96)	N/A	S02011	S02011 (3)	Clear	N/A

[00164] B. *Additional Equipment, Reagents and Labware*

[00165] The equipment used in the method in this example can include Covaris Sonication System, Agilent 2100 Bioanalyzer or materials and equipment for electrophoretic analysis of nucleic acids, Real-time PCR system capable of SYBR Green detection, microcentrifuge for individual 1.5 mL and 0.5 mL tubes, microcentrifuge for individual or strip 0.2 mL tubes, vortexer, thermal cycler with 0.2 mL tube heat block, heated lid, and 100 µL reaction capacity, and appropriate spectrophotometer and cuvettes or Nanodrop UV-Vis Spectrophotometer for quantitation of fragmented DNA.

[00166] The reagents used in the method in this example can include ethanol (e.g., Sigma-Aldrich, Cat. #E7023) for e.g., purification steps, Eva Green Dye (e.g., Biotium, Cat. #31000) or SYBR Green for, e.g., qPCR

[00167] The supplies and labware used in the method in this example can include Amicon Ultra 30K MWCO (e.g., Millipore, Cat. #UFC503024) or vacuum, concentrator, nuclease-free pipette tips, 1.5 mL and 0.5 mL RNase-free microcentrifuge tubes, 0.2 mL individual thin-wall PCR tubes or 8 X 0.2 mL strip PCR tubes or 0.2 mL, thin-wall PCR plates, magnetic bead separation device, disposable gloves, Kimwipes, ice buckets, cleaning solutions such as DNA-OFFTM (e.g., MP Biomedicals, Cat. #QD0500).

[00168] The equipment, reagents, supplies and/or labware can be ordered from companies including Agilent, Inc., Biotium, Covaris, Millipore, MP Biomedicals, and Sigma-Aldrich, Inc.

[00169] III. Planning the Experiment

[00170] A. *Input DNA Requirements*

[00171] Methods, compositions and kits provided herein, e.g., the Ovation Target Enrichment System can work with inputs of 10 ng to 500 ng of fragmented genomic DNA. DNA samples can be free of contaminating proteins, RNA, organic solvents (e.g., phenol and ethanol) and salts. In some cases, use of a commercially available system for genomic DNA isolation is recommended. In some cases, the A260:A280 ratio for DNA samples should be in excess of 1.8. Use of DNA samples with lower ratios can result in low amplification yield. Genomic DNA derived from FFPE samples can work well in the Ovation Target Enrichment System.

[00172] B. *Working with the Target Enrichment 32-Plex and 96-Plex Adaptor Plates*

[00173] Adaptor plates, e.g., the Target Enrichment Adaptor Plates, can contain ligation adaptor mixes, each with a unique eight-base barcode. Each well (first 32 wells (A01–H04) or all 96 wells, respectively) can contain sufficient volume for preparation of a single library. The Target Enrichment Adaptor Plates can be sealed with a foil seal designed to provide for airtight storage. Prior to thawing the adaptor plate, it can be spun for 5 to 10 minutes at room temperature at about 1000 X g in a centrifuge with a rotor appropriate for microwell plates. This can allow the plate to thaw while spinning and help to minimize cross-contamination between wells. To use only a portion of the plate, the seal covering only the portion of the plate you wish to use can be removed. The remaining wells of the plate can remain sealed for use at a later date. Used wells can be covered with a new foil seal (e.g., AlumaSeal II) to prevent any remaining adaptor-containing liquid from contaminating future reactions.

[00174] C. *Enriched, Amplified Library Storage* Amplified libraries may be stored at -20°C.

[00175] IV. Overview

[00176] A. *Overview*

[00177] The library preparation process used in the methods described herein, e.g., the Ovation Target Enrichment System can be performed in four stages: 1. DNA end repair for 0.75 hours; 2. adaptor ligation and purification for 1.75 hours; 3. probe hybridization and extension for from 3 to 12 or more hours (optionally, this step can be performed for about 3 hours); 4. amplification and purification for 2.25 hours. Total time to prepare enriched library can be about 8 to about 20 hours.

[00178] Components in kits provided herein, e.g., the Ovation Target Enrichment System, can be color coded, with each color linked to a specific stage of the process. Performing each stage can require making a master mix then adding it to a reaction, followed by incubation. Master mixes can be prepared by mixing components provided for that stage.

[00179] B. *Protocol Notes*

[00180] A positive control DNA can be routinely used to establish of a baseline of performance and provide the opportunity for an experimenter to become familiar with the protocol steps.

[00181] Water used in this method can be the water provided with a kit (e.g., D1) or an alternate source of nuclease-free water. In some cases, use of DEPC-treated water is not preferred with this protocol.

[00182] Components used in each step can be thawed and immediately placed on ice. In some cases, all reagents are not thawed at once.

[00183] Thawed reagents and reaction tubes can be kept on ice.

[00184] After thawing and mixing buffer mixes, if any precipitate is observed, the mixes can be re-dissolved completely prior to use. The buffer mixes can be gently warmed for 2 minutes at room temperature followed by brief vortexing. In some cases, enzyme mixes are not warmed.

[00185] When placing small amounts of reagents into the reaction mix, pipetting up and down several times can be helpful for transfer.

[00186] Pipetting a mix can be done by gently aspirating and dispensing a volume that is at least half of the total volume of the reaction mix.

[00187] Tubes or plates can be placed in the block after the thermal cycler to reach the initial incubation temperature.

[00188] Master mixes can be prepared by using the minimal amount of extra material to help run the maximal number of reactions using the components provided in a kit.

[00189] In some cases, the Ovation Target Enrichment Systems are not used with components and reagents from other kits, e.g., kits produced by NuGEN.

[00190] The ethanol mixes can be made fresh by measuring both the ethanol and water with pipettes. In some cases, fresh ethanol stocks are used to make 85% ethanol used in the purification steps.

[00191] C. *Agencourt® RNAClean® XP Purification Beads*

[00192] There can be significant modifications to the Agencourt RNAClean XP beads standard procedure; in some cases, it is preferred to follow the protocols outlined in this user guide for the use of these beads. However, the Beckman Coulter user guide may be reviewed to become familiar with the manufacturer's recommendations.

[00193] The bead purification processes used in this kit can consist of the following steps (Figure 6): 1. Binding of DNA to RNAClean XP beads; 2. Magnetic separation of beads from supernatant; 3. Ethanol wash of bound beads to remove contaminants; 4. Elution of bound DNA from beads.

[00194] *Tips and Notes*

[00195] Beads can be removed from 4°C and be left at room temperature for at least 30 minutes before use, and in some case, the reagents can have completely reached room temperature. Cold beads can reduce recovery. Beads can fully resuspended by inverting and tapping before adding to sample. The beads can be separated on the magnet for a full 5 minutes. Binding buffer can be removed before the beads have completely separated will impact DNA yields. After completing the binding step, bead loss can be minimized when removing the binding buffer. Significant loss of beads during the ethanol washes can impact DNA yields, so in some cases it is preferred to minimize bead loss throughout the procedure. The ethanol wash can be freshly prepared from fresh ethanol stocks at the indicated concentration. Lower percent ethanol mixes can reduce recovery. During the ethanol washes, keep the samples on the magnet. In some cases, the beads should not be allowed to disperse; the magnet can keep the beads on the walls of sample wells or tubes in a small ring. All residual ethanol can be removed prior to continuing with the next step. For example, when removing the final ethanol wash, most of the ethanol can first be removed, then the excess can be collected at the bottom of the tube before removing the remaining ethanol. This can reduce the required bead air drying time. After drying the beads for the time

specified in the protocol, each tube can be inspected carefully so that the ethanol has evaporated before proceeding. Strip tubes or partial plates can be firmly placed when used with the magnetic plate. In some cases individual tubes are not used because they are difficult to position stably on the magnetic plates.

[00196] D. *Programming the Thermal Cycler*

[00197] Use a thermal cycler with a heat block designed for PCR plates, equipped with a heated lid, and with a capacity of 100 µL reaction volume. Prepare the programs shown in Table 3 following the operating instructions provided by the manufacturer. For thermal cyclers with an adjustable heated lid, in some cases the lid temperature can be set to 100°C only when sample temperature reaches above 30°C. For thermal cyclers with a fixed temperature heated lid (e.g., ABI GeneAmp® PCR 9600 and 9700 models), the default settings (e.g., 100 to 105°C) can be used.

Table 3. Thermal Cycler Programming

END REPAIR	
Program 1 End Repair/Ligation	25°C – 30 min, 70°C – 10 min, hold at 4°C
DNA REPAIR (OPTIONAL)	
Program 2 FFPE DNA Repair (optional)	16°C – 10 min, 70°C – 10 min, hold at 4°C
HYBRIDIZATION	
Program 3 Probe Hybridization & Extension	95°C – 5 min, 200 cycles (80°C – 10 sec, decrease 0.1°C each cycle), hold at 60°C, 72°C – 10 min, hold at 4°C
AMPLIFICATION	
Program 4 Library Amplification	37°C – 10 min, 95°C – 3 min, 15–30 cycles* (95°C – 30 sec, 62°C – 15 sec, 72°C – 20 sec), 72°C – 2 min, hold at 4°C

*The number of cycles may be adjusted and optimized for desired yield based on qPCR assessment of the unamplified library.

[00198] V. Protocol

[00199] A. *Prepare gDNA*

[00200] Genomic DNA used in the target enrichment protocol can be fragmented to an average of 500 bp in size. Covaris acoustic shearing can be used following the manufacturer's protocols. The gDNA can be concentrated after fragmentation as the protocol requires the gDNA to be in a volume of 10 µL. The volume of the gDNA can be reduced by using a vacuum concentrator (Speedvac) or an Amicon concentration column according to the manufacturer's instructions. Alternatively, Agencourt AMPure XP or RNAClean XP beads can be used to purify

and concentrate the fragmented DNA. Since loss of material can occur during bead purification, quantitation of the samples prior to End Repair can be repeated.

[00201] The 8-sample Cancer Panel and FFPE Cancer Panel kits can be used for multiplex enrichment of 4 samples at a time. Processing fewer than 4 samples can result in insufficient Targeting Probe Pool and PCR reagents to complete the reactions. The 32-sample and 96-sample enrichment kits can be used for multiplex enrichment of 8 samples at a time. Processing fewer than 8 samples can result in insufficient Targeting Probe Pool and PCR reagents to complete the reactions. A heated lid can be used to prevent evaporation. PCR tubes or plates that seal tightly and may not deform when the thermal cycler lid is closed can be used. Insufficient tube sealing can result in sample loss due to evaporation during extended incubation steps.

[00202] Samples intended to be multiplexed together at Target Enrichment (e.g., step F) can be of equivalent mass, quality and average fragment length. Failure to normalize samples that are processed together can result in unequal sample representation in the final target-enriched library.

[00203] B. *End Repair*

[00204] The following end repair process can be used in some cases: 1. The End Repair Buffer Mix (blue: ER1), End Repair Enzyme Mix (blue: ER2) and End Repair Enhancer (blue: ER3) can be obtained from –20°C storage. 2. ER1 can be thawed at room temperature. Mix by vortexing, spin and place on ice. 3. Spin down contents of ER2 and ER3 and place on ice. 4. Prepare a master mix by combining ER1, ER2 and ER3 in a 0.5 mL capped tube, according to the volumes shown in Table 4. 5. Add 5 µL of the End Repair Master Mix to 10 µL of fragmented DNA (10 ng to 500 ng). 6. Mix by pipetting, cap and spin tubes and place on ice. 7. Place the tubes in a pre-warmed thermal cycler programmed to run Program 1 (End Repair/Ligation; see Table 3): 25°C -30 min, 70°C -10 min, and hold at 4°C. 8. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice. 9. Proceed to DNA Repair (optional, Section C) or continue immediately with the Adaptor Ligation protocol (Section D).

Table 4. End Repair Master Mix (Volumes listed are for a single reaction)

END REPAIR BUFFER MIX (BLUE: ER1 VER 9)	END REPAIR ENZYME MIX (BLUE: ER2 VER 4)	END REPAIR ENHANCER (BLUE: ER3 VER 2)
4 µL	0.5 µL	0.5 µL

[00205] C. *FFPE DNA Repair (Optional)*

[00206] This is an optional repair step for use with damaged DNA, e.g., FFPE DNA. 1. Add 1 µL of End Repair FFPE Enhancer (ER5). 2. Mix well by pipetting. 3. Place the tubes in a pre-warmed thermal cycler programmed to run Program 2 (FFPE DNA Repair; see Table 3): 16°C -10 min, 70°C -10 min, and hold at 4°C. 4. Remove the tubes from the thermal cycler, spin briefly and place on ice.

[00207] D. *Adaptor Ligation*

[00208] In some cases, the following adaptor ligation protocol can be used: 1. Remove the Ligation Buffer Mix (yellow: L1), Ligation Enzyme Mix (yellow: L3) and Nuclease-free Water (D1) from –20°C storage. For 8-reaction kits: Remove Ligation Adaptor Mixes (L2V15DR-BC1 to BC8) from -20°C storage. For 32- and 96-reaction kits: Remove Target Enrichment Adaptor Plate (32-plex or 96-plex) from –20°C storage. 2. Thaw L1, Nuclease-free Water and Ligation Adaptor Mixes at room temperature. Mix L1 by vortexing, spin and place on ice. 3. For 8-reaction kits: Spin down Ligation Adaptor Mixes and place on ice. For 32- and 96-reaction kits: Spin down Target Enrichment Adaptor Plate and place on ice. 4. Spin down L3 and place on ice. 5. Add 3 µL of Ligation Adaptor Mix directly to each end-repaired sample. Each tube of Ligation Adaptor Mix and each well of Target Enrichment Adaptor Plates contains adaptors with a unique barcode. 6. Make a master mix by combining the Nuclease-free Water (D1), L1 and L3 in a 0.5 mL capped tube, according to the volumes shown in Table 5. Prepare the master mix by combining water and Ligation Buffer Mix, vortexing to mix, then adding Ligation Enzyme Mix. Mix thoroughly by repeatedly pipetting up and down using a large volume (100–200 µL) pipettor - the larger bore tips used with such pipettors can be used to sufficiently mix the Ligation Master Mix and sample. The L1 Ligation Buffer Mix can be viscous so it can be pipetted slowly. 7. Add 12 µL Ligation Master Mix to each tube of sample/adaptor. Mix thoroughly by repeatedly pipetting up and down using a large volume (100–200 µL) pipettor - the larger bore tips used with such pipettors can be used to sufficiently mix the Ligation Master Mix and sample. 8. Place the tubes in a pre-warmed thermal cycler programmed to run Program 1 (End Repair/Ligation; see Table 3): 25°C – 30 min, 70°C -10 min, and hold at 4°C. 9. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice. 10. Continue immediately with the Ligation Purification protocol.

Table 5. Ligation Master Mix (volumes listed are for a single reaction)

WATER (GREEN: D1)	LIGATION BUFFER MIX (YELLOW: L1 VER 4)	LIGATION ENZYME MIX (YELLOW: L3 VER 4)
4.5 µL	6.0 µL	1.5 µL

[00209] E. *Ligation Purification*

[00210] In some cases, the following ligation purification scheme can be used: Retrieve the Agencourt RNAClean XP Beads from 4°C and ensure they are at room temperature. Also remove the DNA Resuspension Buffer (DR1) from –20°C and thaw at room temperature. Resuspend the beads by inverting and tapping the tube. Ensure the beads can be fully resuspended before adding to the sample. After resuspending, the beads do not need to be spun. (An excess of beads can be provided; therefore, it is not necessary to recover any trapped in the cap.) Add 70 µL of room-temperature Nuclease-free Water (D1) to each ligation reaction. Add 80 µL (0.8 volumes) room temperature bead suspension to each sample. Mix thoroughly by pipetting 10 times. Incubate at room

temperature for 5 minutes. Transfer the tubes to the magnet plate and let stand 5 minutes to completely clear the solution of beads. Remove 160 µL of the binding buffer and discard it. Leaving some of the volume behind can minimize bead loss at this step. Care may be taken to prevent beads from being removed with the binding buffer or the wash. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but, in some cases, no more than 1 minute. Remove the ethanol wash using a pipette. Air dry the beads on the magnet for 10 minutes. Inspect each tube to check that all the ethanol has evaporated. All residual ethanol can be removed prior to continuing. Remove the tubes from the magnet. Add 21 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads. Carefully remove 20 µL of the eluate, ensuring as few beads as possible are carried over. Combine the eluted DNA of 8 samples (combine 4 samples when using the 0400-08 core kit) into a single fresh 1.5 mL tube. Concentrate the DNA by either using an Amicon 0.5 mL 30k MWCO spin column (e.g., Millipore UFC503024) following the manufacturers' recommendations or by vacuum concentration. If final volume falls below 32 µL after the concentration step, add water to bring the volume up to 32 µL. At this point, samples can be stored at –20°C until needed or can be used immediately in the Enrichment protocol.

[00211] F. *Target Enrichment*

[00212] In some cases, the following target enrichment protocol can be performed: 1. Remove the Target Extension Buffer Mix (clear: TX1) and Target Probe Pool (e.g., S02007, S02008, S02009, S02010 or S02011) from -20°C storage. 2. Thaw TX1 and Target Probe Pool at room temperature. Mix by vortexing, spin and place on ice. 3. Prepare a master mix by combining TX1 and Target Probe Pool in a 0.5 mL capped tube, according to the volumes shown in Table 6. 4. Add 17 µL of Target Enrichment Master Mix to 32 µL of pooled, ligated gDNA. Mix by vortexing. Spin down briefly. 5. Place the tubes in a pre-warmed thermal cycler programmed to run Program 3 (Probe Hybridization and Extension; see Table 3): 95°C -5 min, 200 cycles (80°C -10 sec, decrease 0.1°C each cycle), hold at 60°C (after a minimum of 12 hours at 60°C, open lid of thermal cycler, open tube caps and, without removing the tubes from the thermal cycler, add 1 µL of Target Extension Enzyme Mix (clear: TX2). Pipette mix up and down more than 5 times with a volume equal to or greater than 40 µL. Reseal the tubes, close the thermal cycler lid and immediately continue to the next step of the program), 72°C for 10 min, and hold at 4°C. 6. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice. 7. Continue immediately with the Enriched Library Purification protocol.

Table 6. Target Enrichment Mix (volumes listed are for a single reaction)

TARGET EXTENSION BUFFER MIX (CLEAR: TX1)	TARGET PROBE POOL (CLEAR: N/A)
7 µL	10 µL

[00213] Step 5 can also be performed by Alternative Quick Hybridization and Extension to run the quick protocol to complete the target enrichment and library preparation in a single day. A larger amount of input gDNA can be used for the Quick Protocol. Lower input amounts can exhibit lower on-target reads and elevated duplication rates. Alternative step 5. Place the tubes in a pre-warmed thermal cycler programmed to run Program 3 (Probe Hybridization and Extension; see Table 3): 95°C for 5 min, 200 cycles (80°C for 10 sec, decrease 0.1°C each cycle), hold at 60°C (After a minimum of 2 hours at 60°C, open lid of thermal cycler, open tube caps and, without removing tubes from the thermal cycler, add 1 µL of Target Extension Enzyme Mix (clear: TX2). Pipette mix up and down more than 5 times with a volume equal to or greater than 40 µL. Reseal the tubes, close the thermal cycler lid and immediately continue to the next step of the program), 72°C – 10 min, 10 cycles of (95°C for 20 sec, 60°C for 30 sec, 72°C for 30 sec) 72°C for 10 min, hold at 4°C.

[00214] G. *Enriched Library Purification*

[00215] In some cases, the following enriched library purification protocol can be used: This procedure can involve two sequential bead purifications. 1. Retrieve the Agencourt RNAClean XP Beads from 4°C and ensure they are at room temperature before using. Also remove the DNA Resuspension Buffer (DR1) from -20°C and thaw at room temperature. 2. Resuspend the beads by inverting and tapping the tube. Ensure the beads are fully resuspended before adding to the sample. After resuspending, the beads do not need to be spun. (An excess of beads can be provided; therefore, it is not necessary to recover any trapped in the cap.) 3. Add 50 µL of room-temperature Nuclease-free Water (D1) to each Target Extension reaction. 4. Add 80 µL (0.8 volumes) room temperature bead suspension to each sample. Mix thoroughly by pipetting 10 times. 5. Incubate at room temperature for 5 minutes. 6. Transfer the tubes to the magnet plate and let stand 5 minutes to completely clear the solution of beads. 7. Remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind can minimize bead loss at this step. Note: The beads can stay on the walls of the tubes. Significant loss of beads at this stage can impact the amount of DNA carried into Library Amplification, so in some cases ensure beads are not removed with the binding buffer or the wash. 8. With the plate still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute. 9. Remove the ethanol wash using a pipette. 10. Air dry the beads on the magnet for a minimum of 10 minutes. Inspect each tube to check that all the ethanol has evaporated. All residual ethanol can be removed prior to continuing. 11. Remove the tubes from the magnet. 12. Add 100 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature. 13. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads.14. Carefully remove 100 µL of the eluate to a new tube. 15. Add 80 µL (0.8 volumes) room temperature bead suspension to the 100 µL eluate. Mix thoroughly by pipetting 10 times. 16. Incubate at room temperature for 5 minutes. 17. Transfer the tubes to the magnet and let stand 5 minutes to completely clear the solution of beads. 18. Carefully remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step. Note: The beads may stay on the walls of the tubes. Significant loss of beads at this stage can impact the amount of DNA carried into Library Amplification, so in some cases ensure beads are not removed with the

binding buffer or the wash. 19. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute. 20. Remove and discard ethanol wash. 21. Air dry the beads on the magnet for at least 10 minutes at room temperature. In some cases, inspect each tube to ensure that all the ethanol has evaporated. All residual ethanol can be removed prior to continuing. 22. Remove the tubes from the magnet. 23. Add 25 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature. 24. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads. 25. Remove 24 µL of the eluate, ensuring as few beads as possible are carried over. 26. At this point, samples can be stored at –20°C until needed or can be used immediately for qPCR (if being performed) or in the Library Amplification protocol.

[00216] H. *qPCR (Optional)*

[00217] When using compositions or kits provided herein, e.g., the Ovation Target Enrichment System for the first time with a new probe set or sample type, a qPCR assay can be performed in order to determine the number of cycles necessary for library amplification.

[00218] I. *Library Amplification*

[00219] In some cases, the following library amplification protocol can be used: 1. Remove the Library Amplification Buffer Mix (P1) Amplification Primer Mix (P2), Amplification Enzyme Mix (red: P3) and Amplification Enhancer Mix (P6) from –20°C storage. 2. Spin down P3 and place on ice. 3. Thaw P1 and P2 at room temperature. Mix each by vortexing, spin and place on ice. 4. Prepare a master mix by combining P1, P2, P3, P6 and water in a 0.5 mL capped tube, according to the volumes shown in Table 7. 5. Add 78 µL of Library Amplification Master Mix to 22 µL of purified DNA from step G.25 above. 6. Mix well by pipetting. 7. Place the tubes in a pre-warmed thermal cycler programmed to run Program 4 (Library Amplification; see Table 3): 37°C for 10 min, 95°C for 3 min, 15-30 cycles* (95°C for 30 sec, 62°C for 15 sec, 72°C for 20 sec; number of cycles determined by qPCR), 72°C for 2 min, and hold at 4°C. 8. Remove the tubes from the thermal cycler, spin to collect condensation and place on ice. 9. Continue immediately with the Amplified Library Purification protocol.

