



- (51) International Patent Classification:
C12Q 1/68 (2006.01) *C40B 20/04* (2006.01)
- (21) International Application Number:
PCT/US2012/048892
- (22) International Filing Date:
30 July 2012 (30.07.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/513,474 29 July 2011 (29.07.2011) US
61/601,514 21 February 2012 (21.02.2012) US
- (71) Applicant (for all designated States except US): **BIO-RAD LABORATORIES, INC.**, [US/US]; A Delaware Corporation, 1000 Alfred Nobel Drive, Hercules, CA 94547 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **SAXONOV, Serge** [US/US]; 6676 Heartwood Drive, Oakland, CA 94611 (US). **TZONEV, Svilen, S.** [US/US]; 121 Rocky Creek Place, Pleasanton, CA 94566 (US). **LUCERO, Michael, Y.** [US/US]; 634 Pine Terrace, South San Francisco, CA 94080 (US). **KOEHLER, Ryan** [US/US]; 1000 Alfred Nobel Drive, Hercules, CA 94547 (US). **HINDSON, Benjamin, J.** [AU/US]; 1039 Bannock Street, Livermore, CA 94551 (US).
- (74) Agent: **ABNEY, James, R.**; Kolisch Hartwell, PC, 520 SW Yamhill Street, Suite 200, Portland, Oregon 97204-1324 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report (Art. 21(3))

(54) Title: LIBRARY CHARACTERIZATION BY DIGITAL ASSAY

(57) Abstract: Methods of characterizing a nucleic acid library by digital assay.



WO 2013/019751 A1

LIBRARY CHARACTERIZATION BY DIGITAL ASSAY**Cross-References to Priority Applications**

This application is based upon and claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Patent Application Serial No. 61/513,474, filed July 29, 2011; and U.S. Provisional Patent Application Serial No. 61/601,514, filed February 21, 2012. Each of these priority applications is incorporated herein by reference in its entirety for all purposes.

Cross-References to Other Materials

This application incorporates by reference in their entireties for all purposes the following materials: U.S. Patent No. 7,041,481, issued May 9, 2006; U.S. Patent Application Publication No. 2010/0173394 A1, published July 8, 2010; U.S. Patent Application Publication No. 2011/0217712 A1, published September 8, 2011; U.S. Patent Application Publication No. 2012/0152369 A1, published June 21, 2012; U.S. Patent Application Serial No. 13/251,016, filed September 30, 2011; U.S. Patent Application Serial No. 13/548,062, filed July 12, 2012; and Joseph R. Lakowicz, PRINCIPLES OF FLUORESCENCE SPECTROSCOPY (2nd Ed. 1999).

Introduction

DNA sequencing determines the order of nucleotide bases in a DNA molecule. The ability to obtain sequence information quickly is crucial to many fields, such as biological research, clinical diagnostics, pharmacogenomics, forensics, and environmental studies. Due to the demand for improved sequencing technologies, the speed of sequence acquisition has increased dramatically over the past several decades.

The predominant first-generation sequencing technology is a chain-termination method developed by Frederick Sanger in 1977. The Sanger method performs a sequencing reaction for each sample in a separate reaction vessel and resolves reaction products according to size by electrophoresis in a gel or capillary. The ability to scale up the Sanger method for a very large number of samples is limited by the space and individual

manipulations needed for each sample (e.g., transferring the reacted sample from its reaction vessel to a gel or capillary).

Next-generation sequencing technologies, such as pyrosequencing (Roche Diagnostics), sequencing by synthesis (Illumina), and sequencing by
5 oligonucleotide ligation and detection (Life Technologies), overcome the major limitations of the first-generation approach. Sequencing reactions can be performed in parallel with a very large number of different samples (templates) immobilized in an array in the same flow cell. The density of samples per unit area can be very high, and the total number of samples can
10 be increased by enlarging the array. The samples can be exposed to a series of sequencing reagents in parallel in a shared fluid volume inside the flow cell. Also, the samples in the array can be monitored with a camera to record sequence data from all of the samples in real time as the sequencing reactions proceed in parallel with cyclical exposure to reagents passing
15 through the flow cell. Next-generation sequencing technologies are responsible for a dramatic increase in sequencing speed—orders of magnitude—over the past decade.

First-generation methods generally utilize conventional libraries to produce a sufficient amount of each template for sequencing. A first-
20 generation library may be composed of a collection of DNA fragments inserted into a vector, such as a plasmid or a bacteriophage vector. Each inserted fragment is cloned by placing the vector in a suitable host organism, such as a bacterium, which can replicate the vector and the fragment to make many clonal copies. In contrast, next-generation technologies increase
25 throughput dramatically by providing the capability to sequence in vitro libraries constructed exclusively in vitro by the action of one or more enzymes. In vitro libraries do not contain or need a vector for replication in vivo because each fragment is cloned by amplification in vitro, such as through the polymerase chain reaction (PCR). Accordingly, in vitro libraries can be
30 constructed from very small amounts of nucleic acid and permit sequencing of rare species (e.g., rare mutations) that occur at a very low frequency in a sample.