Table 7. Amplification Master Mix (volumes listed are for a single reaction)

AMPLIFICATION BUFFER MIX (RED: P1 VER 4)	AMPLIFICATION PRIMER MIX (RED: P2 VER 14)	AMPLIFICATION ENHANCER MIX (RED: P6)	AMPLIFICATION ENZYME MIX (RED: P3 VER 2)	WATER (GREEN: D1)
20 µL	8 µL	4 µL	2 µL	44 µL

[00220] J. *Amplified Library Purification*

[00221] In some cases, the following amplified library purification protocol can be used: This procedure involves two sequential bead purifications. 1. Retrieve the Agencourt RNAClean XP Beads from 4°C and ensure they are at room temperature before using. Also remove the DNA Resuspension Buffer (DR1) from -20°C and thaw at

room temperature. 2. Resuspend the beads by inverting and tapping the tube. Ensure the beads are fully resuspended before adding to the sample. After resuspending, the beads do not need to be spun. (An excess of beads can be provided; therefore, it is not necessary to recover any trapped in the cap.) 3. Add 80 µL (0.8 volumes) room temperature bead suspension to each sample. Mix thoroughly by pipetting 10 times. 4. Incubate at room temperature for 5 minutes. 5. Transfer the tubes to the magnet plate and let stand 5 minutes to completely clear the solution of beads. 6. Remove 160 µL of the binding buffer and discard it. Leaving some of the volume behind can minimize bead loss at this step. Note: The beads can stay on the walls of the tubes. Care may be taken to prevent beads from being removed with the binding buffer or the wash. 7. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute. 8. Remove and discard the ethanol wash. 9. Air dry the beads on the magnet for at least 10 minutes at room temperature. Inspect each tube to check that all the ethanol has evaporated. Remove ethanol prior to continuing. 10. Remove the tubes from the magnet. 11. Add 100 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature. 12. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads. 13. Remove 100 µL of the eluate to a new tube, attempting to transfer as few beads as possible. 14. Add 80 µL (0.8 volumes) room temperature bead suspension to the 100 µL eluate. Mix thoroughly by pipetting 10 times. 15. Incubate at room temperature for 5 minutes. 16. Transfer the tubes to the magnet and let stand 5 minutes to completely clear the solution of beads. 17. Remove only 160 µL of the binding buffer and discard it. Leaving some of the volume behind minimizes bead loss at this step. The beads can stay on the walls of the tubes. Care can be taken to prevent the beads from being removed with the binding buffer or the wash. 18. With the tubes still on the magnet, add 180 µL of freshly prepared 85% ethanol and allow to stand for at least 30 seconds but no more than 1 minute. 19. Remove and discard ethanol wash. 20. Air dry the beads on the magnet for at least 10 minutes at room temperature. Inspect each tube carefully to ensure that all the ethanol has evaporated. Remove residual ethanol prior to continuing. 21. Remove the tubes from the magnet. 22. Add 25 µL of DNA Resuspension buffer (DR1) to the dried beads. Mix thoroughly by pipetting to ensure all the beads are resuspended. Incubate for 2 minutes at room temperature. 23. Transfer the tubes to the magnet and let stand for 3 minutes to completely clear the solution of beads. 24. Carefully remove 24 µL of the eluate, ensuring as few beads as possible are carried over. 25. The library is now ready for quantification and sequencing and can be stored at -20°C until needed.

[00222] K. *Quantitative and Qualitative Assessment of the Library*

[00223] 1. Run 0.5–1 µL of the sample on a Bioanalyzer High Sensitivity DNA Chip (Agilent Technologies). A typical fragment distribution is shown in Figure 7, below. 2. Validate the library as described in the most recent Illumina user guides for DNA library construction, e.g., Genomic DNA Sample Prep Manual.

[00224] VI Guidelines for Optional qPCR

[00225] When using methods ,compositions or kits as provided herein, e.g., the Ovation Target Enrichment System for the first time with a new probe set or sample type, a qPCR can be performed to assay after protocol

step G.25 in order to determine the number of cycles necessary for subsequent library amplification. In some cases, the following protocol is used: 1. Remove the Library Amplification Buffer Mix (P1) Amplification Primer Mix (P2), Amplification Enzyme Mix (P3) and Amplification Enhancer Mix (P6) from –20°C storage. 2. Spin down P3 and place on ice. 3. Thaw P1 and P2 at room temperature. Mix each by vortexing, spin and place on ice. 4. Prepare a master mix by combining P1, P2, P3, P6, water and dye (Eva Green or SYBR Green reagent) in a 0.5 mL capped tube, according to the volumes shown in Table 8. 5. Add 8 µL of qPCR Master Mix to 2 µL of purified library DNA (or water for the negative control) in a qPCR plate and mix well. 6. Perform qPCR using the following protocol: 37°C for 10 min, 95°C for 3 min, 35 cycles (95°C for 30 sec, 62°C for 15 sec, 72°C for 20 sec). The cycle number used for subsequent library amplification can be within the exponential phase of the amplification plot (17 or 18 cycles in Figure 8 as an example). In some cases, there is sufficient reagent included to perform a single test qPCR reaction for each of the multiplexed libraries.

Table 8. qPCR Master Mix (volumes listed are sufficient for 2 reactions and a small overage)

AMPLIFICATION BUFFER MIX (RED: P1 vial 4)	AMPLIFICATION PRIMER MIX (RED: P2 vial 14)	AMPLIFICATION ENHANCER MIX (RED: P6)	AMPLIFICATION ENZYME MIX (RED: P3 vial 2)	QPCR DYE AT 1X FINAL REACTION CONC.	WATER (GREEN: D1)
5 µL	2 µL	1 µL	0.5 µL	X	To total volume of 20 µL

[00226] VII Sequencing Recommendations and Guidlines

[00227] Standard Illumina single end 50–150 base sequencing with a 14-base Indexing read can be used for most applications using methods, compositions and kits provided herein, e.g., the Ovation Target Enrichment System. For detecting translocations it can be beneficial to perform paired end sequencing. First, in some cases, trim the first 65 bases of the reverse read as this sequence is primarily probe derived and not useable for identifying variants. The second consideration with paired end sequencing is that the first 15 bases of all probe sequences are the same sequence and on the HiSeq platform canresult in reverse read failures due to lack of diversity. This is not the case on the MiSeq which uses the color matrix generated in read 1 for all subsequent reads. However, the HiSeq recalculates the color matrix for the reverse read. To mitigate this problem, the user can specify a different lane of the flow cell that contains sufficient diversity to be used for lane alignment and error rate calculations. Please refer to Illumina's documented guidance on using a separate lane of sequence for lane alignment and error rate calculation or consult with Illumina technical support. The system uses 8-base barcodes and the same approach to multiplexing found in the standard Illumina method. The resulting libraries can be sequenced using the Illumina protocol for multiplex sequencing of 8-base barcodes, but with an additional 6 bases added to the index read for a total of 14 bases. The additional 6 bases can be used for true duplicate read determination using the Duplicate Marking Software Tool. The 8-base barcode sequences can be accessed through the barcodes link

in the left hand menu of the Ovation Target Enrichment product page at www.nugen.com and may be entered into the sample sheet for the Illumina sequencing run prior to the analysis. The additional 6 random bases can be entered as 'N's.

[00228] VIII. Optional Duplicate Determination

[00229] The index reads can consist of eight bases for library identification followed immediately by six random bases. By increasing the indexing read length settings on the sequencer to 14 bases and using the NuGEN Technologies Duplicate Marking Software tool, you can be able to accurately identify PCR duplicates in the data set. The 8-base barcodes followed by NNNNNN can be entered into the sample sheet to enable proper multiplex library parsing. The Duplicate Marking tool and instructions for use can be obtained by contacting NuGEN Technical Support (techserv@nugen.com) or from the NuGEN website, www.nugen.com.

[00230] IX. Barcode Sequences and Guidelines for Multiplex Experiments

[00231] The Ovation Target Enrichment System 32-reaction kit can contain 32 different barcodes and the 96-reaction kit contains 96 different barcodes that can be used for multiplex sequencing to interrogate several independent libraries on a single lane of the Illumina NGS platforms. Barcode sequences and multiplex guidelines for adaptors used in Ovation Target Enrichment Systems can be found in Table 9. All barcode sequences can be separated by an edit distance of three. For further details on the barcode design strategy, please refer to Faircloth BC, Glenn TC (2012) Not All Sequence Tags Are Created Equal: Designing and Validating Sequence Identification Tags Robust to Indels. PLoS ONE 7(8): e42543. doi:10.1371/journal.pone.0042543, which is incorporated herein by reference in its entirety.

Table 9. The 32-reaction kit can contain the 32 barcodes shown in plate positions A01–H04. The 96-reaction kit can contain 96 barcodes in plate positions as shown. Barcodes can be color-balanced in sets of 8 by column on the plate. (e.g., A01–H01 is a color-balanced set of barcodes.)

PLATE POSITION	BARCODE SEQUENCE	PLATE POSITION	BARCODE SEQUENCE	PLATE POSITION	BARCODE SEQUENCE
A01	CACGTCTA	A05	ATGGCGAT	A09	CAAGAAGC
B01	AGCTAGTG	B05	AACGCCTT	B09	CTGATGAG
C01	ACTATCGC	C05	GTAAGGTG	C09	CTCGAACA
D01	GCGTATCA	D05	TGTCGACT	D09	TCGACAAG
E01	ACTCTCCA	E05	ACTCTGAG	E09	AGTGCATC
F01	CGTCCATT	F05	GATGGAGT	F09	TGGCTACA
G01	AGCCGTAA	G05	CTAGCTCA	G09	GCACACAA
H01	GAGTAGAG	H05	CTGTACCA	H09	GCATAGTC
A02	ACGTCGTT	A06	CCTGTCAA	A10	AACACGCT
B02	GTCCTGTT	B06	GGTCGTAT	B10	TTCACGGA
C02	AGAAGCCT	C06	CGCTGATA	C10	TGCTGTGA
D02	GAAGATCC	D06	TAGCTTCC	D10	CCTCGAAT
E02	TAGCTGAG	E06	CAAGTCGT	E10	TGTGGCTT
F02	ACGTCCAA	F06	GTCTCATC	F10	CTGTGGTA
G02	CACACATC	G06	ACCAAGCA	G10	TCACTCGA
H02	CGGATCAA	H06	AGTCAGGT	H10	ACTCCTAC
A03	TCAGCCTT	A07	TATCGCGA	A11	CCACAACA
B03	AAGGCTCT	B07	TAGCAGGA	B11	CCGCTTAA
C03	TGTTCCGT	C07	AGAAGGAC	C11	GTGGTATG
D03	GGAATGTC	D07	TGAGCTGT	D11	GGTGTACA
E03	CATCCAAG	E07	CAGAGTGA	E11	TCTAGGAG
F03	GTCAACAG	F07	AGGTTCCT	F11	TGGAAGCA
G03	TCGCTATC	G07	AGACCTTG	G11	AACACCAC
H03	AGCCTATC	H07	CTTCCTTC	H11	CATACGGA
A04	TCGGATTC	A08	CAGGTTCA	A12	CTCTCAGA
B04	CGGAGTAT	B08	ACTGGTGT	B12	AAGCTGGT
C04	GAACCTTC	C08	GGATTCAC	C12	GTGTCCTT
D04	AGAGGATG	D08	CACGATTC	D12	AGGTCTGT
E04	ACGCTTCT	E08	AGACATGC	E12	CATTCGTC
F04	CACAGGAA	F08	GACACAGT	F12	CTCACCAA
G04	ACGAATCC	G08	CCAGTTGA	G12	GACTACGA
H04	CCTTCCAT	H08	CATGGATC	H12	ATACGCAG

[00232] X Other considerations.

[00233] The custom kits can be available as 32 or 96 reaction size with the corresponding number of unique barcodes.

[00234] The kit can have a shelf life of 6 months from receipt.

[00235] A member of a Technical Services team can help to obtain the correct information for a custom design. RefSeq IDs or genomic coordinates will generally suffice.

[00236] If the targets are provided in the appropriate format, it can take no more than a couple of days for the design to be complete. For non-human sequence this process can take longer in order for the correct genome sequence to be identified and added to the design pipeline. The estimated turnaround time can be 3–4 weeks between obtaining customer approval on a finalized design and shipping the custom kit.

[00237] The genomic coordinates and 50 bases of sequence information are also provided that includes the probe.

[00238] Covaris focused acoustic shearing can be used to shear the DNA, and other methods (e.g., fragmentase) can work.

[00239] The Ovation Target Enrichment System features a simple add and incubate workflow, so it can be automated.

[00240] FFPE material can works well in the system and can work best with a probe design optimized to accommodate the smaller expected fragment sizes of FFPE genomic DNA. Because the DNA is modified and degraded, starting with a larger amount of input material (more than 100 ng) can be preferred.

[00241] In some cases, cell-free DNA can require an optimized probe design that addresses smaller fragments, similar to FFPE material. If there is sufficient plasma DNA as input, this may be possible.

[00242] In some cases, genomic DNA from a single cell can require a preamplification step prior to the library prep and enrichment.

[00243] Probes can be designed for regions that target the differences between genes and pseudogenes.

[00244] In design, probes can be placed outside the target region. However, for larger target regions a probe can be tiled within the target region in order to make sure that longer exons are accurately represented.

[00245] Reads up to 150 bp can work. Longer reads can be performed and can require a larger average fragment size and the read may extend into the probe.

[00246] Single end reads can be sufficient and advised for many applications. Paired end sequencing can be performed to look specifically for rearrangements like inversions and translocations.

[00247] Due to the probe extension technology employed by the target enrichment system, the first 65 bases of reverse read sequence are probe-derived and can be trimmed prior to alignment.

Example 6 – Enrichment cDNA from 100ng of UHR RNA

[00248] 100ng of universal human reference (UHR) RNA was used as input into the Ovation target enrichment cDNA module and libraries were prepared with the Gene Fusion Panel of probes following the manufacturers protocol. Sequencing libraries were quantitated by Bioanalyzer (Agilent Technologies, USA) and Kapa Library Quantification Kit (KAPA Biosytems, USA). Libraries were sequenced on the Illumina Hiseq as paired end

reads, 100 bases in length. Resulting sequences were analyzed for the presence of gene fusions with SOAPFuse v1.22 software using default settings (//soap.genomics.org.cn/soapfuse.html). Standard RNAseq data were obtained for comparison at //sra.dnanexus.com/studies/SRP028705. The targeted reads and untargeted reads are shown in Figure 9.

Example 7 – Selective target enrichment using a priming reaction to append forward and reverse adaptors

[00249] In this example, both forward and reverse adaptors are appended to a nucleic acid sequence by priming reactions. As depicted in FIG. 10, nucleic acids are fragmented to generate a plurality of fragmented nucleic acids **1001**. The nucleic acid fragments are denatured and adaptor 1 is appended to the nucleic acid fragments by annealing primer A with adaptor 1 at the 5' end to the nucleic acid fragments **1003, 1005**. Primer A can anneal to either strand of the nucleic acid fragments **1003, 1005**. Primer A is extended with a polymerase to generate an oligonucleotide extension product **1003, 1005**. After extension, primer B is added **1007, 1009** that anneals to a sequence on the oligonucleotide extension product. Primer B has an adaptor 2 sequence at the 5' end and a sequence that anneals to the oligonucleotide extension product **1007, 1009**. Primer B is extended with a polymerase **1007, 1009** to generate a second oligonucleotide extension product with the complement of adaptor 1 at a first end, the target sequence of interest, and adaptor 2 at a second end. Primer A may be random when primer B is target specific or primer A may be target specific when primer B is random. The second oligonucleotide extension product is then amplified using primers that anneal to the complement of adaptors 1 and 2 **1011, 1013**.

Example 8 – Selective target enrichment using a reverse protocol to append forward and reverse adaptors

[00250] As depicted in FIG. 11, target-specific sequences can be enriched using a protocol that appends a first adaptor to a nucleic acid fragment by annealing and extending a primer and then ligating a second adaptor to the opposite end of the nucleic acid fragment. Nucleic acids are fragmented to generate fragmented nucleic acid sequences **1101**. The first adaptor is appended to the nucleic acids by annealing a primer comprising a target-specific portion at the 3' end and adaptor 1 at the 5' end **1103**. The target-specific portion of the primer will anneal to a target-specific sequence in the nucleic acid fragments **1103**. The primer is extended by addition of a polymerase **1103** generating an oligonucleotide extension product **1103** comprising a sequence complementary to the target-specific sequence and adaptor 1 at a first end **1103**. Double-stranded adaptor 2 is ligated to the double-stranded blunt-end extension product **1105**. The double-stranded extension product is denatured and a primer comprising adaptor 2 sequence is annealed and extended **1107**. The target library is amplified with primers that are specific for adaptor 1 and adaptor 2 **1109**.

Example 9 – Selective target enrichment using transposon-tagging

[00251] As depicted in FIG. 12, target-specific sequences can be enriched using transposon-tagging. In this example, a transposome complex containing a transposase enzyme and a transposon sequence are combined with a nucleic acid **1201**. The transposome complex randomly binds to the nucleic acid and simultaneously fragments the nucleic acid and tags the nucleic acid with the transposon sequence **1203 (TS1 and TS2)**. The known transposon sequence then acts as a primer-binding site for a primer that has an adaptor 1 sequence at the 5' end

and a sequence complementary to the transposon sequence at the 3' end **1205**. The primer is annealed to the nucleic acid fragment and extended with a polymerase **1205** generating an oligonucleotide extension product with a sequence complementary to the nucleic acid fragment and adaptor 1 at a first end **1207**. A target-specific primer that contains a sequence that anneals to the complement of a target nucleic acid sequence at the 3' end and an adaptor 2 at the 5' end is then annealed to the oligonucleotide extension product **1207** and extended by a polymerase to generate a second oligonucleotide extension product with the complement of adaptor 1 at a first end **(C1)**, the target nucleic acid sequence and the second adaptor at a second end. The target nucleic acid sequence is then amplified with primers that anneal to the complement of adaptor 1 **1209** and the complement of adaptor 2 **1211** thereby generating a double-stranded nucleic acid molecule with adaptor 1 at one end and adaptor 2 at the other end **1213**.

Example 10 – Selective target enrichment using transposon-tagging

[00252] As depicted in FIG. 13, target-specific sequences can be enriched using transposon-tagging. In this example, a transposome complex containing a transposase enzyme and a transposon sequence are combined with a nucleic acid **1301**. The transposome complex randomly binds to the nucleic acid and simultaneously fragments the nucleic acid and tags the nucleic acid with the transposon sequence **1303 (TS1 and TS2)**. The tagged fragments are denatured and a primer that contains a sequence that anneals to the complement of a target nucleic acid sequence at the 3' end and an adaptor 1 sequence at the 5' end is annealed to a nucleic acid fragment **1305, 1307** and extended by a polymerase to generate an oligonucleotide extension product. The double-stranded oligonucleotide extension product is then denatured and the target nucleic acid sequence is amplified using primers that anneal to the complement of adaptor 1 and the transposon sequence **1309, 1311** thereby generating a double-stranded nucleic acid molecule with adaptor 1 at one end and the transposon sequence at the other end **1313, 1315**.

Example 11 – Selective target enrichment using a priming reaction to append forward and reverse adaptors

[00253] As depicted in FIG. 14, both forward and reverse adaptors can be appended to a nucleic acid sequence by priming reactions. Double-stranded nucleic acids containing a nucleic acid sequence of interest are denatured to produce single-stranded nucleic acid molecules. Primers containing random sequence (**RS**) at the 3' end and adaptor 1 sequence (**A1**) at the 5' end are annealed to the single-stranded nucleic acid molecules **1401**. The random primers are composite primers (e.g., the random sequence includes DNA and adaptor 1 sequence includes RNA). The primers are extended with a polymerase to generate a first oligonucleotide extension product **1403**. The first oligonucleotide extension product is denatured and target-specific primers (**TS**) are annealed to the first oligonucleotide extension product **1405**, wherein the target-specific primers contain a 3' end that specifically anneals to the nucleic acid sequence of interest within the first oligonucleotide extension product and a 5' end that contains adaptor 2 sequence (**A2**). The target-specific primers are extended with a polymerase to generate a second oligonucleotide extension product forming an RNA/DNA heteroduplex between adaptor 1 (RNA) and the complement of adaptor 1 (DNA) **1407**. The adaptor 1 sequence is digested with RNase H **1409** to generate a pool

of tagged nucleic acid molecules containing the nucleic acid sequence of interest **1411**. The nucleic acid sequence of interest is then amplified and enriched with primers (**A1** and **A2**) that bind to the complement of the adaptor 2 sequence (**A2'**) and the complement of the adaptor 1 sequence (**A1'**) **1411, 1413**.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a) annealing an oligonucleotide to the nucleic acid sequence of interest in a nucleic acid fragment, wherein the oligonucleotide comprises a 3' portion complementary to the nucleic acid sequence of interest and a 5' portion comprising a first adaptor sequence;

b) extending the oligonucleotide annealed to the nucleic acid sequence of interest in the nucleic acid fragment with a polymerase, thereby generating an oligonucleotide extension product comprising sequence complementary to the nucleic acid sequence of interest and the first adaptor sequence at a first end;

c) ligating a sequence complementary to a second adaptor sequence to the oligonucleotide extension product, thereby generating an oligonucleotide extension product comprising the first adaptor sequence at the first end, the sequence complementary to the nucleic acid sequence of interest, and a sequence complementary to the second adaptor sequence at a second end; and

d) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest, thereby generating an enriched nucleic acid sequence of interest.

2. The method of claim 1, further comprising, before step a), fragmenting a nucleic acid in the sample, thereby generating the nucleic acid fragment of step a) comprising the sequence of interest.

3. The method of claim 1, wherein the nucleic acid fragment is double-stranded prior to step a).

4. The method of claim 1, wherein the nucleic acid fragment comprises deoxyribonucleic acid (DNA).

5. The method of claim 4, wherein the DNA is genomic DNA.

6. The method of claim 4, wherein the DNA is cDNA.

7. The method of claim 3, further comprising denaturing the double-stranded nucleic acid fragment prior to step b), wherein the nucleic acid fragment comprising the nucleic acid sequence of interest in step b) is single-stranded prior to annealing to the oligonucleotide.

8. The method of claim 1, wherein the 3' portion of the oligonucleotide comprises a random sequence.

9. The method of claim 1, wherein the 3' portion of the oligonucleotide comprises a sequence designed to anneal to the nucleic acid sequence of interest.

10. The method of claim 1, wherein ligating the sequence complementary to the second adaptor sequence comprises ligating the sequence complementary to the second adaptor sequence to a 3' end of the oligonucleotide extension product.

11. The method of claim 10, wherein the second adaptor sequence is within a strand of a double-stranded adaptor.

12. The method of claim 1, wherein the first adaptor sequence and the second adaptor sequence are distinct.

13. The method of claim 1, wherein the first adaptor sequence and/or the second adaptor sequence comprise barcode sequence.

14. The method of claim 1, wherein the first adaptor sequence and/or the second adaptor sequence comprise sequence that can anneal to a sequencing primer, or complement of a sequence that can anneal to a sequencing primer.

15. The method of claim 1, wherein the polymerase is a DNA polymerase.

16. The method of claim 1, further comprising:

(i) annealing a plurality of oligonucleotides to different nucleic acid sequences of interest in nucleic acid fragments, wherein each of the plurality of oligonucleotides comprises a 3' portion complementary to a nucleic acid sequence of interest among the nucleic acid sequences of interest and a 5' portion comprising a first adaptor sequence;

(ii) extending the plurality of oligonucleotides annealed to the different nucleic acid sequences of interest in the nucleic acid fragments with a polymerase, thereby generating a plurality of oligonucleotide extension products comprising sequence complementary to the nucleic acid sequences of interest and the first adaptor sequence at a first end;

(iii) ligating a sequence complementary to a second adaptor sequence to the plurality of oligonucleotide extension products, thereby generating a plurality of oligonucleotide extension products comprising the first adaptor sequence at the first end, the sequence complementary to the nucleic acid sequence of interest, and a sequence complementary to the second adaptor sequence at a second end; and

(iv) amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequences of interest, thereby generating enriched nucleic acid sequences of interest.

17. The method of claim 1, wherein the amplifying in step d) comprises a polymerase chain reaction.

18. The method of claim 1, further comprising enriching a plurality of nucleic acid sequences of interest using the steps of claim 1.

19. The method of claim 18, further comprising sequencing the plurality of enriched nucleic acid sequences of interest.

20. The method of claim 1, wherein the nucleic acid sequence of interest comprises a gene.

21. The method of claim 20, wherein the gene is a cancer gene.

22. The method of claim 1, further comprising sequencing the enriched nucleic acid sequence of interest.

23. The method of claim 22, wherein the sequencing comprises use of a reversible-dye terminator.

24. A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a) combining a transposome comprising a transposase and a transposon sequence with a nucleic acid in the sample, thereby generating a nucleic acid fragment comprising the nucleic acid sequence of interest and a transposon sequence appended to a first end of the nucleic acid fragment;

b) annealing a first oligonucleotide to the nucleic acid fragment comprising the transposon sequence at the first end, wherein the first oligonucleotide comprises a 3' portion that is complementary to the transposon sequence and a 5' portion comprising a first adaptor sequence;

c) extending the first oligonucleotide annealed to the transposon sequence with a polymerase, thereby generating a first oligonucleotide extension product comprising sequence complementary to the nucleic acid fragment and a first adaptor sequence at the first end;

d) annealing a second oligonucleotide to a sequence complementary to the nucleic acid sequence of interest in the nucleic acid fragment comprising the first adaptor sequence at the first end, wherein the second oligonucleotide comprises a 3' portion that anneals to the complement of the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence;

e) extending the second oligonucleotide annealed to the complement of the nucleic acid sequence of interest in the nucleic acid fragment comprising the first adaptor sequence at the first end with a polymerase, thereby generating a second oligonucleotide extension product comprising sequence complementary to the first adaptor sequence at the first end, the nucleic acid sequence of interest, and the second adaptor sequence at the second end; and

f) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest.