Next-generation technologies currently on the market rely on in vitro libraries having a particular construction. The various fragments to be sequenced are each flanked by adapters to form library members. The adapters provide primer binding sites for clonal amplification of each library member on a support, such as on a flat surface or beads. The adapters can introduce binding sites that enable amplification of all members of the library with the same primer or pair of adapter-specific primers. Also, one or both of the adapters can provide a binding site for a sequencing primer. Furthermore, an adapter can introduce a library-specific index sequence that permits members of different libraries to be pooled and sequenced together in the same flow cell, without losing track of the library of origin for each member.

A set of libraries can be constructed in parallel, such as in different wells of a multi-well plate, from different nucleic acid samples. However, despite the best efforts to achieve uniform reaction conditions among the wells, the concentration and quality of the libraries can vary widely.

Summary

The present disclosure provides methods of characterizing a nucleic acid library by digital assay.

Brief Description of the Drawings

Figure 1 is schematic representation of an exemplary library that may be characterized according to the present disclosure, with the library being constructed with a pair of different adapters and including well-formed and malformed members, in accordance with aspects of the present disclosure.

Figure 2 is a flowchart of selected aspects of an exemplary method of characterizing the library of Figure 1, in accordance with aspects of the present disclosure.

Figure 3 is a schematic illustration of selected aspects of a library characterization performed according to Figure 2 and exemplifying amplification data that can be collected from different reaction sites (e.g., discrete fluid volumes, such as distinct droplets), in accordance with aspects of the present disclosure.

Figure 4 is an exemplary reaction diagram illustrating an exemplary approach for constructing members of the library of Figure 1, in accordance with aspects of the present disclosure.

5 Figure 5 is another exemplary reaction diagram illustrating another exemplary approach for constructing members of the library of Figure 1, in accordance with aspects of the present disclosure.

Figure 6 is yet another exemplary reaction diagram illustrating yet another exemplary approach for constructing members of the library of Figure 1, in accordance with aspects of the present disclosure.

10 Figure 7 is a schematic representation of exemplary amplification primers and probes for use in a digital amplification assay to quantify library members containing both of the different adapters of Figure 1, in accordance with aspects of the present disclosure.

15 Figure 8 is a schematic representation of exemplary amplification primers and probes for use in a digital amplification assay to quantify empty and filled library members, in accordance with aspects of the present disclosure.

20 Figure 9 is a schematic representation of exemplary amplification primers, a probe, and an intercalating reporter for use in a digital amplification assay to quantify empty and filled library members, in accordance with aspects of the present disclosure.

25 Figure 10 is a plot of exemplary amplification data obtained with the methods of Figures 2 and 3 using a library constructed according to Figure 5 and tested with the primers and probes of Figure 7, in accordance with aspects of the present disclosure.

Figure 11 is a schematic representation of the plot of Figure 10, in accordance with aspects of the present disclosure.

30 Figure 12 is another plot of exemplary amplification data obtained with the methods of Figures 2 and 3 using a library constructed according to Figure 5 and tested with the primers and probes of Figure 7, in accordance with aspects of the present disclosure.

Detailed Description

The present disclosure provides methods of characterizing a nucleic acid library by digital assay.

5 An exemplary method of library characterization is provided. In the method, a nucleic acid library may be obtained. The library may include members each having a first adapter region and a second adapter region. At least a subset of the members may have an insert disposed between the first and second adapter regions. At least a portion of the library may be divided into partitions. A digital assay may be performed on the partitions with an
10 adapter region probe to generate data indicating whether a library member is present in each partition. A characteristic of the library may be determined based on the data.

Another exemplary method of library characterization is provided. In the method, a nucleic acid library may be obtained. The library may include
15 members each having a first constant region and a second constant region. At least a subset of the members may have a variable region disposed between the first and second constant regions. Droplets containing members of the library at limiting dilution may be formed. Members of the library may be amplified in the droplets using a primer for each constant region. Amplification
20 data may be collected from a constant region probe in the droplets. A level of members of the library may be determined based on the amplification data.

Library characterization before sequencing can be problematic. Only properly formed library members containing both adapters in the correct relative orientation produce clonal populations that can be interrogated
25 reliably by sequencing. Malformed members in the library, such as members flanked by two copies of only one of the adapters, can be difficult to distinguish from those that are well-formed. However, the malformed members generally cannot be amplified on a support, a prerequisite to sequence acquisition, or do not have a binding site for the sequencing primer,
30 or both. As a result, malformed members can take up space and consume reagents and can reduce the amount of useful sequence information

produced by a next-generation sequencing run, in direct proportion to the fraction of malformed members in the library.