25. The method of claim 24, wherein the transposase is a Tn5 transposase.

26. The method of claim 24, wherein the nucleic acid fragment of step a) is double stranded.

27. The method of claim 24, further comprising denaturing the double stranded nucleic acid fragment prior to step b), wherein the nucleic acid fragment in step b) is single-stranded prior to annealing to the oligonucleotide.

28. The method of claim 24, wherein the first oligonucleotide extension product is annealed to the nucleic acid fragment after the extension step c).

29. The method of claim 28, further comprising denaturing the first oligonucleotide extension product from the nucleic acid fragment prior to step d), whereby the first oligonucleotide extension product in step d) is single stranded.

30. The method of claim 24, wherein the 3' portion of the second oligonucleotide comprises a random sequence.

31. The method of claim 24, wherein the 3' portion of the second oligonucleotide comprises a sequence designed to anneal to the nucleic acid sequence of interest.

32. The method of claim 24, wherein the first adaptor sequence and the second adaptor sequence are distinct.

33. The method of claim 24, wherein the first adaptor sequence and/or the second adaptor sequence comprise a barcode sequence.

34. The method of claim 24, wherein the first adaptor sequence and/or the second adaptor sequence comprise sequence designed to anneal to a sequencing primer.

35. The method of claim 24, wherein the polymerase is a DNA polymerase.

36. The method of claim 24, further comprising:

(i) using steps a) –c) to produce a plurality of first oligonucleotide extension products comprising sequence complementary to nucleic acid fragments and a first adaptor sequence at the first end; annealing each of a plurality of second oligonucleotides to a sequence complementary to a nucleic acid sequence of interest in the nucleic acid fragments comprising the first adaptor sequence at the first end, wherein the plurality of second oligonucleotides comprise a 3' portion that anneals to a complement of a nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence;

(ii) extending the plurality of second oligonucleotides annealed to a complement of a nucleic acid sequence of interest in a nucleic acid fragment comprising the first adaptor sequence at the first end with a polymerase, thereby generating a plurality of second oligonucleotide extension products comprising sequence complementary to the first adaptor sequence at the first end, a nucleic acid sequence of interest, and the second adaptor sequence at the second end; and

(iii) amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest.

37. The method of claim 24, wherein the amplification of step f) comprises polymerase chain reaction.

38. The method of claim 24, wherein the annealing of step d) comprises annealing a plurality of second oligonucleotides to a plurality of nucleic acid fragments, wherein the plurality of second oligonucleotides comprise a 3' portion that anneals to the complement of a different nucleic acid sequence of interest in the nucleic acid fragments and a 5' portion comprising the second adaptor sequence.

39. The method of claim 24, further comprising enriching a plurality of nucleic acid sequences of interest using the steps of claim 24.

40. The method of claim 39, further comprising sequencing the plurality of enriched nucleic acid sequences of interest.

41. The method of claim 40, wherein the sequencing comprises use of a reversible-dye terminator.

42. A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a) combining a transposome comprising a transposase and a transposon sequence with a nucleic acid in the sample, thereby generating a nucleic acid fragment comprising the nucleic acid sequence of interest and a transposon sequence appended to a first end of the nucleic acid fragment;

b) annealing an oligonucleotide to the nucleic acid fragment comprising the transposon sequence at the first end, wherein the oligonucleotide comprises a 3' portion that is complementary to the nucleic acid sequence of interest in the nucleic acid fragment and a 5' portion comprising a first adaptor sequence;

c) extending the oligonucleotide annealed to the nucleic acid sequence of interest with a polymerase, thereby generating an oligonucleotide extension product comprising sequence complementary to the transposon sequence at a first end, sequence complementary to the nucleic acid sequence of interest, and a first adaptor sequence at a second end; and

d) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the transposon sequence and a second primer that anneals to the complement of the first adaptor sequence to enrich for the nucleic acid sequence of interest.

43. The method of claim 42, wherein the transposase is a Tn5 transposase.

44. The method of claim 42, wherein the nucleic acid fragment of step a) is double stranded.

45. The method of claim 42, further comprising denaturing the double stranded nucleic acid fragment prior to step b), wherein the nucleic acid fragment in step b) is single-stranded prior to annealing to the oligonucleotide.

46. The method of claim 42, wherein the oligonucleotide extension product is annealed to the nucleic acid fragment after the extension of step c).

47. The method of claim 46, further comprising denaturing the oligonucleotide extension product from the nucleic acid fragment prior to step d), whereby the oligonucleotide extension product in step d) is single stranded.

48. The method of claim 42, wherein the first adaptor sequence comprises a barcode sequence.

49. The method of claim 42, wherein the first adaptor sequence comprises sequence designed to anneal to a sequencing primer.

50. The method of claim 42, wherein the polymerase is a DNA polymerase.

51. The method of claim 42, further comprising:

(i) generating a plurality of nucleic acid fragments comprising a nucleic acid sequence of interest and a transposon sequence appended to a first end of the nucleic acid fragment;
annealing a plurality of oligonucleotides to the nucleic acid fragments comprising the transposon sequence at the first end, wherein each of the plurality of oligonucleotides comprises a 3' portion that is complementary to the nucleic acid sequence of interest in a nucleic acid fragment and a 5' portion comprising a first adaptor sequence;

(ii) extending the plurality of oligonucleotides annealed to the nucleic acid sequences of interest with a polymerase, thereby generating a plurality of oligonucleotide extension products comprising sequence

complementary to the transposon sequence at a first end, sequence complementary to the nucleic acid sequence of interest, and a first adaptor sequence at a second end; and

(iii) amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the transposon sequence and a second primer that anneals to the complement of the first adaptor sequence to enrich for the nucleic acid sequences of interest.

52. The method of claim 42, wherein the amplification of step f) comprises polymerase chain reaction.

53. The method of claim 42, further comprising enriching a plurality of nucleic acid sequences of interest using the steps of claim 42.

54. The method of claim 42, further comprising sequencing the plurality of enriched nucleic acid sequences of interest.

55. The method of claim 54, wherein the sequencing comprises use of a reversible-dye terminator.

56. A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a) annealing a first oligonucleotide to a nucleic acid fragment comprising the nucleic acid sequence of interest, wherein the first oligonucleotide comprises a 3' portion complementary to a sequence in the nucleic acid fragment and a 5' portion comprising a first adaptor sequence;

b) extending the first oligonucleotide annealed to the nucleic acid fragment comprising the nucleic acid sequence of interest with a polymerase, thereby generating a first oligonucleotide extension product comprising sequence complementary to the nucleic acid sequence of interest and the first adaptor sequence at a first end;

c) annealing a second oligonucleotide to the complement of the nucleic acid sequence of interest in the first oligonucleotide extension product, wherein the second oligonucleotide comprises a 3' portion complementary to the sequence complementary to the nucleic acid sequence of interest and a 5' portion comprising a second adaptor sequence;

d) extending the second oligonucleotide annealed to the complement of the nucleic acid sequence of interest with a polymerase, thereby generating a second oligonucleotide extension product comprising a sequence complementary to the first adaptor sequence at a first end, the nucleic acid sequence of interest, and the second adaptor sequence at a second end; and

e) amplifying the nucleic acid sequence of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest, thereby generating an enriched nucleic acid sequence of interest.

57. The method of claim 56, further comprising, before step a), fragmenting a nucleic acid in the sample, thereby generating the nucleic acid fragment of step a) comprising the nucleic acid sequence of interest.

58. The method of claim 56, wherein the nucleic acid fragment is double-stranded prior to step a).

59. The method of claim 56, wherein the nucleic acid fragment comprises deoxyribonucleic acid (DNA) or ribonucleic acid (RNA).

60. The method of claim 59, wherein the DNA is genomic DNA.

61. The method of claim 59, wherein the DNA is cDNA.

62. The method of claim 56, wherein the 3' portion of the first oligonucleotide comprises DNA and the 5' portion of the first oligonucleotide comprises RNA.

63. The method of claim 58, further comprising denaturing the double-stranded nucleic acid fragment prior to step a), wherein the nucleic acid fragment comprising the nucleic acid sequence of interest in step a) is single-stranded prior to annealing to the first oligonucleotide.

64. The method of claim 56, wherein the 3' portion of the first oligonucleotide comprises a random sequence.

65. The method of claim 56, wherein the 3' portion of the second oligonucleotide comprises a sequence designed to anneal to the nucleic acid sequence of interest.

66. The method of claim 56, wherein the first adaptor sequence and the second adaptor sequence are distinct.

67. The method of claim 56, wherein the first adaptor sequence and/or the second adaptor sequence comprise barcode sequence.

68. The method of claim 56, wherein the first adaptor sequence and/or the second adaptor sequence comprise sequence that can anneal to a sequencing primer, or complement of a sequence that can anneal to a sequencing primer.

69. The method of claim 56, wherein the polymerase is a DNA polymerase.

70. The method of claim 56, further comprising:

(i) annealing a plurality of first oligonucleotides to different nucleic acid fragments comprising different nucleic acid sequences of interest, wherein each of the plurality of first oligonucleotides comprises a 3' portion complementary to a sequence in the nucleic acid fragment among the nucleic acid fragments and a 5' portion comprising a first adaptor sequence;

(ii) extending the plurality of first oligonucleotides annealed to the different nucleic acid fragments comprising different nucleic acid sequences of interest with a polymerase, thereby generating a plurality of first oligonucleotide extension products each comprising sequence complementary to the nucleic acid sequences of interest and the first adaptor sequence at a first end;

(iii) annealing a plurality of second oligonucleotides to the complement of the nucleic acid sequences of interest in the plurality of first oligonucleotide extension products, wherein each of the plurality of second oligonucleotides comprises a 3' portion complementary to the sequence complementary to the nucleic acid sequence of interest among the nucleic acid sequences of interest and a 5' portion comprising a second adaptor sequence;

(iv) extending the plurality of second oligonucleotides annealed to the complement of the nucleic acid sequences of interest in the first oligonucleotide extension products, thereby generating a plurality of second oligonucleotide extension products each comprising a sequence complementary to the first adaptor sequence at a first end, the nucleic acid sequence of interest among the nucleic acid sequences of interest, and the second adaptor sequence at a second end; and

(v) amplifying the nucleic acid sequences of interest using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequences of interest, thereby generating enriched nucleic acid sequences of interest.

71. The method of claim 56, further comprising prior to step e), digesting the second oligonucleotide extension product with RNase H.

72. The method of claim 56, wherein the amplifying in step e) comprises a polymerase chain reaction.

73. The method of claim 56, further comprising enriching a plurality of nucleic acid sequences of interest using the steps of claim 56.

74. The method of claim 73, further comprising sequencing the plurality of enriched nucleic acid sequences of interest.

75. The method of claim 56, wherein the nucleic acid sequence of interest comprises a gene.

76. The method of claim 75, wherein the gene is a cancer gene.

77. The method of claim 56, further comprising sequencing the enriched nucleic acid sequence of interest.

78. The method of claim 77, wherein the sequencing comprises use of a reversible-dye terminator.

79. A method for enriching for a nucleic acid sequence of interest in a sample comprising nucleic acids, the method comprising:

a) fragmenting the nucleic acids thereby generating nucleic acid fragments, wherein the nucleic acid fragments comprise the nucleic acid sequence of interest;

b) ligating a first adaptor sequence to a 5' end of each of the nucleic acid fragments thereby generating nucleic acid fragments comprising the first adaptor sequence at the 5' end;

c) annealing one or more oligonucleotides to the nucleic acid fragments comprising the first adaptor sequence at the 5' end, wherein each of the one or more oligonucleotides comprise a 3' portion that is complementary to a nucleic acid sequence of interest present in one or more of the nucleic acid fragments, and a 5' portion comprising a second adapter sequence;

d) extending the one or more oligonucleotides with a polymerase thereby generating one or more oligonucleotide extension products comprising a sequence complementary to the first adaptor sequence at a first end, a sequence complementary to the nucleic acid sequence of interest and the second adaptor sequence at a second end; and

e) amplifying the one or more oligonucleotide extension products using a first primer that anneals to the complement of the first adaptor sequence and a second primer that anneals to the complement of the second adaptor sequence to enrich for the nucleic acid sequence of interest;

wherein the nucleic acid sequence of interest comprises a gene fusion or a gene rearrangement.

80. The method of claim 79, wherein the nucleic acids in the sample comprise RNA.

81. The method of claim 80, wherein the RNA is converted to double-stranded cDNA prior to step b).

82. The method of claim 79, wherein the gene fusion or gene rearrangement is in a cancer gene.

83. The method of claim 82, wherein the gene rearrangement is a translocation or an inversion.

84. The method of claim 82, wherein the one or more oligonucleotides anneals to a portion of at least one gene in the gene fusion.

85. The method of claim 84, wherein the at least one gene in the gene fusion is BCR, ABL, ALK, RET, or ROS1.

86. The method of claim 79, further comprising sequencing the nucleic acid sequence of interest after step e).

87. The method of claim 79, further comprising denaturing the nucleic acid fragments prior to step c), thereby generating single-stranded nucleic acid fragments comprising the first adaptor sequence at the 5' end.

88. The method of claim 79, wherein the first adaptor sequence, the second adaptor sequence or a combination thereof comprises barcode sequence.

ABSTRACT OF THE DISCLOSURE

Provided herein are methods, compositions and kits for targeted nucleic acid sequence enrichment in a nucleic acid sample and for high efficiency nucleic acid library generation for next generation sequencing (NGS). The methods, compositions and kits provided herein can be useful for the production and capture of amplification-ready, target-specific and strand-specific regions of interest from nucleic acid samples containing complex DNA.

FIG. 1

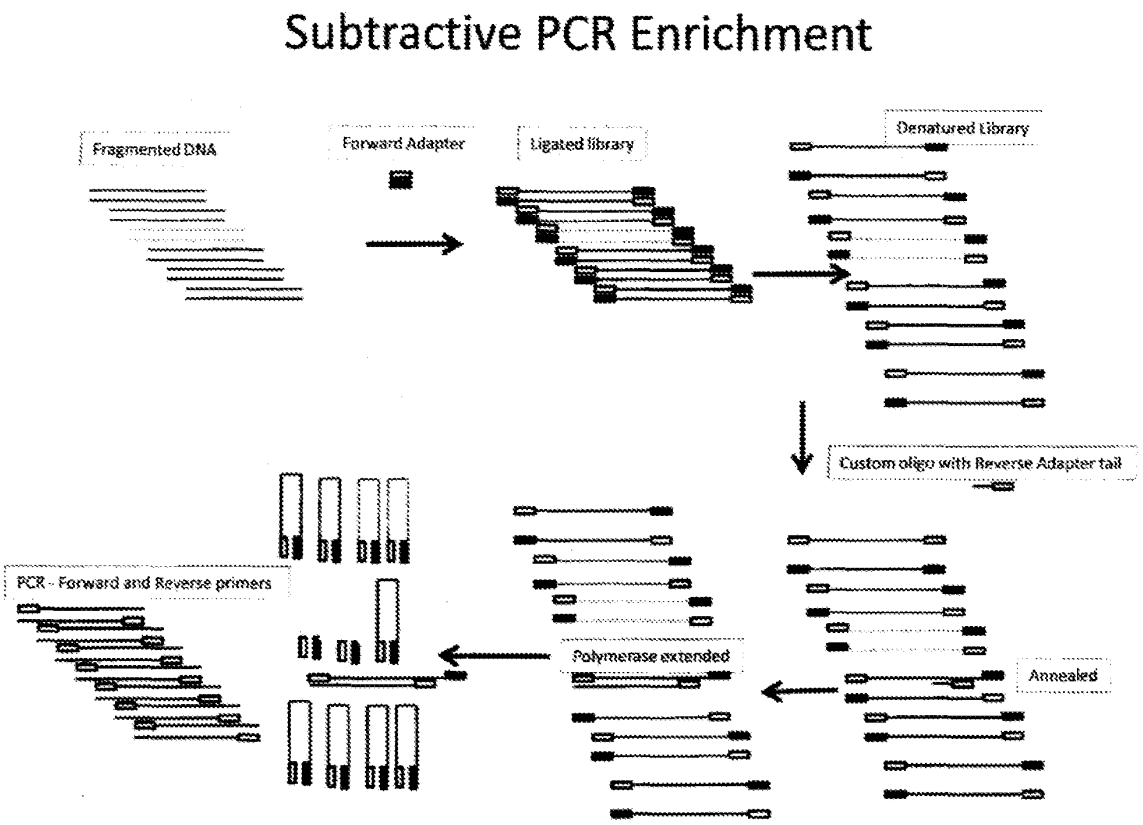

Subtractive PCR Enrichment
Fragmented DNA
Forward Adapter
Ligated library
Denatured Library
Custom oligo with Reverse Adapter tail
Annealed
Polymerase extended
PCR - Forward and Reverse primers

FIG. 2

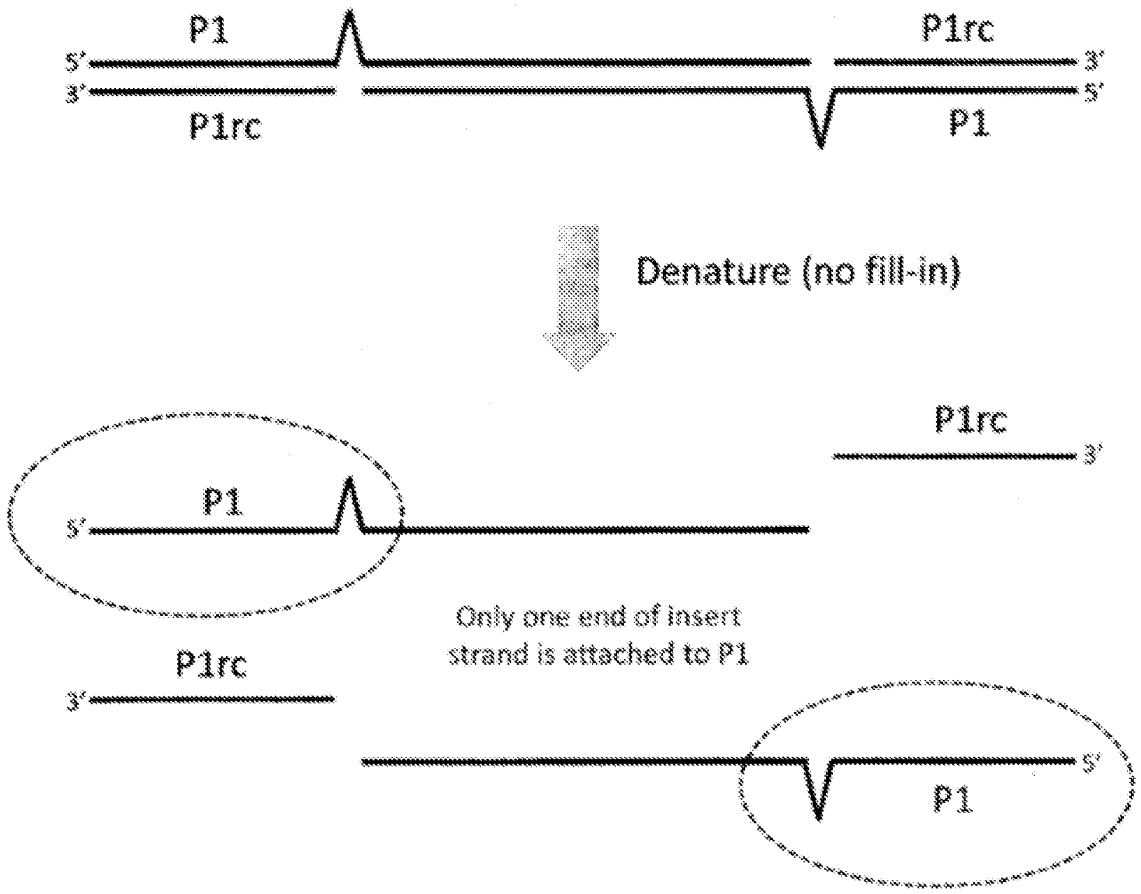

P1
P1rc
5'
3'
3'
5'
P1rc
P1
Denature (no fill-in)
P1rc
3'
P1
5'
Only one end of insert
strand is attached to P1
P1rc
3'
5'
P1

FIG. 3

Primer Directed Targeted Enrichment of Sequences of Interest in a Complex Library

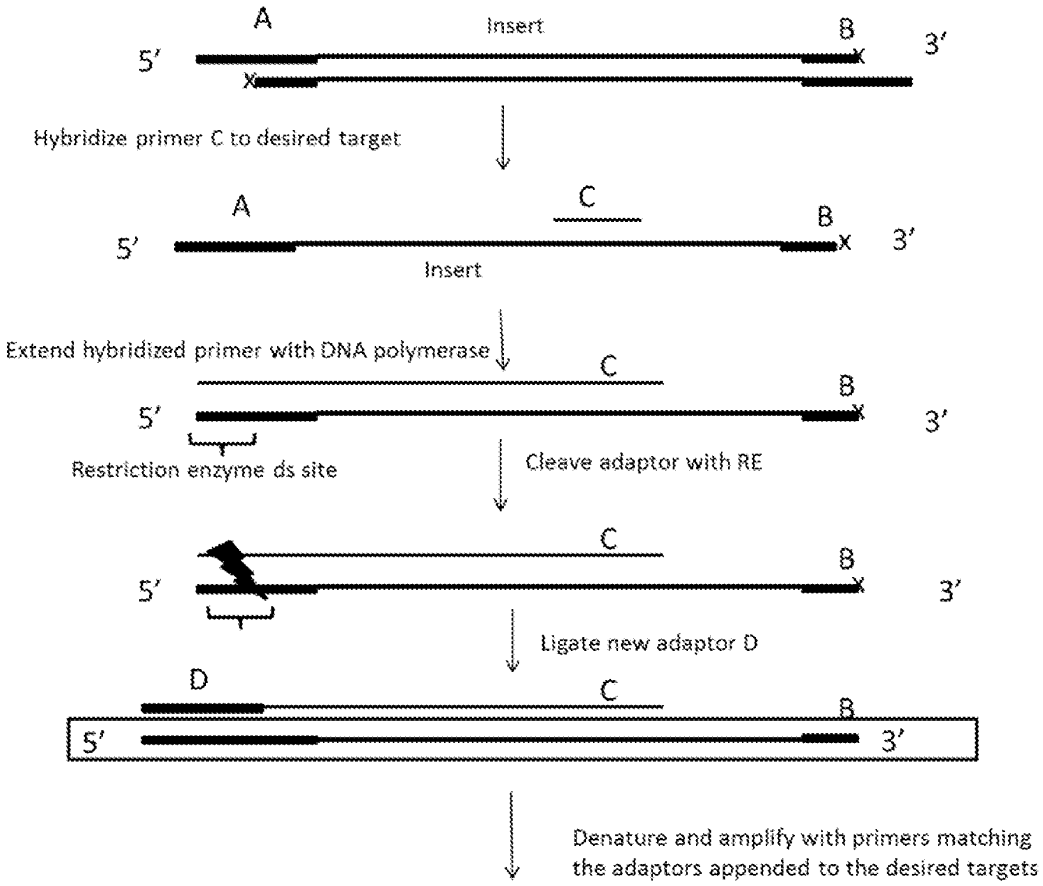


FIG. 4

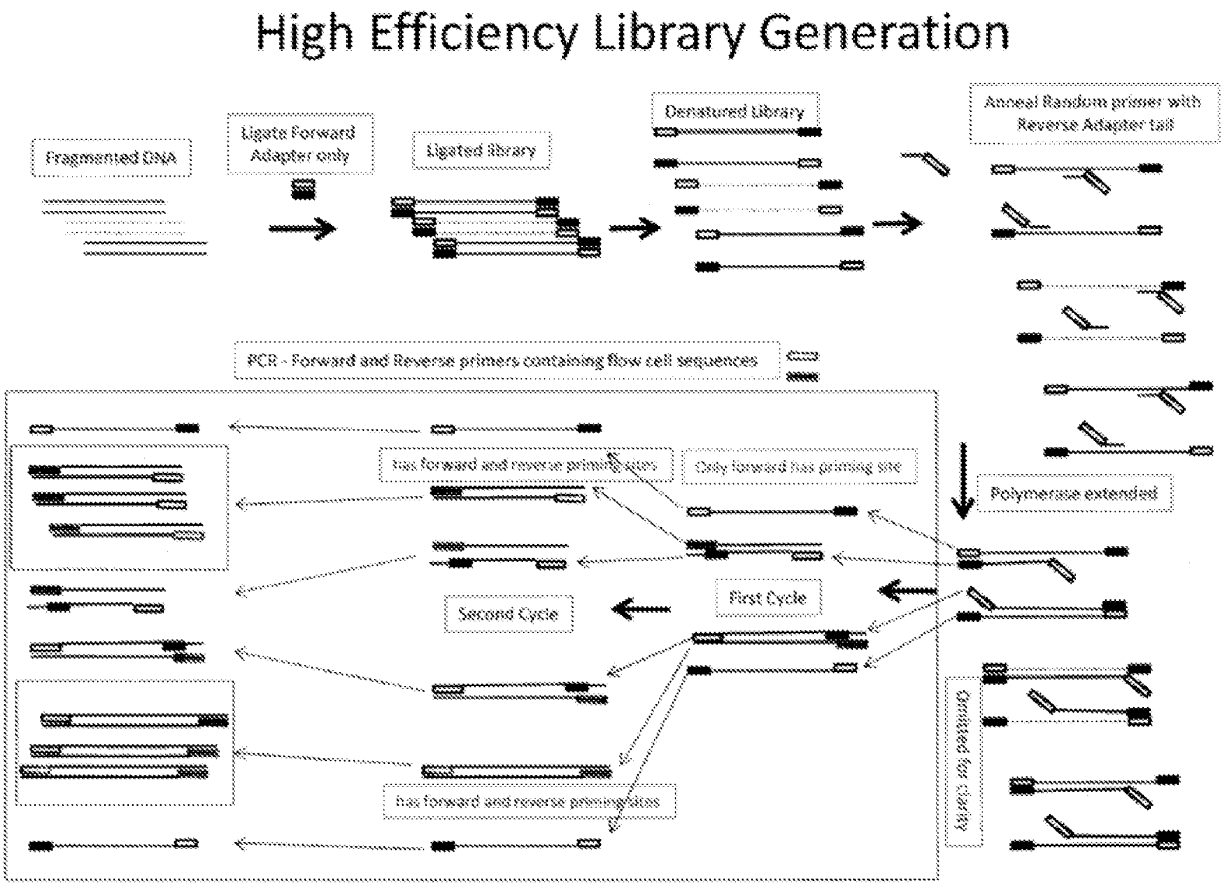

High Efficiency Library Generation
Fragmented DNA
Ligate Forward Adapter only
Ligated library
Denatured Library
Anneal Random primer with Reverse Adapter tail
PCR - Forward and Reverse primers containing flow cell sequences
Has forward and reverse priming sites
Only forward has priming site
Polymerase extended
First Cycle
Second Cycle
Has forward and reverse priming sites
Denatured for Cycle

FIG. 5

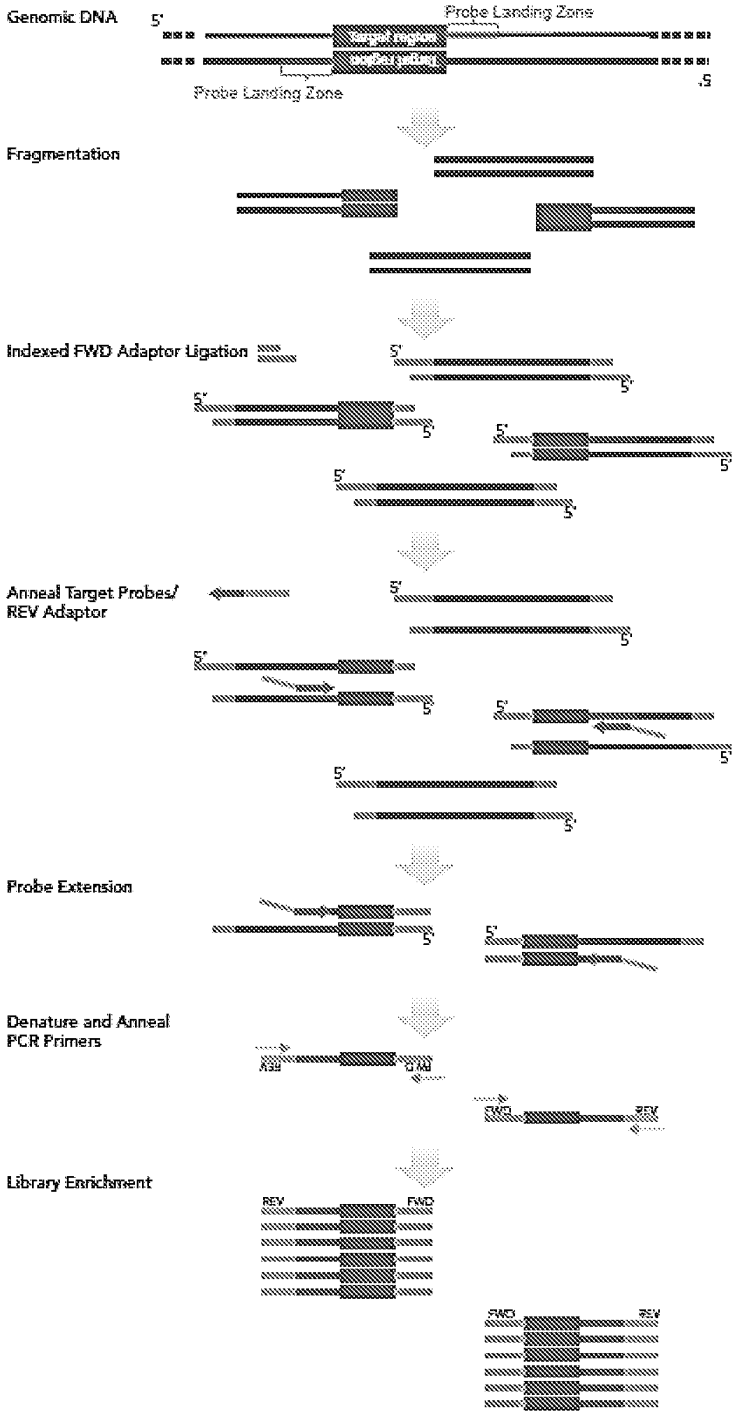
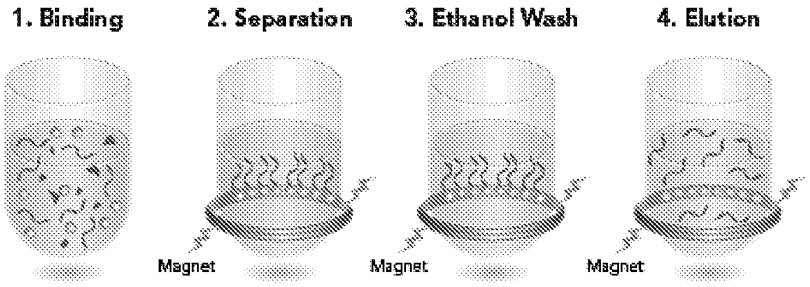

Genomic DNA
5'
Probe Landing Zone
Probe Landing Zone
3'
Fragmentation
Indexed FWD Adaptor Ligation
Anneal Target Probes/
REV Adaptor
Probe Extension
Denature and Anneal
PCR Primers
REV
FWD
Library Enrichment
REV
FWD
FWD
REV

FIG. 6



Reproduced from original picture from Agencourt/Beckman Coulter Genomics

FIG. 7

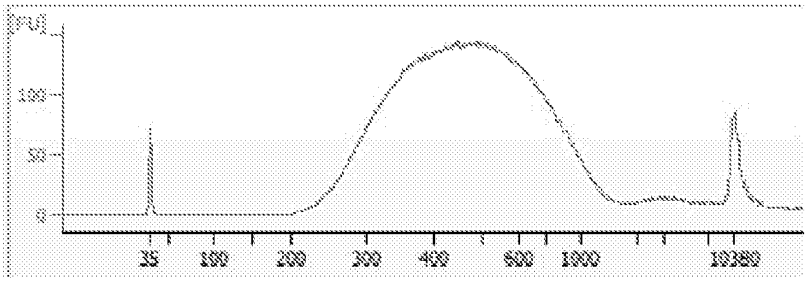


FIG. 8

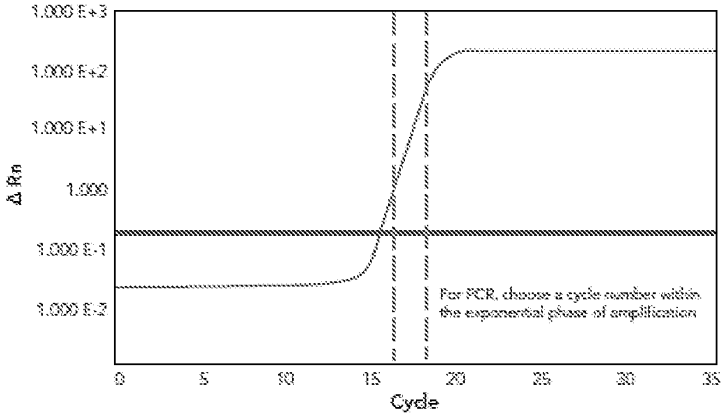

1.000 E+3
1.000 E+2
1.000 E+1
1.000
1.000 E-1
1.000 E-2
Δ Rn
For PCR, choose a cycle number within the exponential phase of amplification
0
5
10
15
20
25
30
35
Cycle

FIG. 9

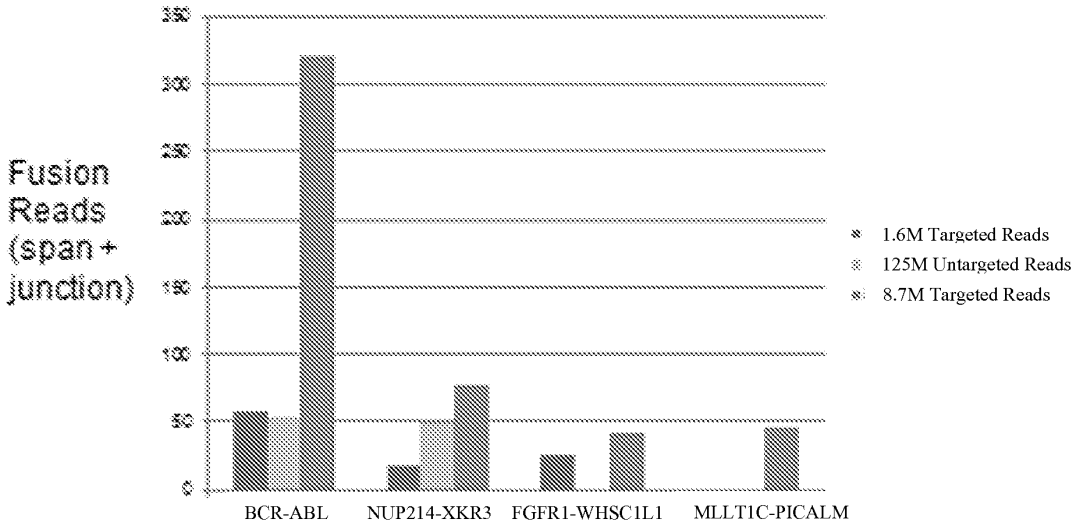

Fusion Reads (span + junction)
0
50
100
150
200
250
300
BCR-ABL
NUP214-XKR3
FGFR1-WHSC1L1
MLLT1C-PICALM
1.6M Targeted Reads
125M Untargeted Reads
8.7M Targeted Reads

FIG. 10

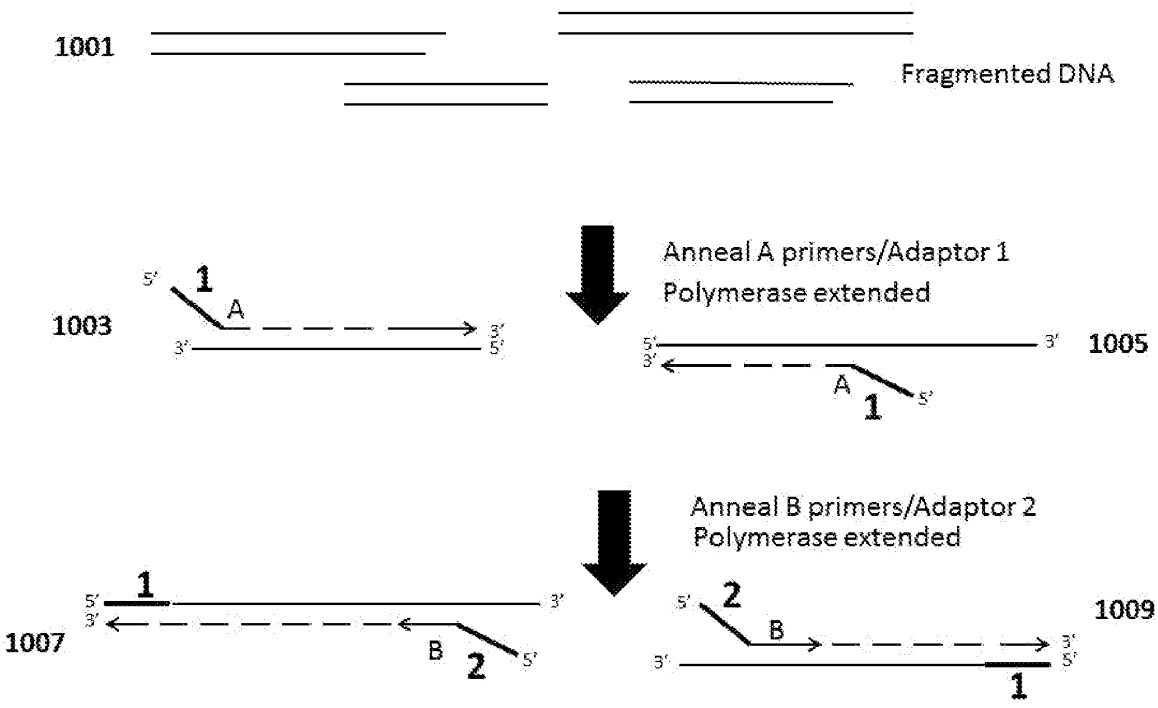

1001
Fragmented DNA
1003
Anneal A primers/Adaptor 1
Polymerase extended
1005
Anneal B primers/Adaptor 2
Polymerase extended
1007
1009

FIG. 10 (cont.)

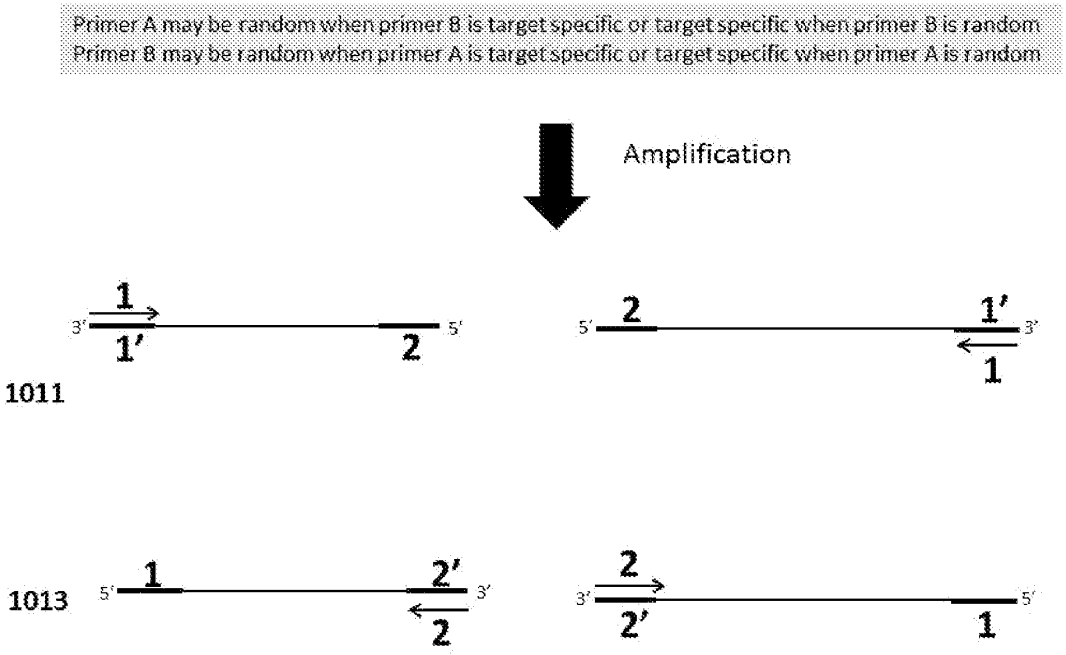

Primer A may be random when primer B is target specific or target specific when primer B is random
Primer B may be random when primer A is target specific or target specific when primer A is random
Amplification
1
1'
2
1011
2
1'
1
1013
1
2'
2
2
2'
1

FIG. 11

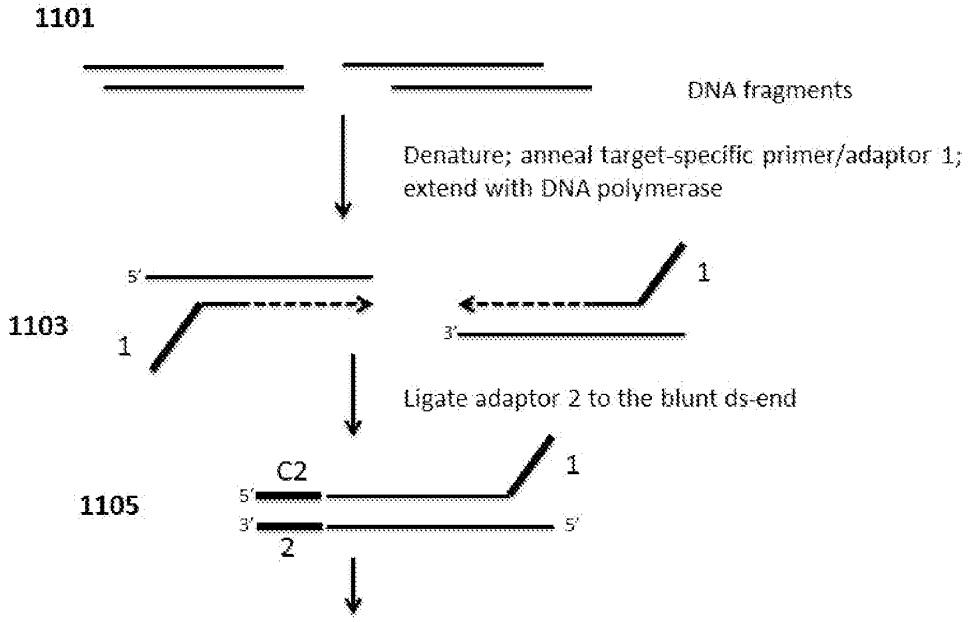

1101
DNA fragments
Denature; anneal target-specific primer/adaptor 1;
extend with DNA polymerase
1103
5'
1
1
3'
Ligate adaptor 2 to the blunt ds-end
1105
C2
1
5'
3'
5'
2

FIG. 11 (cont.)

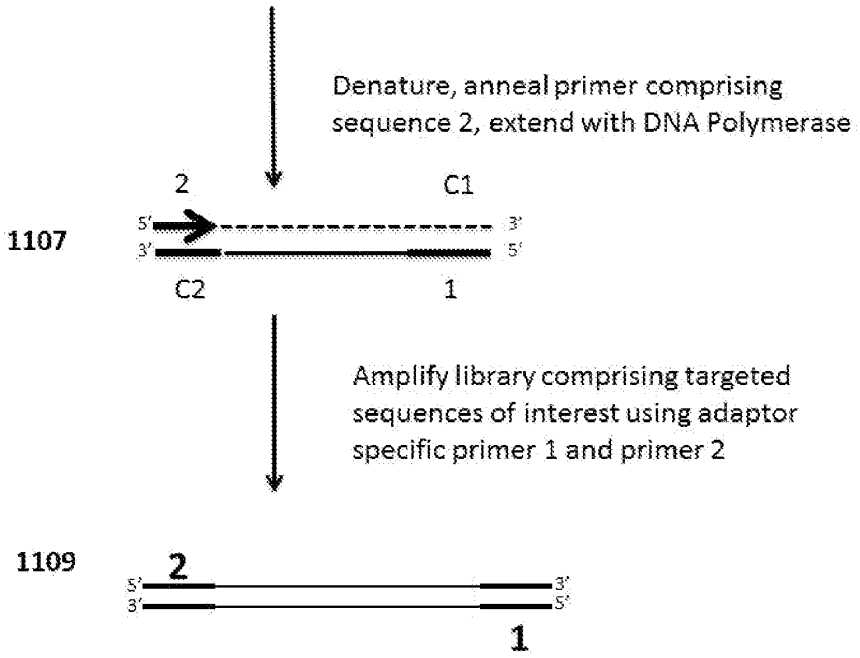

Denature, anneal primer comprising sequence 2, extend with DNA Polymerase
2
C1
1107
5'
3'
3'
5'
C2
1
Amplify library comprising targeted sequences of interest using adaptor specific primer 1 and primer 2
1109
2
5'
3'
3'
5'
1

FIG. 12

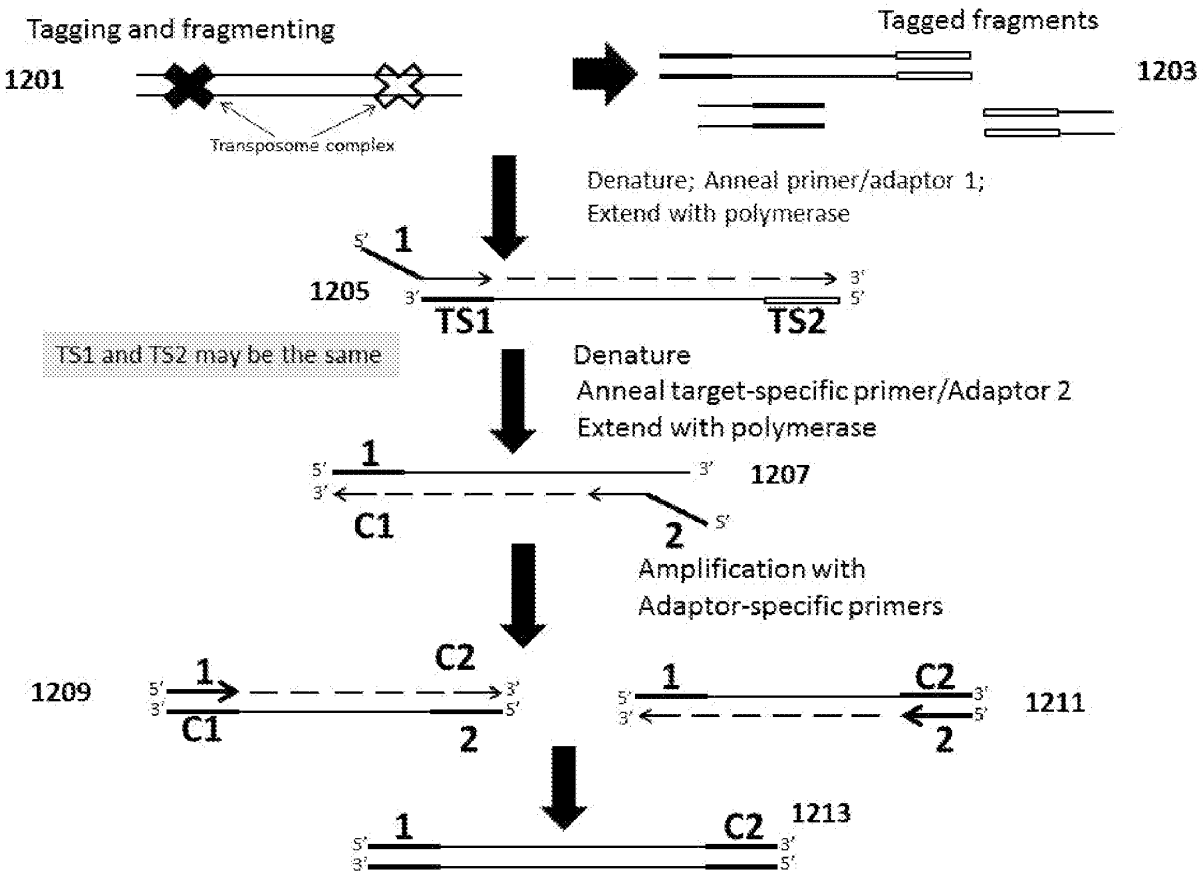

Tagging and fragmenting
1201
Transposome complex
Tagged fragments
1203
Denature; Anneal primer/adaptor 1;
Extend with polymerase
1205
TS1
TS2
TS1 and TS2 may be the same
Denature
Anneal target-specific primer/Adaptor 2
Extend with polymerase
1207
C1
2
Amplification with
Adaptor-specific primers
1209
1211
C2
1213

FIG. 13

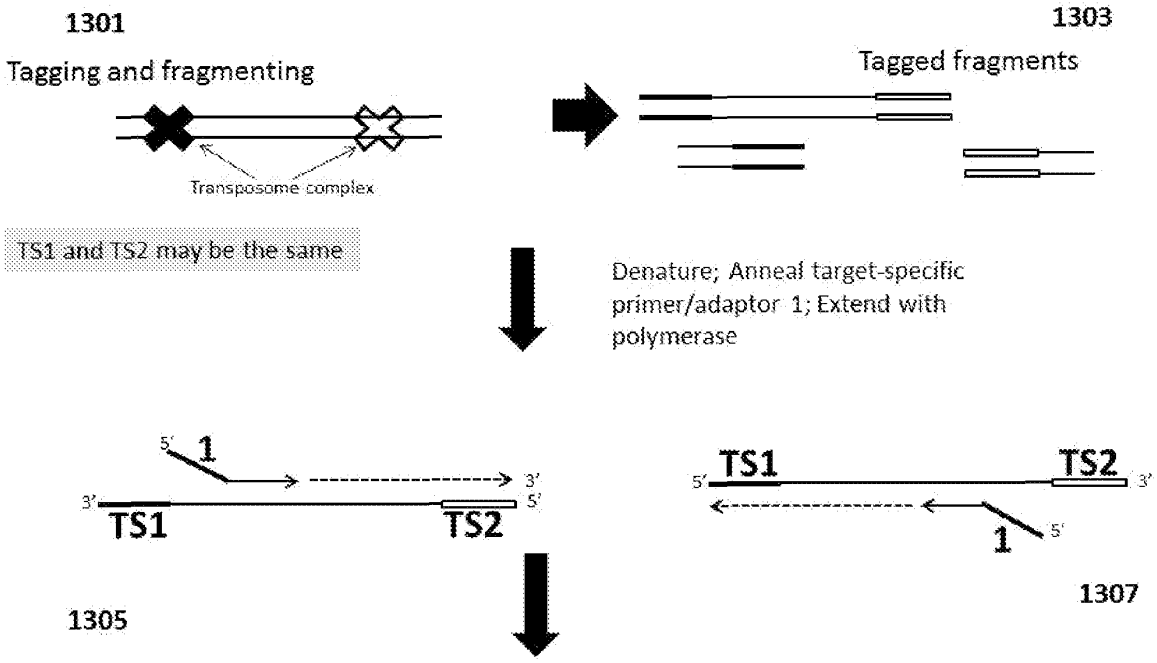

1301
Tagging and fragmenting
Transposome complex
1303
Tagged fragments
TS1 and TS2 may be the same
Denature; Anneal target-specific primer/adaptor 1; Extend with polymerase
5'
1
3'
3'
5'
TS1
TS2
1305
TS1
TS2
5'
3'
1
5'
1307

FIG. 13 (cont.)

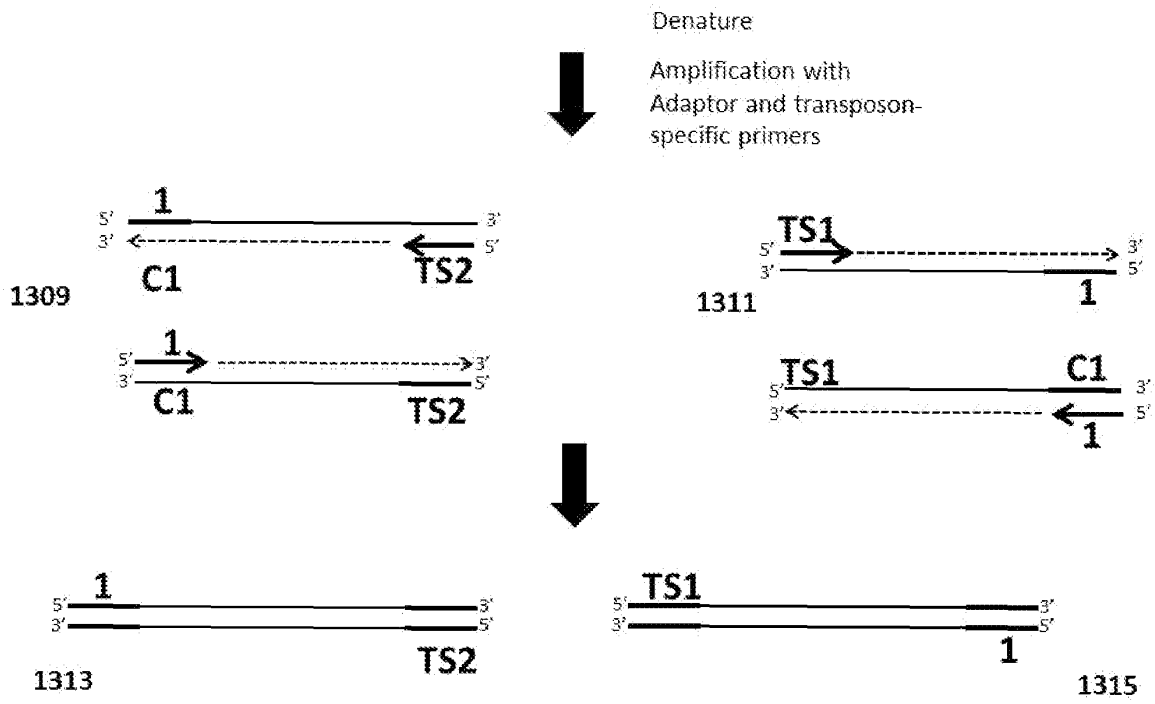

Denature
Amplification with
Adaptor and transposon-
specific primers
1
C1
TS2
1309
TS1
1
1311
1
C1
TS2
TS1
C1
1
1
TS2
1313
TS1
1
1315

FIG. 14

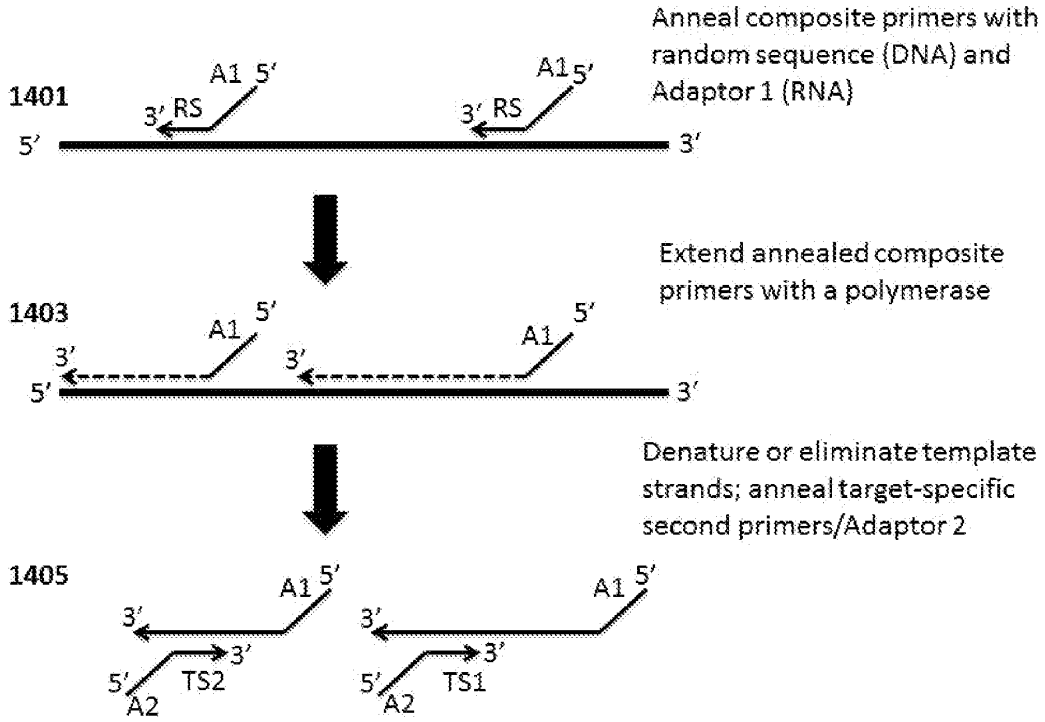
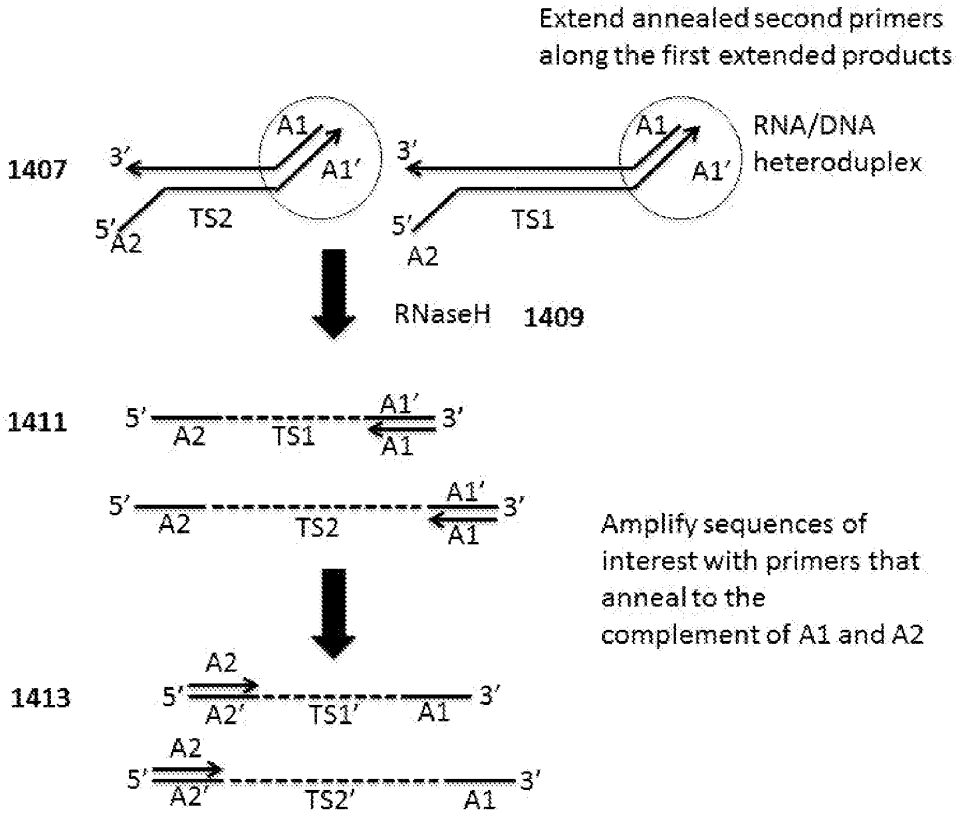

Anneal composite primers with random sequence (DNA) and Adaptor 1 (RNA)
1401
A1
5′
3′
RS
5′
3′
Extend annealed composite primers with a polymerase
1403
Denature or eliminate template strands; anneal target-specific second primers/Adaptor 2
1405
TS2
A2
TS1

FIG. 14 (cont.)


Extend annealed second primers along the first extended products
1407
A1
A1′
3′
5′
A2
TS2
TS1
RNA/DNA heteroduplex
RNaseH
1409
1411
A1′
TS1
TS2
Amplify sequences of interest with primers that anneal to the complement of A1 and A2
1413
A2
A2′
TS1′
TS2′

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
14/877,075	10/07/2015	1634	0.00	25115-704.307	13	7

CONFIRMATION NO. 4814

12754
Wilson Sonsini Goodrich & Rosati/NUGEN
650 Page Mill Road
Palo Alto, CA 94304

FILING RECEIPT

OC000000078094334

Date Mailed: 10/19/2015

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Nurith KURN, Palo Alto, CA;

Applicant(s)

NUGEN TECHNOLOGIES, INC., SAN CARLOS, CA;

Power of Attorney: None

Domestic Priority data as claimed by applicant

This application is a CON of 13/922,146 06/19/2013
which is a CON of 13/282,732 10/27/2011 PAT 8492095
which is a CON of 12/615,958 11/10/2009 PAT 8071311
which is a CON of 12/020,434 01/25/2008 PAT 7771946
which is a CON of 10/934,890 09/03/2004 PAT 7354717
which is a CON of 10/100,321 03/11/2002 PAT 6946251
which claims benefit of 60/274,550 03/09/2001

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.) - None.
Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

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

If Required, Foreign Filing License Granted: 10/16/2015

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 14/877,075**

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

Non-Publication Request: No

Early Publication Request: No
Title

METHODS AND COMPOSITIONS FOR AMPLIFICATION OF RNA SEQUENCES

Preliminary Class

435

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications: No

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit http://www.SelectUSA.gov or call +1-202-482-6800.

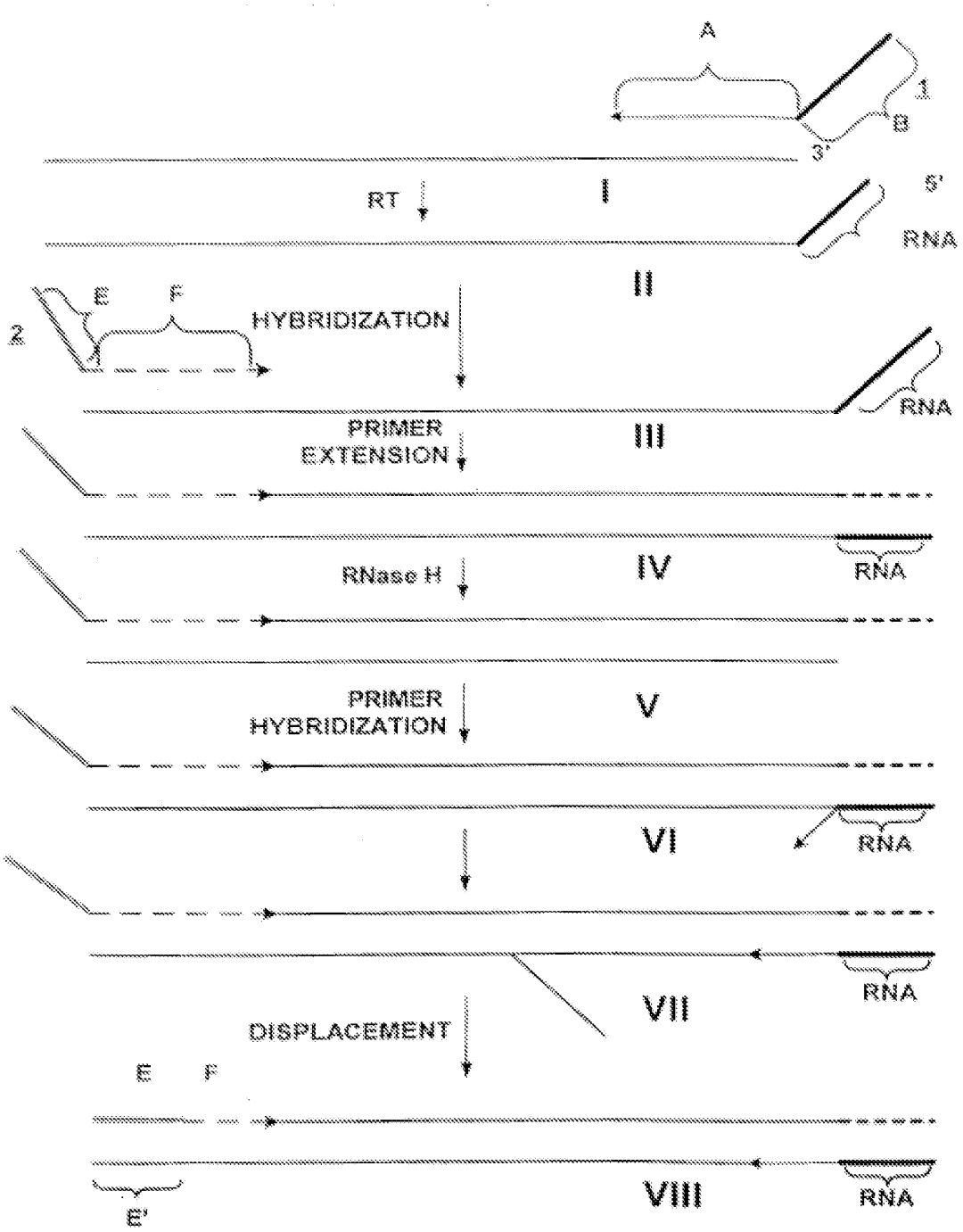

A
1
B
3'
RT
I
5'
RNA
II
E
F
2
HYBRIDIZATION
RNA
PRIMER
EXTENSION
III
RNase H
IV
RNA
PRIMER
HYBRIDIZATION
V
VI
RNA
RNA
VII
DISPLACEMENT
E
F
E'
VIII
RNA

FIG. 1A

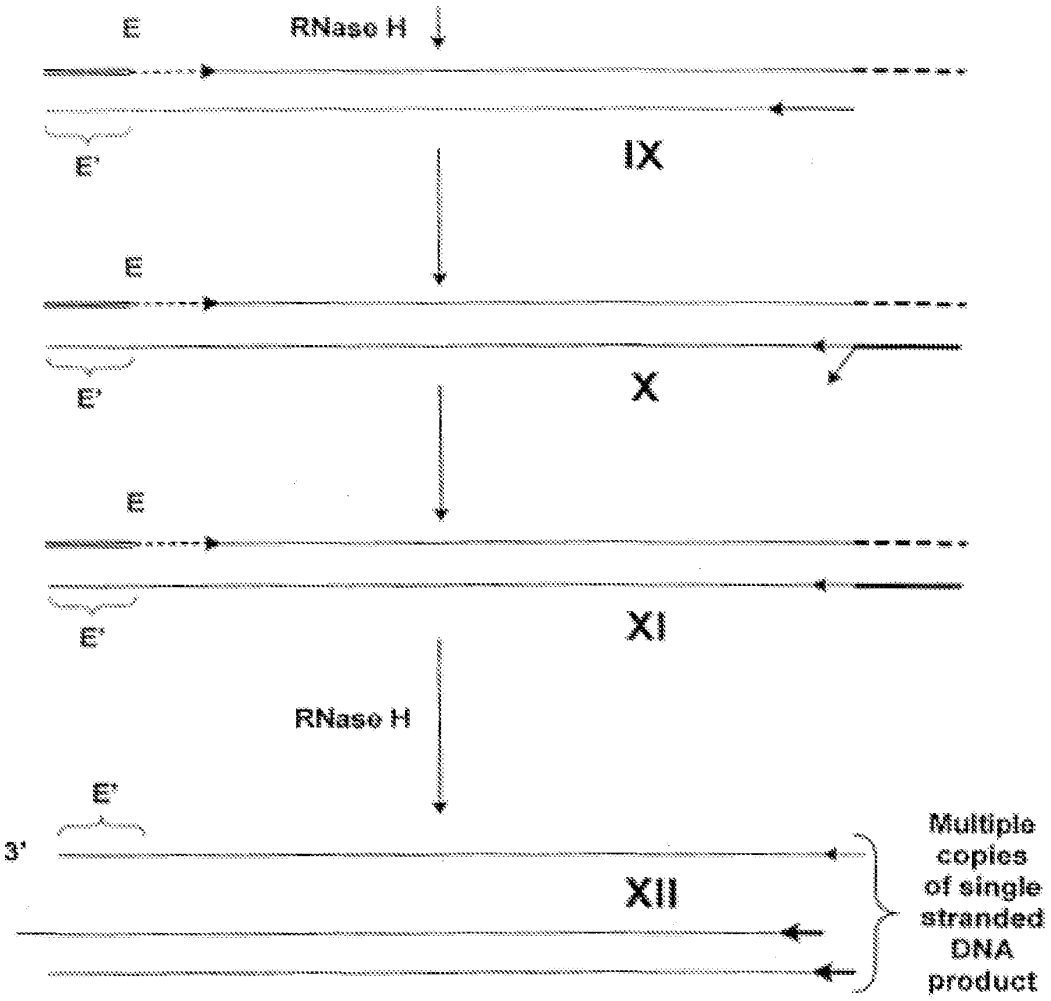

E
RNase H
E'
IX
E
E'
X
E
E'
XI
RNase H
E'
3'
XII
Multiple copies of single stranded DNA product

FIG. 1B

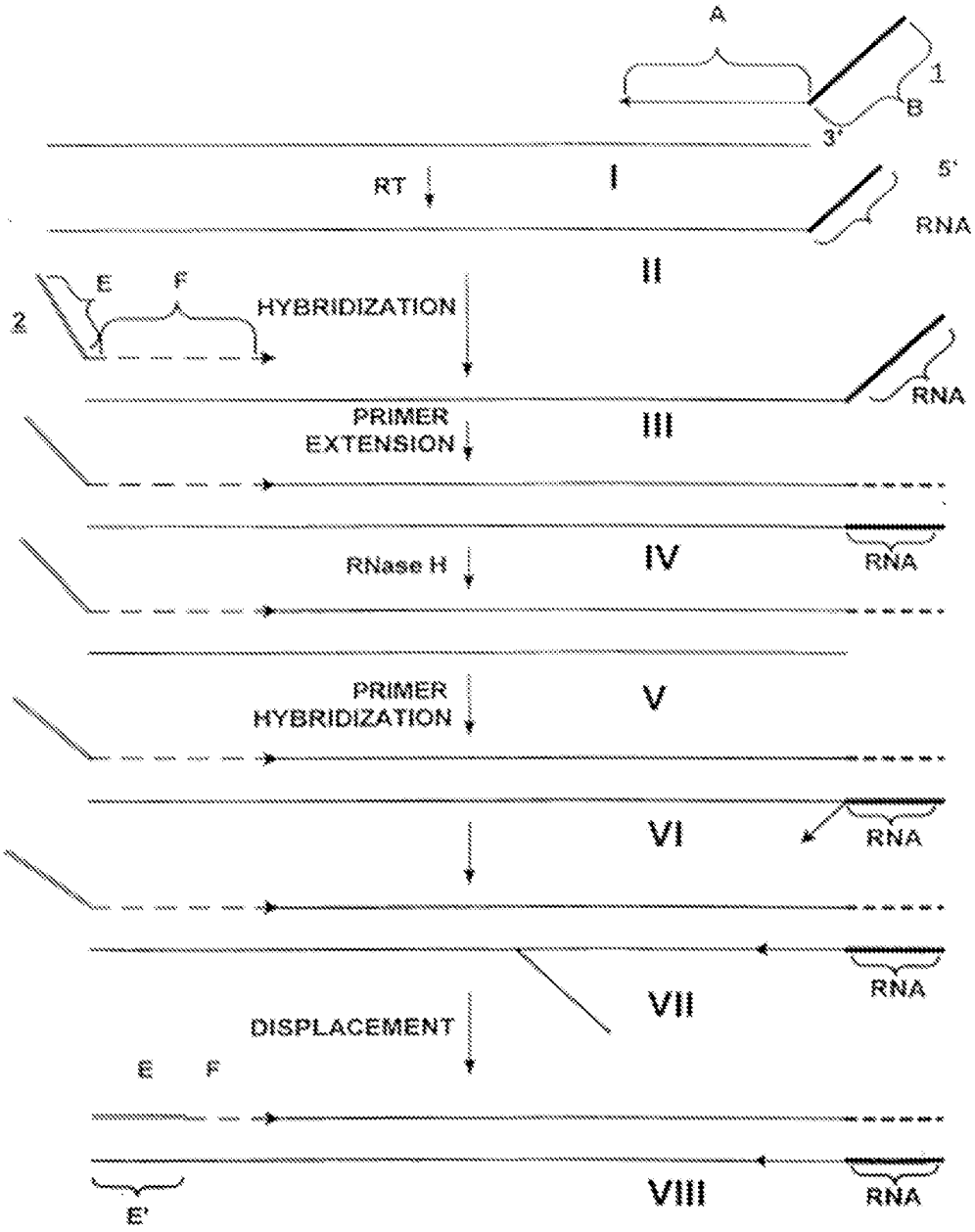

A
1
B
3'
RT
I
5'
RNA
II
E
F
2
HYBRIDIZATION
RNA
PRIMER
EXTENSION
III
RNase H
IV
RNA
PRIMER
HYBRIDIZATION
V
VI
RNA
RNA
VII
DISPLACEMENT
E
F
E'
VIII
RNA

FIG. 2A

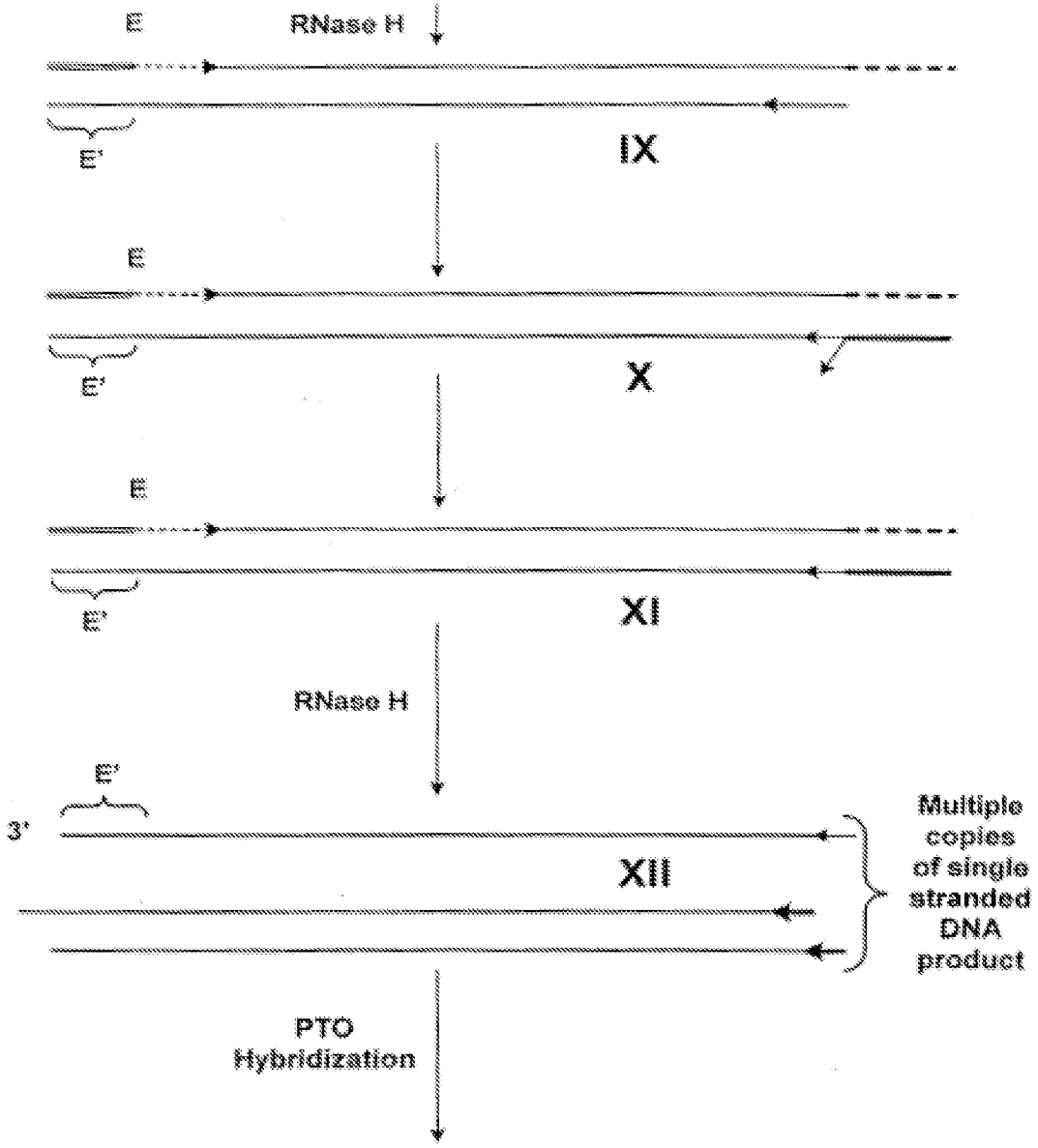

E
RNase H
E'
IX
E
E'
X
E
E'
XI
RNase H
E'
3'
XII
Multiple copies of single stranded DNA product
PTO Hybridization

FIG. 2B

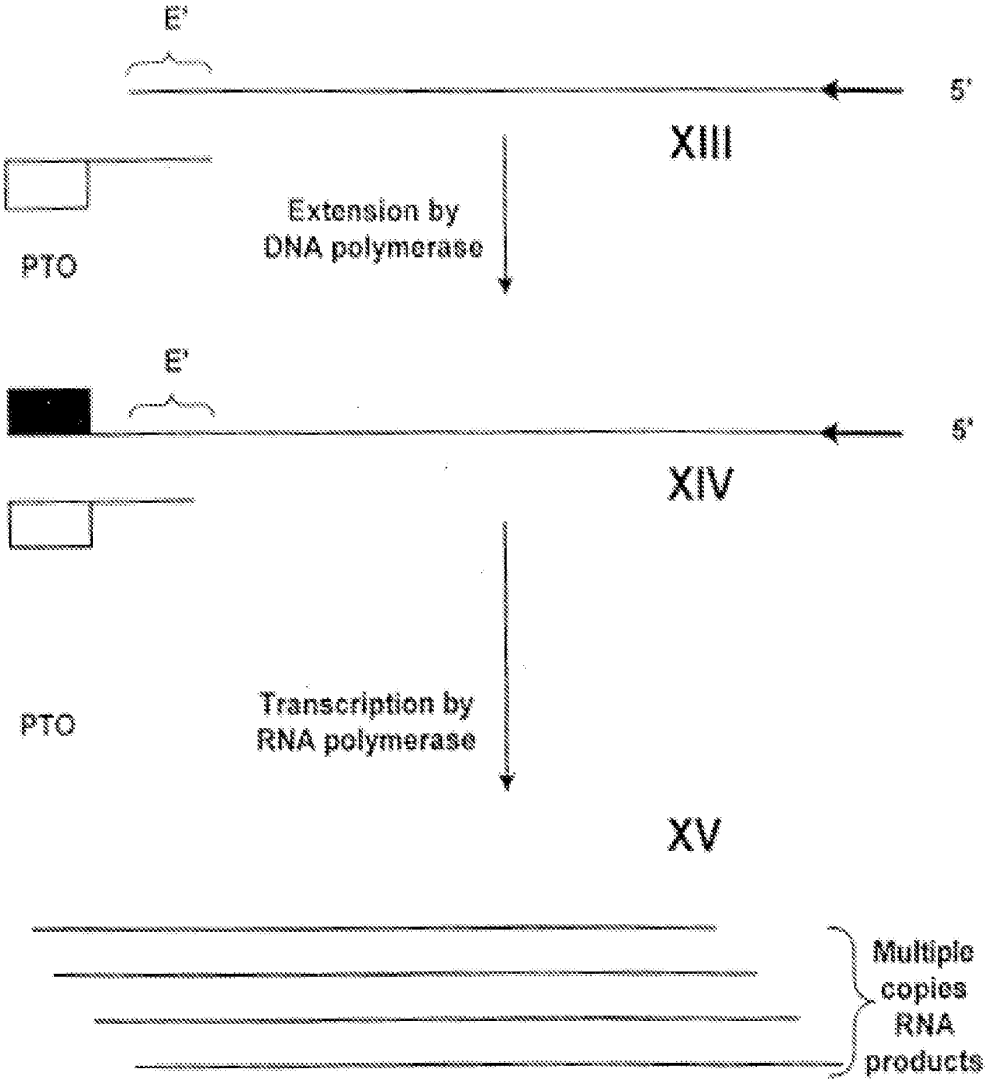

E'
5'
XIII
Extension by
DNA polymerase
PTO
E'
5'
XIV
PTO
Transcription by
RNA polymerase
XV
Multiple
copies
RNA
products

FIG. 2C

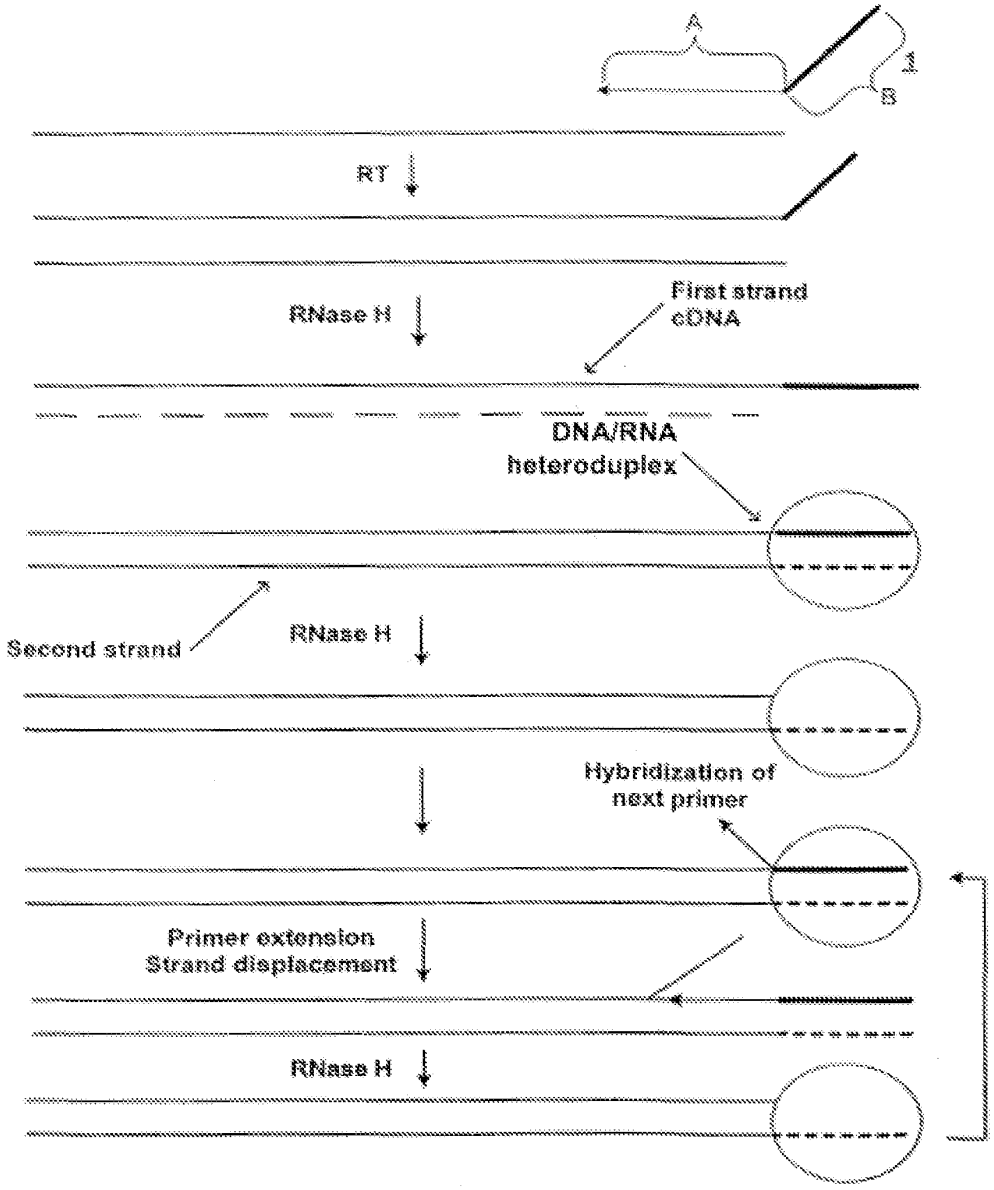

A
1
B
RT
RNase H
First strand
cDNA
DNA/RNA
heteroduplex
Second strand
RNase H
Hybridization of
next primer
Primer extension
Strand displacement
RNase H

FIG. 3A

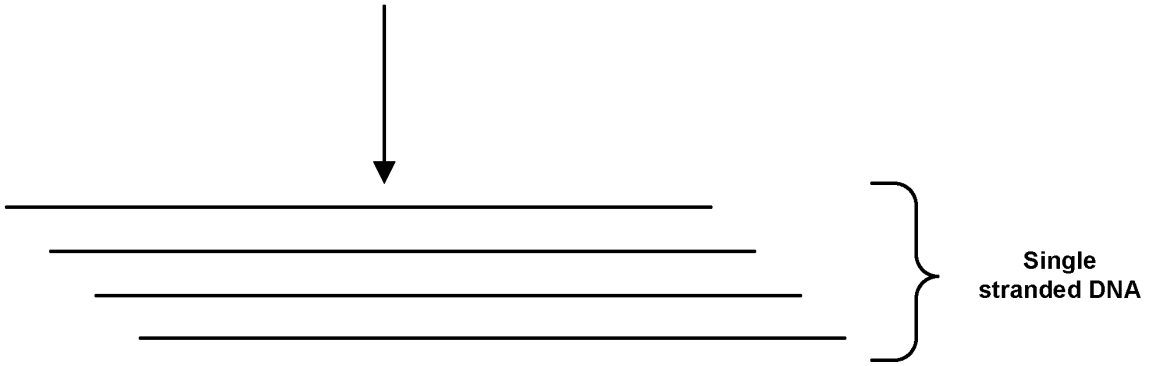


FIG. 3B

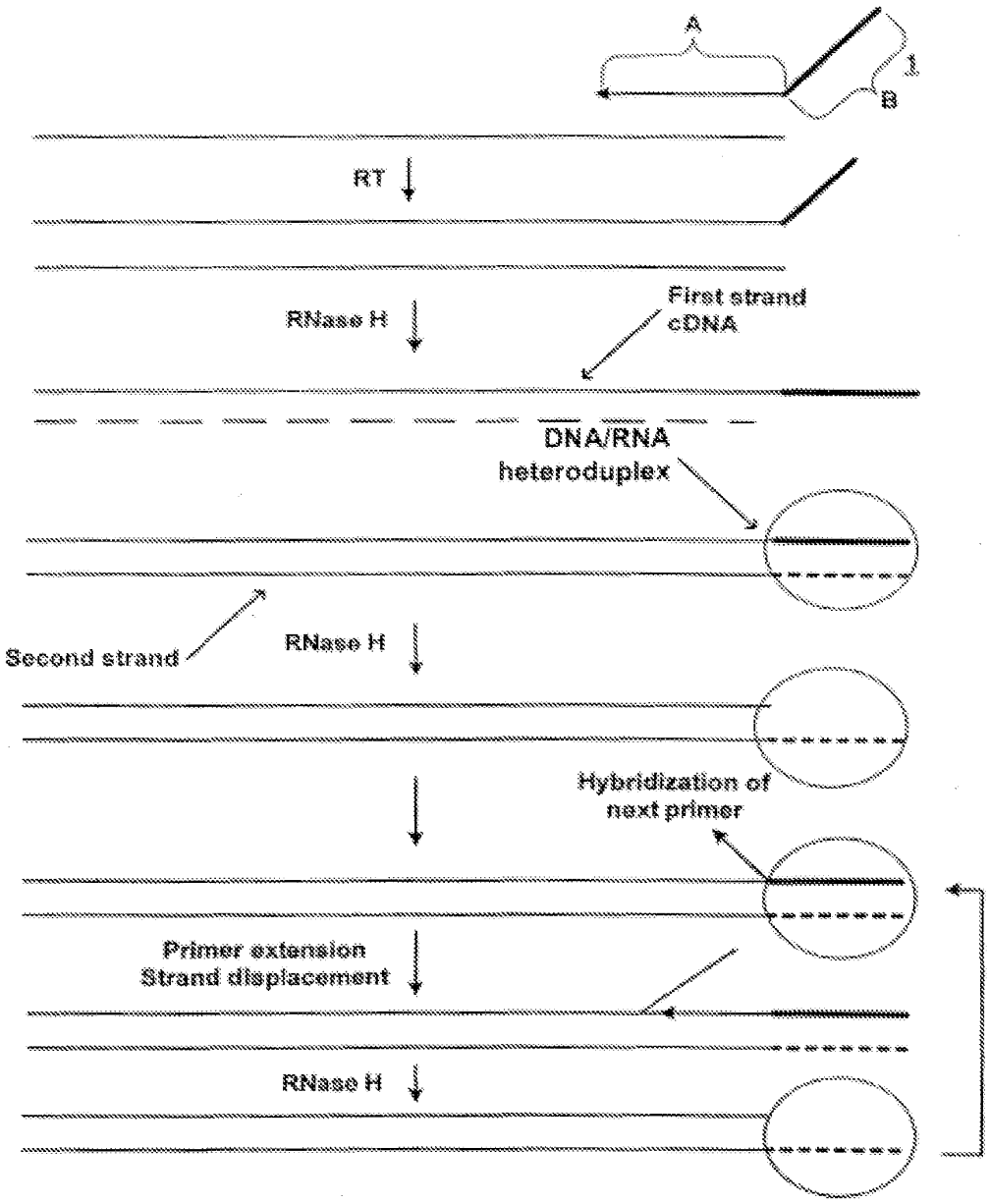

A
B
RT
RNase H
First strand
cDNA
DNA/RNA
heteroduplex
Second strand
RNase H
Hybridization of
next primer
Primer extension
Strand displacement
RNase H

FIG. 4A

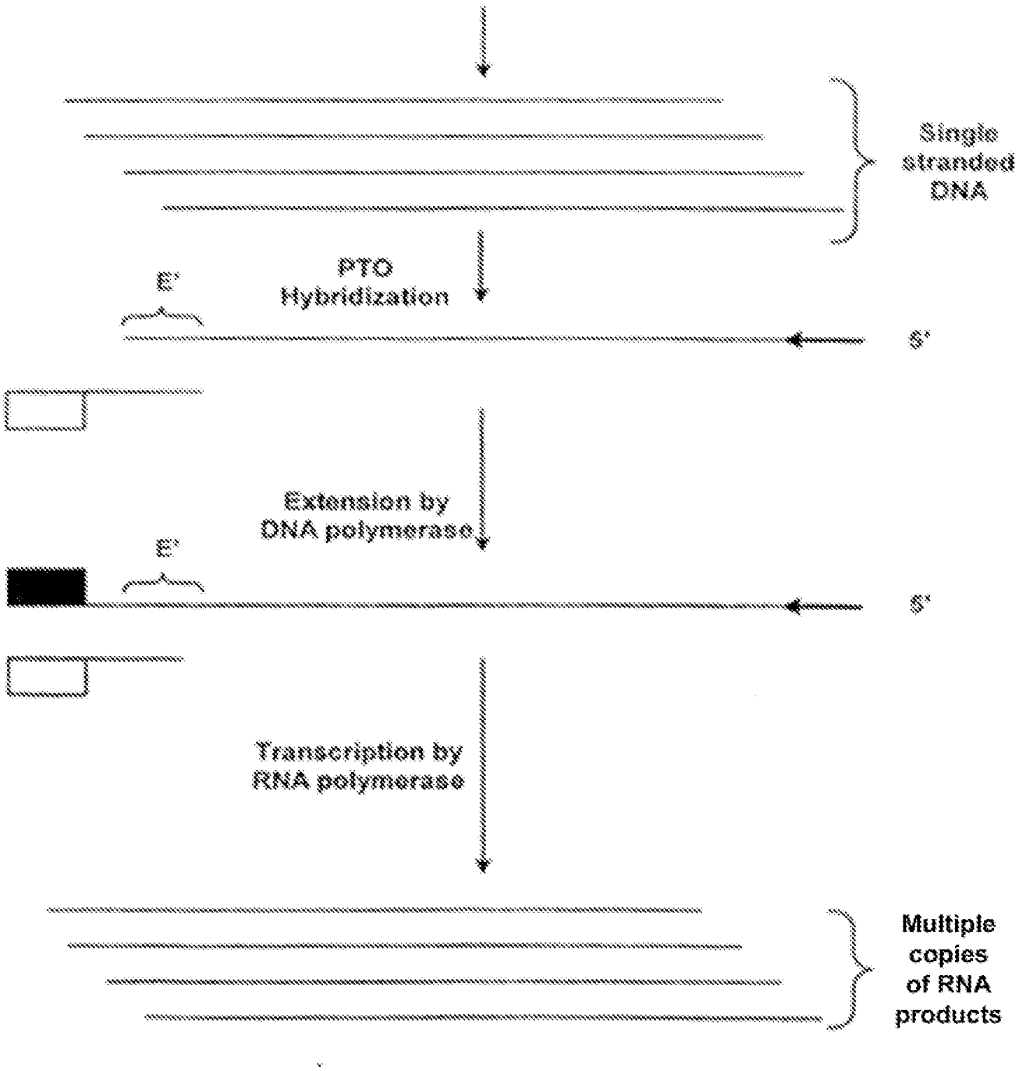

Single stranded DNA
PTO Hybridization
E'
5'
Extension by DNA polymerase
E'
5'
Transcription by RNA polymerase
Multiple copies of RNA products

FIG. 4B

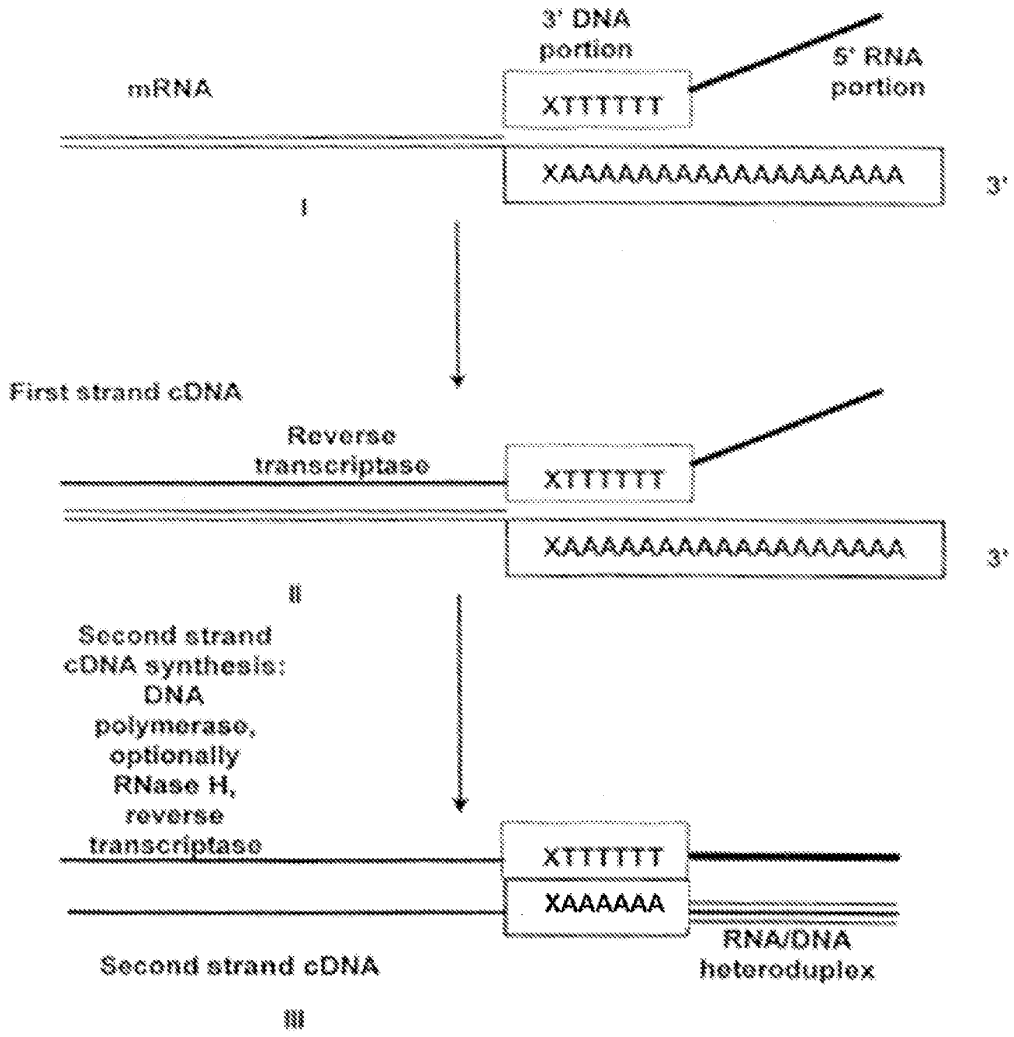

3' DNA portion
5' RNA portion
mRNA
XTTTTTT
XAAAAAAAAAAAAAAAAAA
3'
I
First strand cDNA
Reverse transcriptase
XTTTTTT
XAAAAAAAAAAAAAAAAAA
3'
II
Second strand cDNA synthesis: DNA polymerase, optionally RNase H, reverse transcriptase
XTTTTTT
XAAAAAA
RNA/DNA heteroduplex
Second strand cDNA
III

FIG. 5A

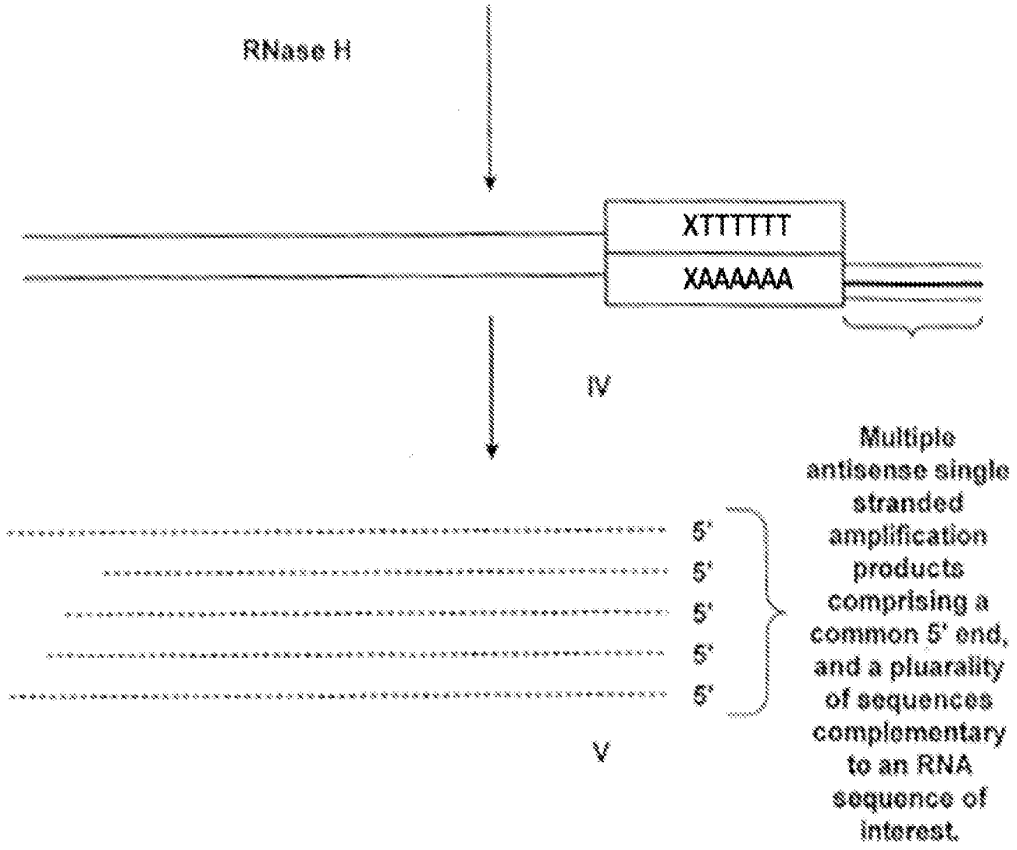

RNase H
XTTTTTT
XAAAAAA
IV
5'
5'
5'
5'
5'
V
Multiple antisense single stranded amplification products comprising a common 5' end, and a pluarality of sequences complementary to an RNA sequence of interest.

FIG. 5B

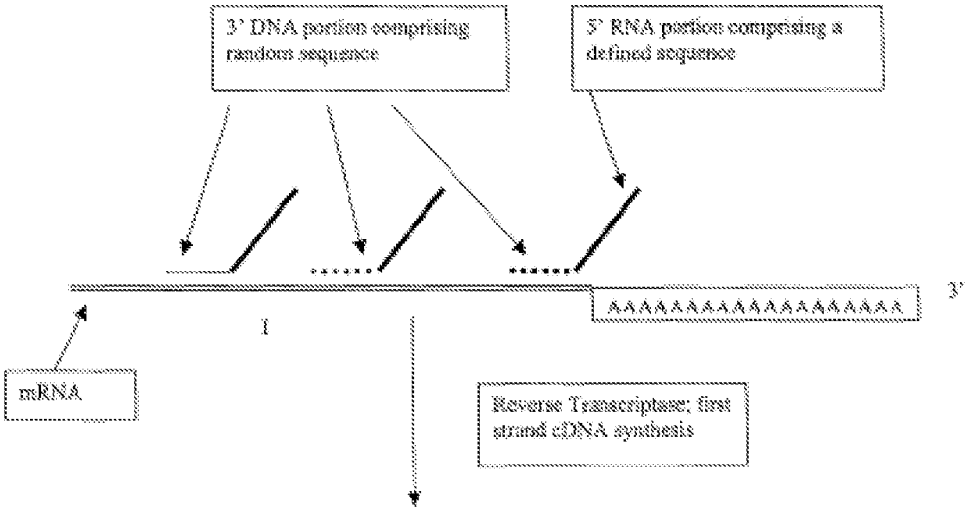


FIG. 6A

FIG. 6B

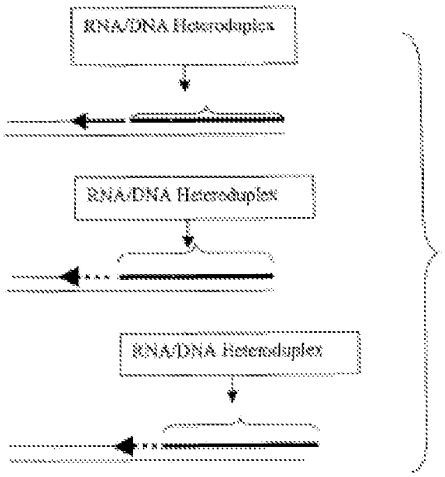
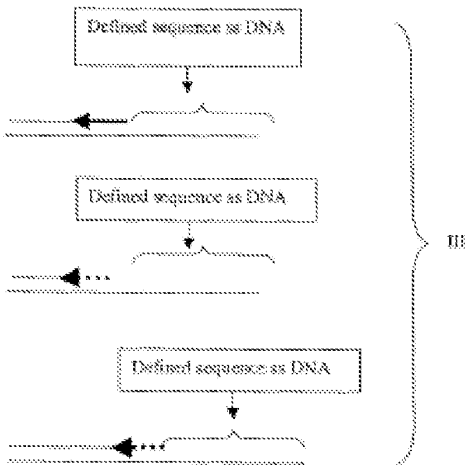


FIG. 6C



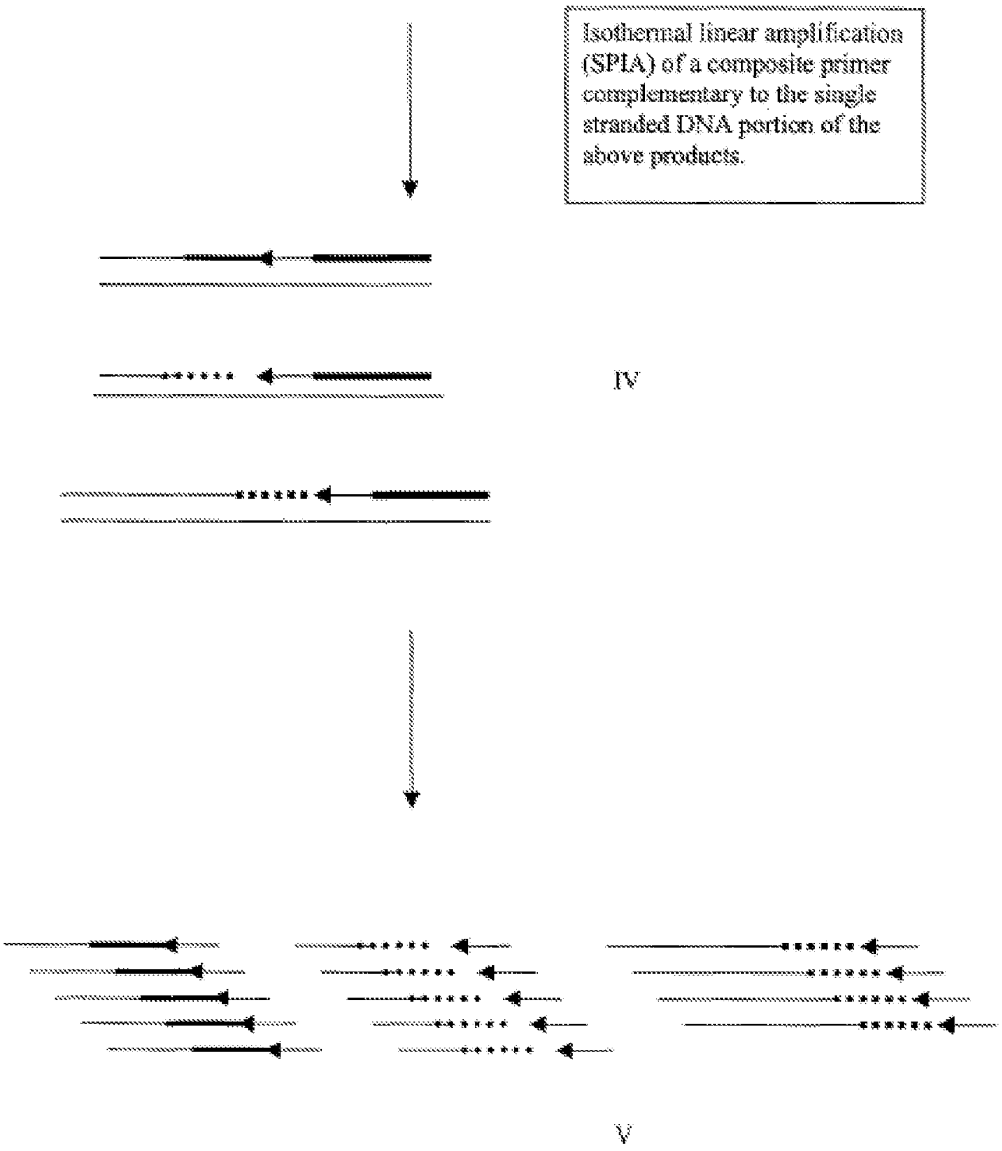

Isothermal linear amplification (SPIA) of a composite primer complementary to the single stranded DNA portion of the above products.
IV
V

FIG. 6D

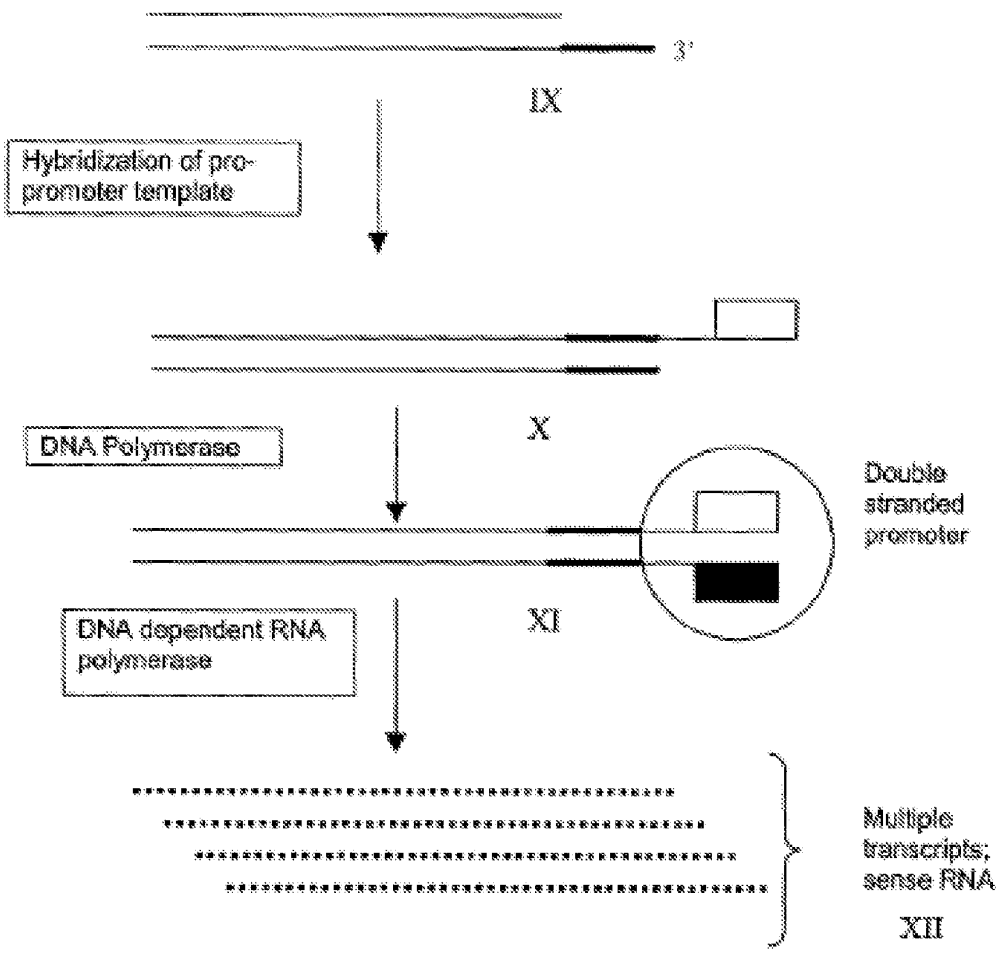

3'
IX
Hybridization of pro-promoter template
X
DNA Polymerase
Double stranded promoter
XI
DNA dependent RNA polymerase
Multiple transcripts; sense RNA
XII

FIG. 7

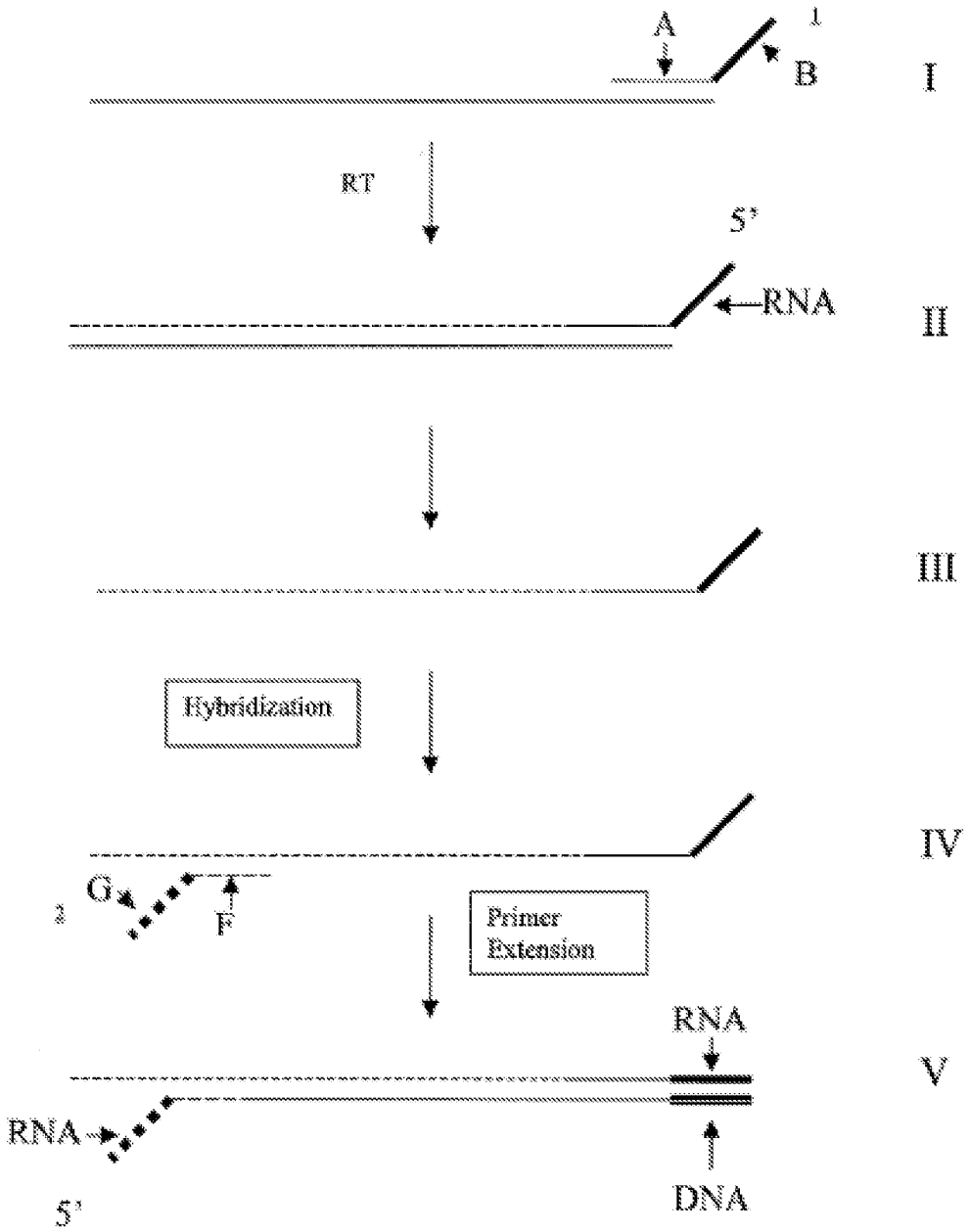

A
1
B
I
RT
5'
RNA
II
III
Hybridization
IV
G
F
Primer
Extension
RNA
V
RNA
5'
DNA

FIG. 8A

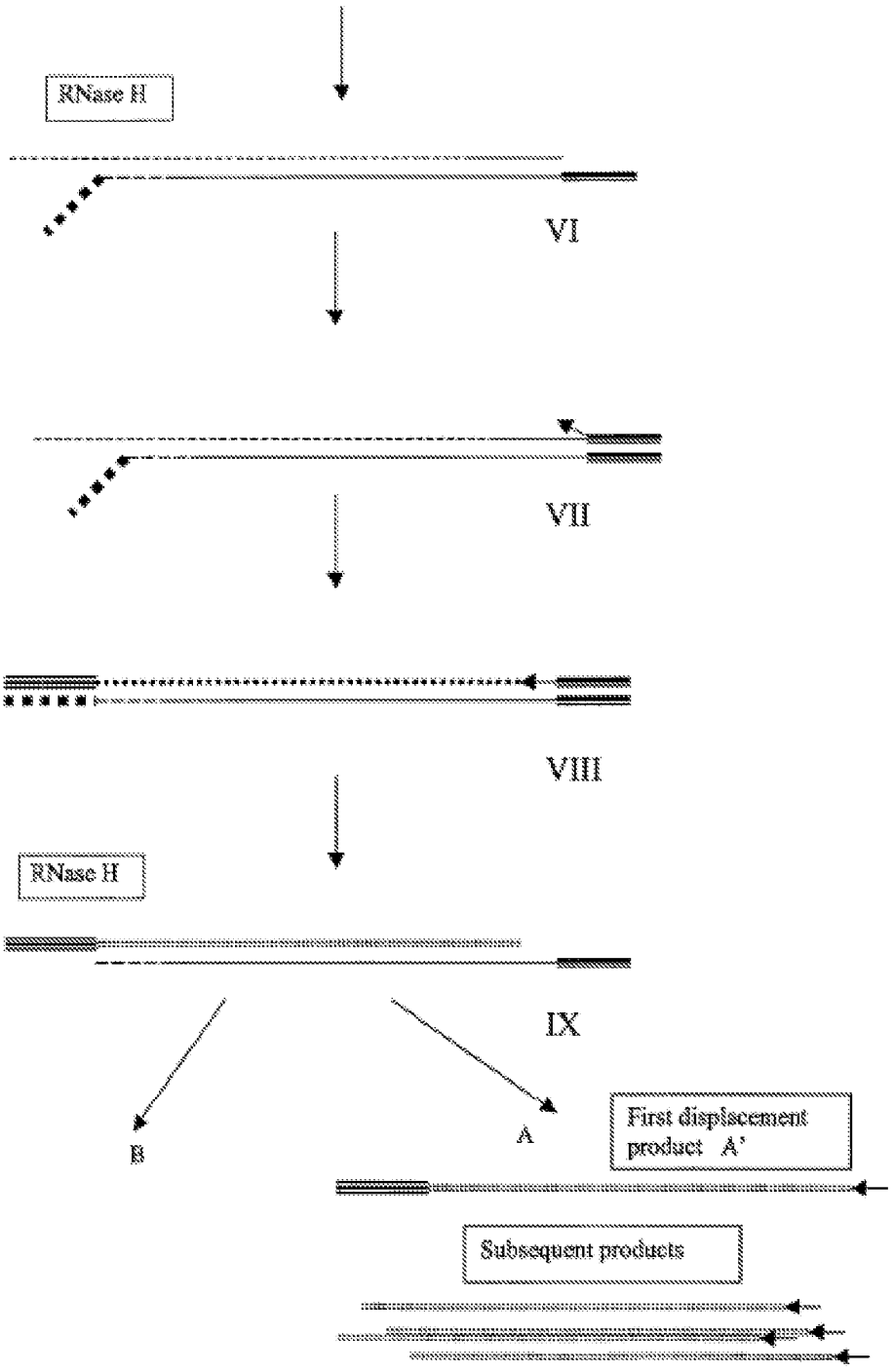

RNase H
VI
VII
VIII
RNase H
IX
B
A
First displacement product A'
Subsequent products

FIG. 8B

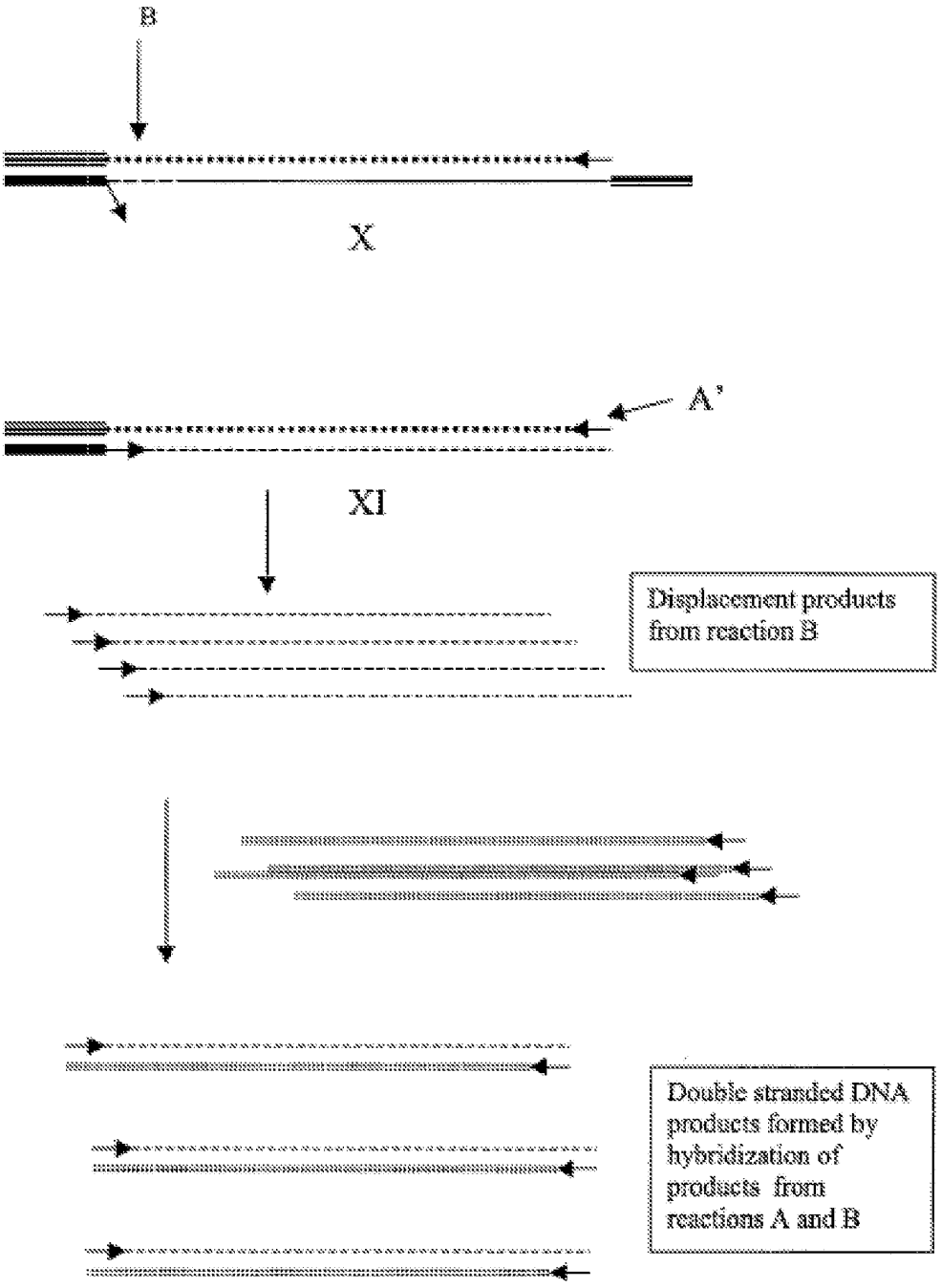

B
X
A'
XI
Displacement products from reaction B
Double stranded DNA products formed by hybridization of products from reactions A and B

FIG. 8C

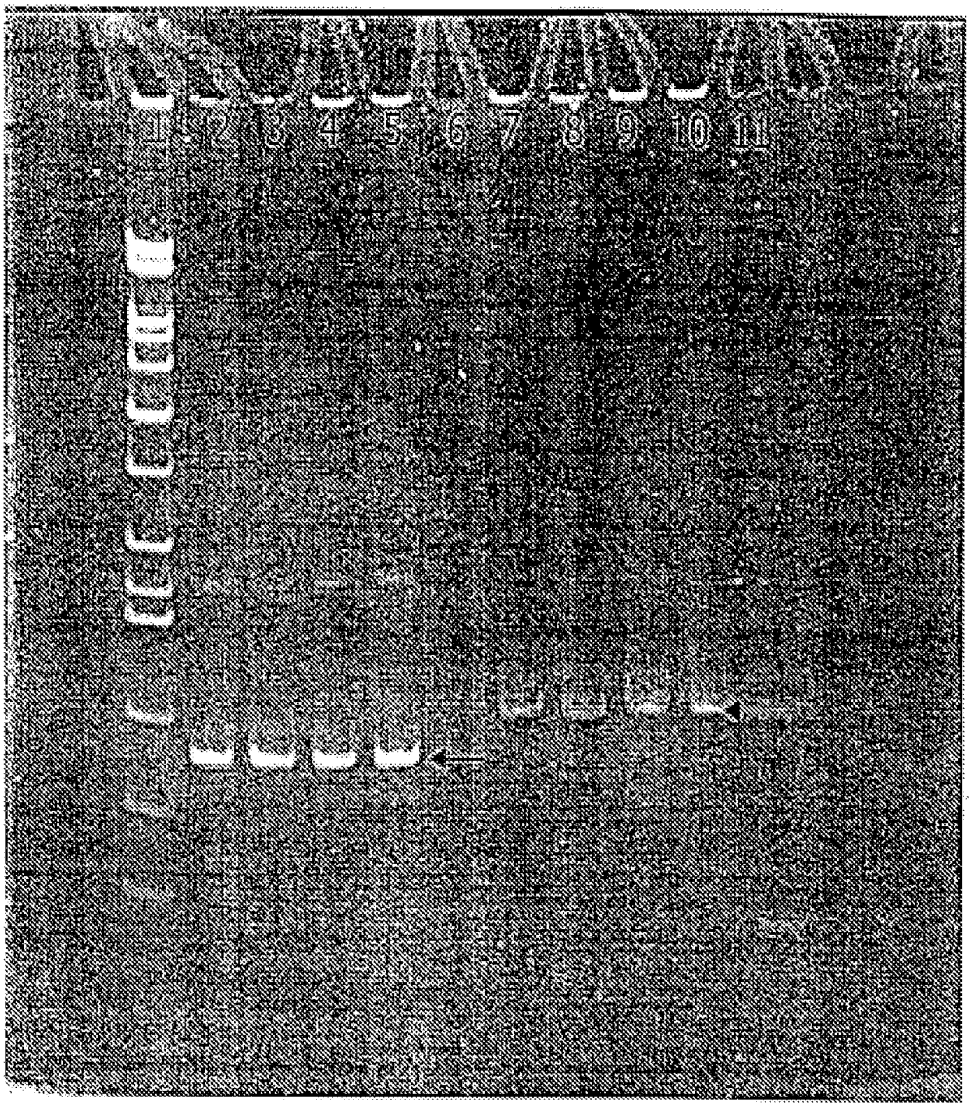


FIG. 9

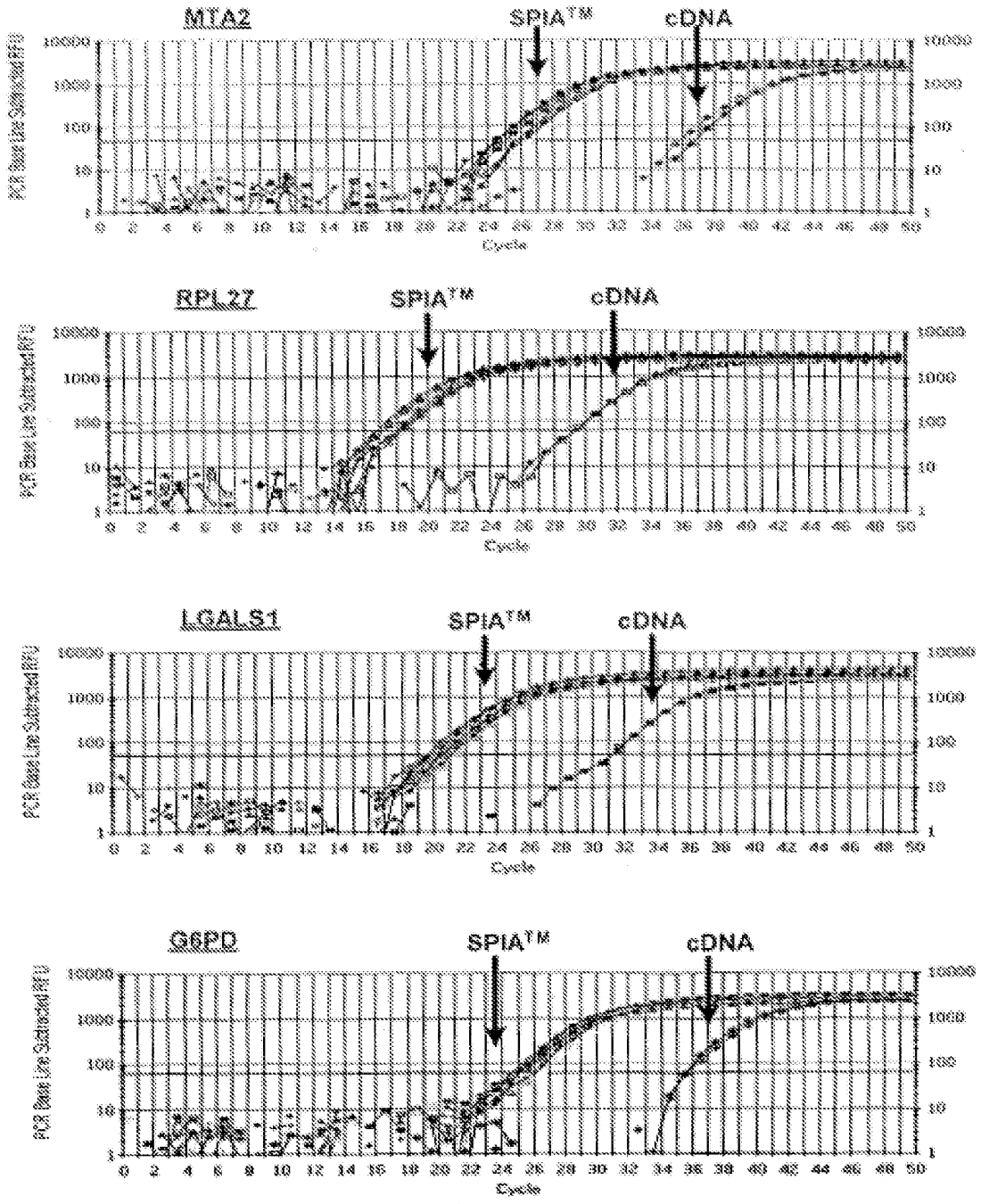

MTA2
SPIA™
cDNA
RPL27
SPIA™
cDNA
LGALS1
SPIA™
cDNA
G6PD
SPIA™
cDNA
PCR Base Line Subtracted RFU
Cycle

FIG. 10

WSGR Docket No. 25115-704.307

PATENT APPLICATION

METHODS AND COMPOSITIONS FOR AMPLIFICATION OF RNA SEQUENCES

Inventor(s): Nurith Kurn,
Citizen of Israel, Residing at
2876 Ramona Street
Palo Alto, CA 94306

Assignee: NuGEN Technologies, Inc.,
201 Industrial Road, Suite 310,
San Carlos, CA 94070

a Delaware Corporation

Entity: Undiscounted business concern

WSGR

Wilson Sonsini Goodrich & Rosati
PROFESSIONAL CORPORATION

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

Filed Electronically on: October 7, 2015

METHODS AND COMPOSITIONS FOR AMPLIFICATION OF RNA SEQUENCES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of U.S. Serial No. 13/922,146, filed June 19, 2013, which is a continuation of U.S. Serial No. 13/282,732, filed October 27, 2011, now U.S. Patent 8,492,095, which is a continuation of U.S. Serial No. 12/615,958, filed November 10, 2009, now U.S. Patent 8,071,311, which is a continuation of U.S. Serial No. 12/020,434, filed January 25, 2008, now U.S. Patent 7,771,946, which is a continuation of U.S. Serial No. 10/934,890, filed September 3, 2004, now U.S. Patent 7,354,717, which is a continuation of U.S. Serial No. 10/100,321, filed March 11, 2002, now U. S. Patent 6,946,251, which claims the priority benefit of provisional patent application U.S. Serial No. 60/274,550, filed March 9, 2001, all of which are incorporated herein by reference in their entirety.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on October 5, 2015 is named 25115704307_SeqList.txt, and is 7 kb in size.

TECHNICAL FIELD

The invention relates to the field of polynucleotide amplification. More particularly, the invention provides methods, compositions and kits for amplifying (i.e., making multiple copies) RNA sequences of interest which employ an RNA/DNA composite primer, and, optionally transcription.

BACKGROUND ART

The ability to amplify ribonucleic acid (RNA) is an important aspect of efforts to elucidate biological processes. To date, RNA (generally, mRNA) amplification is most commonly performed using the reverse transcriptase-polymerase chain reaction

(RT-PCR) method and variations thereof. These methods are based on replication of RNA by reverse transcriptase to form single stranded DNA complementary to the RNA (cDNA), which is followed by polymerase chain reaction (PCR) amplification to produce multiple copies of double stranded DNA. Although these methods are most commonly used, they have some significant drawbacks: a) the reactions require thermocycling; b) the products are double stranded, thus rendering them less accessible to binding to probes; c) the reactions are prone to contamination with products of prior amplification, thus requiring strict containment of reaction mixtures; and d) the exponential nature of amplification of these methods renders them prone to generate pools of products which do not truly reflect the representation of the various RNA sequences in the input total RNA sample, due to unequal efficiency of amplification of different sequences, and the nature of exponential amplification which is based on replication of amplification products rather than on continued replication of the input target RNAs.

Total cellular mRNA represents gene expression activity at a defined time. Gene expression is affected by cell cycle progression, developmental regulation, response to internal and external stimuli and the like. The profile of expressed genes for any cell type in an organism reflects normal or disease states, response to various stimuli, developmental stages, cell differentiation, and the like.

Various methods for the analysis of gene expression have been developed in recent years. See, for example, U.S. Pat. Nos. 5,744,308; 6,143,495; 5,824,517; 5,829,547; 5,888,779; 5,545,522; 5,716,785; 5,409,818; EP 0971039A2; EP0878553A2. These include quantification of specific mRNAs, and the simultaneous quantification of a large number of mRNAs, as well as the detection and quantification of patterns of expression of known and unknown genes. The analysis of gene expression profiles is currently one of the most powerful tools in the study of cellular differentiation and cellular development, and in the investigation of normal and disease states of various organisms, in particular in human. This analysis is crucial for gene discovery, molecular medicine and drug discovery processes.

Amplification of the total cellular mRNAs prepared from any cell or tissue is generally critical for gene expression profiling. Although analysis of non-amplified

mRNA is feasible, a significant amount of starting mRNA would be required. However, the total amount of sample mRNA that is available is frequently limited by the amount of biological sample from which it is derived. Biological samples are often limited in amount and precious. Moreover, the amount of the various mRNA species is not equal; some species are more abundant than others, and these are more likely and easier, to analyze. The ability to amplify mRNA sequences enables the analysis of less abundant, rare mRNA species. The ability to analyze small samples, by means of nucleic acid amplification, is also advantageous for design parameters of large scale screening of effector molecule libraries, for which reduction in sample volume is a major concern both for the ability to perform very large scale screening or ultra high throughput screening, and in view of the limiting amounts of library components.

Therefore, there is a need for improved RNA amplification methods that overcome drawbacks in existing methods. The invention provided herein fulfills this need and provides additional benefits.

All references cited herein, including patent applications and publications, are incorporated by reference in their entirety.

DISCLOSURE OF THE INVENTION

The invention provides methods, compositions, and kits for RNA amplification, as well as applications of the amplification methods.

Accordingly, in one aspect, the invention provides methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest, said method comprising the steps of: (a) extending a first primer hybridized to a target RNA with an RNA-dependent DNA polymerase, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion, whereby a complex comprising a first primer extension product and the target RNA is produced; (b) cleaving RNA in the complex of step (b) with an enzyme that cleaves RNA from an RNA/DNA hybrid; (c) extending a second primer hybridized to the first primer extension product with a DNA-dependent DNA polymerase, whereby a second

primer extension product is produced to form a complex of first and second primer extension products; (d) cleaving RNA from the composite primer in the complex of first and second primer extension products with an enzyme that cleaves RNA from an RNA/DNA hybrid such that a composite primer hybridizes to the second primer extension product, wherein the composite primer comprises an RNA portion and a 3' DNA portion; (e) extending the composite primer hybridized to the second primer extension product with a DNA-dependent DNA polymerase; whereby said first primer extension product is displaced, and whereby multiple copies of a polynucleotide sequence complementary to the RNA sequence of interest are generated.

In another aspect, the invention provides methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest, said method comprising the steps of: (a) extending a second primer hybridized to a first primer extension product with a DNA-dependent DNA polymerase, wherein the first primer extension product comprises an RNA portion at the 5' end, said first primer extension product comprising a sequence complementary to an RNA sequence, whereby a second primer extension product is produced to form a complex of first and second primer extension products; (b) cleaving RNA in the complex of first and second primer extension products with an enzyme that cleaves RNA from an RNA/DNA hybrid such that a composite primer hybridizes to the second primer extension product, wherein the composite primer comprises an RNA portion and a 3' DNA portion; (c) extending the composite primer hybridized to the second primer extension product with a DNA-dependent DNA polymerase; whereby said first primer extension product is displaced, and whereby multiple copies of a polynucleotide sequence complementary to the RNA sequence of interest are generated. In some embodiments of the invention, the first primer extension product is produced by extension of a first primer hybridized to a target RNA with a RNA-dependent DNA polymerase, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion.

In another aspect, the invention provides methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest,

said method comprising the steps of: (a) cleaving RNA from a complex of first and second primer extension products with an enzyme that cleaves RNA from an RNA/DNA hybrid such that a composite primer hybridizes to the second primer extension product, wherein the composite primer comprises an RNA portion and a 3' DNA portion, wherein the first primer extension product is produced by extension of a first primer hybridized to a target RNA with a RNA-dependent DNA polymerase, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion; (b) hybridizing a composite primer to the second primer extension product and extending the composite primer with a DNA-dependent DNA polymerase; whereby said first primer extension product is displaced, and whereby multiple copies of a polynucleotide sequence complementary to the RNA sequence of interest are generated.

In another aspect, the invention provides methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest, said method comprising the step of: extending a composite primer in a complex comprising: (i) a complex of a first and second primer extension products, wherein the first primer extension product is produced by extension of a first primer hybridized to a target RNA with a RNA-dependent DNA polymerase, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion, wherein the second primer extension product is generated by extension of a second primer hybridized to the first primer extension product, and wherein RNA from the complex of first and second primer extension products is cleaved with an enzyme that cleaves RNA from an RNA/DNA hybrid; and (ii) a composite primer, said composite primer comprising an RNA portion and a 3' DNA portion, wherein the composite primer is hybridized to the second primer extension product, and wherein the composite primer may be the same or different from the first primer; whereby said first primer extension product is displaced, and whereby multiple copies of a polynucleotide sequence complementary to the RNA sequence of interest are generated.

In another aspect, the invention provides methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest,

said method comprising the steps of: (a) extending a composite primer hybridized to a second primer extension product, wherein said primer extension product (i) comprises a complement or a sequence of a first primer extension product generated by extension of a first primer hybridized to template RNA, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion; and (ii) is hybridized to a second primer extension product generated by extension of a second primer hybridized to the first primer extension product, and cleavage of RNA from the first primer with an enzyme that cleaves RNA from an RNA/DNA hybrid; whereby said first primer extension product is displaced, and whereby multiple copies of a polynucleotide sequence complementary to the RNA sequence of interest are generated.

In another aspect, the invention provides methods of generating multiple copies of an RNA sequence of interest, said method comprising the steps of:

hybridizing the displaced first primer extension product from any of the methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest described herein, with a polynucleotide comprising a propromoter and a region which is hybridizable to the displaced first primer extension product under conditions which allow transcription to occur by RNA polymerase, such that RNA transcripts are produced comprising sequences complementary to the displaced first primer extension product, whereby multiple copies of the RNA sequence of interest are generated.

In another aspect, the invention provides methods of generating multiple copies of an RNA sequence of interest, said method comprising the steps of: hybridizing a first primer extension product from any of the methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest described herein, with a polynucleotide comprising a propromoter and a region which is hybridizable to the displaced first primer extension product under conditions which allow transcription to occur by RNA polymerase, such that RNA transcripts are produced comprising sequences complementary to the displaced first primer extension product, wherein the primer extension product is a displaced primer extension product generated by: (a) extending a first primer hybridized to a target

RNA with an RNA-dependent DNA polymerase, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion, whereby a complex comprising a first primer extension product and the target RNA is produced; (b) cleaving RNA in the complex of step (b) with an enzyme that cleaves RNA from an RNA/DNA hybrid; (c) extending a second primer hybridized to the first primer extension product with a DNA-dependent DNA polymerase, whereby a second primer extension product is produced to form a complex of first and second primer extension products; (d) cleaving RNA from the composite primer in the complex of first and second primer extension products with an enzyme that cleaves RNA from an RNA/DNA hybrid such that a composite primer hybridizes to the second primer extension product, wherein the composite primer comprises an RNA portion and a 3' DNA portion; (e) extending said composite primer hybridized to the second primer extension product with a DNA-dependent DNA polymerase, whereby said first primer extension product is displaced; whereby multiple copies of the RNA sequence of interest are generated.

In another aspect, the invention provides methods of generating multiple copies of (amplifying) a polynucleotide sequence complementary to an RNA sequence of interest, said method comprising the steps of: (a) hybridizing a first primer to a target RNA, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion; (b) extending the first primer with an RNA-dependent DNA polymerase, whereby a complex comprising a first primer extension product and the target RNA is produced; (c) cleaving RNA in the complex of step (b) with an agent (such as an enzyme) that cleaves RNA from an RNA/DNA hybrid; (d) hybridizing a second primer to the first primer extension product; (e) extending the second primer with a DNA-dependent DNA polymerase, whereby a second primer extension product is produced to form a complex of first and second primer extension products; (f) cleaving RNA from the composite primer in the complex of first and second primer extension products with an agent (such as an enzyme) that cleaves RNA from an RNA/DNA hybrid such that a composite primer (which may or may not be the same as the first primer) hybridizes to the second primer extension product, wherein said composite primer comprises an RNA portion and a 3' DNA

portion; (g) extending said composite primer of step (f) with a DNA-dependent DNA polymerase, whereby said first primer extension product is displaced, and whereby multiple copies of a polynucleotide sequence complementary to the RNA sequence of interest are generated.

In some aspects, the invention provides methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest, said method comprising the steps of: combining and reacting: the complex of first and second primer extension products of step (e) described above; a composite primer (which may or may not be the same as the first primer) that is hybridizable to the second primer extension product, wherein said composite primer comprises an RNA portion and a 3' DNA portion; a DNA-dependent DNA polymerase; and an agent (such as an enzyme) that cleaves RNA from an RNA/DNA hybrid; wherein the mixture is incubated under conditions (which include necessary substrates and buffer conditions) that permit primer hybridization, RNA cleavage, and displacement of the first primer extension product from the complex of step (e) described above when its RNA is cleaved and a composite primer binds to the second primer extension product in the complex.

As is clear to one skilled in the art, reference to production of copies of an RNA sequence of interest or copies of a polynucleotide sequence complementary to an RNA sequence of interest refers to products that may contain, comprise or consist of such sequences. As is evident to one skilled in the art, aspects that refer to combining and incubating the resultant mixture also encompasses method embodiments which comprise incubating the various mixtures (in various combinations and/or subcombinations) so that the desired products are formed.

In another aspect, the invention provides methods of generating multiple copies of a polynucleotide sequence complementary to an RNA sequence of interest comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target RNA; (b) a first primer that is hybridizable to the target RNA, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion; (c) a second primer that is hybridizable to an extension product of the first primer; (d) an RNA-dependent DNA polymerase; (e) a DNA-dependent DNA polymerase; (f) an

enzyme that cleaves RNA from an RNA/DNA hybrid; wherein the incubation is under conditions that permit primer hybridization, extension from the first primer and the second primer to form a complex comprising first and second primer extension products, RNA cleavage, and displacement of the first primer extension product from the complex when its RNA is cleaved and a composite primer binds to the second primer extension product in the complex and is extended, whereby multiple copies of a polynucleotide sequence complementary to the RNA sequence of interest are generated.

In another aspect, the invention provides methods of generating multiple copies of (amplifying) an RNA sequence of interest, said method comprising the steps of: (a) hybridizing a first primer to a target RNA, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion; (b) extending the first primer with an RNA-dependent DNA polymerase, whereby a complex comprising a first primer extension product and the target RNA is produced; (c) cleaving RNA in the complex of step (b) with an agent (such as an enzyme) that cleaves RNA from an RNA/DNA hybrid; (d) hybridizing a second primer to the first primer extension product; (e) extending the second primer with a DNA-dependent DNA polymerase, whereby a second primer extension product is produced to form a complex of first and second primer extension products; (f) cleaving RNA from the composite primer in the complex of first and second primer extension products with an agent (such as an enzyme) that cleaves RNA from an RNA/DNA hybrid such that a composite primer (which may or may not be the same as the first primer) hybridizes to the second primer extension product, wherein said composite primer comprises an RNA portion and a 3' DNA portion; (g) extending said composite primer of step (f) with a DNA-dependent DNA polymerase whereby said first primer extension product is displaced; (h) hybridizing the displaced first primer extension product with a polynucleotide comprising a propromoter and a region which is hybridizable to the displaced first primer extension product under conditions which allow transcription to occur by RNA polymerase, such that RNA transcripts are produced comprising sequences complementary to the displaced primer extension products, whereby multiple copies of the RNA sequence of interest are generated. In some embodiments,

the invention provides methods of generating multiple copies of an RNA sequence of interest, said method comprising the steps of: (a) combining: a first complex, wherein the first complex is the complex of first and second primer extension products of step (e) described above in this paragraph; a composite primer (which may or may not be the same as the first primer) that is hybridizable to the second primer extension product, wherein said composite primer comprises an RNA portion and a 3' DNA portion; a DNA-dependent DNA polymerase; an RNA polymerase; a propromoter polynucleotide comprising a propromoter and a region which hybridizes to a second primer extension product; and an agent (such as an enzyme) that cleaves RNA from an RNA/DNA hybrid; and (b) incubating the mixture of step (a) under conditions (which include necessary substrates and buffer conditions) that permit primer hybridization, RNA cleavage, displacement of the first primer extension product from the first complex when its RNA is cleaved and a composite primer binds to the second primer extension product in the first complex, hybridization of the propromoter polynucleotide to the displaced first primer extension product to form a second complex comprising the displaced primer extension product and the propromoter polynucleotide, and RNA transcription from said second complex.

In another aspect, the invention provides methods of generating multiple copies of an RNA sequence of interest comprising incubating a reaction mixture, said reaction mixture comprising: (a) a target RNA; (b) a first primer that is hybridizable to the target RNA, wherein the first primer is a composite primer comprising an RNA portion and a 3' DNA portion; (c) a second primer that is hybridizable to an extension product of the first primer; (d) an RNA-dependent DNA polymerase; (e) a DNA-dependent DNA polymerase; (f) an RNA polymerase; (g) a propromoter polynucleotide comprising a propromoter and a region which hybridizes to an extension product of the second primer; and (h) an enzyme that cleaves RNA from an RNA/DNA hybrid; wherein the incubation is under conditions that permit primer hybridization, extension from the first primer and the second primer to form a first complex comprising first and second primer extension products, RNA cleavage, displacement of the first primer extension product from the first complex when its RNA is cleaved and a composite primer binds to the second primer extension product