The methods for library characterization disclosed herein may have numerous advantages over other approaches. These advantages may include
5 the ability to obtain more information about library quality (e.g., quantification of both well-formed and malformed library members, quantification of empty and filled library members, qualitative indication of library complexity, or the like), fewer sequencing runs wasted, increased speed, less library material used for analysis, and/or more accurate concentration estimates, among
10 others. Also, the ability to quantify well-formed library members enhances significantly the chance of optimal loading of libraries prior to sequencing.

Further aspects of the present disclosure are presented in the following sections: (I) library overview, (II) methods of library characterization, (III) library construction, (IV) primers and probes, and (V) examples.

15 **I. Library Overview**

Figure 1 shows an exemplary in vitro, nucleic acid library 40 that may be characterized according to the methods disclosed herein. Members of the library each may include one or more adapter regions 42, 44 ("A", "B"), which also or alternatively may be termed adapters or constant regions, and an
20 insert 46, which also or alternatively may be termed a variable region and/or a variable insert. Each insert may be disposed between adapter regions 42, 44, such that the insert is flanked by the adapter regions (i.e., attached at each opposing end to an adapter region). Inserts 46 of the library may be supplied by fragments, which may be attached at each end to an adapter that provides
25 one of the adapter/constant regions.

Inserts 46 of the library are of interest for sequencing analysis and may vary substantially in sequence among members of the library. For example, the inserts may provide a variable region corresponding to a diverse collection of fragments generated from a source material for a shotgun sequencing
30 strategy. However, in some cases, such as in a deep-sequencing approach that looks for rare mutations, the inserts may have a low frequency of

variability. In any event, the adapter regions provide binding sites for primers and/or probes at the ends of each well-formed member 48 of the library

Adapters that provide adapter regions 42, 44 may be attached to inserts 46 and to each other during library construction in various combinations to create desired, well-formed members 48 (only one is shown in Fig. 1) and malformed members, such as members 50-54. The well-formed members have the correct structure, with a different adapter region 42 or 44 attached to each end of the insert (i.e., "A" at one end and "B" at the other end), and in the correct relative orientation of the adapter regions. The well-formed library members, due to the presence of both adapter regions 42, 44 in the correct relative orientation, are capable of being amplified clonally on a solid support with a pair of adapter primers as a preparatory step in a sequencing protocol.

Malformed members of the library are formed incorrectly and may have a variety of different structures, such as those shown in Figure 1. The malformed members illustrated here are each capable of being amplified in solution, in the presence of the same pair of adapter primers that can amplify well-formed library members.

Malformed members 50, 52 have a copy of the same adapter region attached to each end of the insert (i.e., a copy of "A" at both ends or a copy of "B" at both ends). The copies may be arranged as inverted repeats (i.e., rotated 180 degrees relative to one another in the drawing), which is represented in Figure 1 by the rightward "A" copy and the leftward "B" copy being upside down and backwards in members 50 and 52, respectively. Furthermore, an empty library member 54 may be created if the different adapters attach directly to each other in the correct relative orientation but with no intervening insert. In any event, malformed members generally do not yield any useful sequencing data. For example, malformed members may not be amplifiable clonally on a primer-coated support (e.g., a primer-coated bead). Alternatively, or in addition, malformed members may lack the binding site for a sequencing primer used for the well-formed members, or may have more than one instance of the binding site, such that sequence reads are

superimposed on one other. Furthermore, malformed library members may not carry a sequence of interest (e.g., empty member 54). The proportion of malformed members in a library can vary substantially based, for example, on the integrity and concentration of the DNA fragments that provide inserts 46, the ratio of adapters to insert fragments used for ligation, the presence of inhibitors or other contaminants, and the like.

Inserts 46 may be formed with fragments of DNA, such as pieces of genomic DNA, mitochondrial DNA, chloroplast DNA, cDNA, or the like, from any suitable source. The fragments may have any suitable length, such as about 10 to 10,000, or 20 to 2,000 nucleotides, among others. The fragments may or may not be size-selected before attachment to the adapters. Fragments may be generated from a source nucleic acid material by any suitable approach, such as shearing, chemical digestion, enzymatic digestion, amplification with one or more primers, reverse transcription, end-polishing, or any combination thereof, among others. The fragments may have flush or overhanging ends, and may be at least predominantly double-stranded or single-stranded.

Each adapter (or adapter region) may have any suitable structure before and/or after attachment to inserts. The adapter before attachment may include a nucleic acid or nucleic acid analog. Each adapter may be formed by one or more oligonucleotide strands each having any suitable length, such as at least about 6, 8, 10, 15, 20, 30, or 40 nucleotides, among others, and/or less than about 200, 100, 75, or 50 nucleotides, among others. The adapter may be provided by one or more oligonucleotides that are chemically synthesized in vitro. The adapter may be configured to be attached to inserts at only one of its two ends. In some cases, the adapter may be partially or completely single-stranded before attachment to inserts, such as if the adapter is provided by a primer that attaches to inserts via primer extension.

Library 40 may include any suitable medium in which library members (such as members 48-54) are disposed. The medium may be an aqueous phase 56, which may include salt, buffer, surfactant, at least one enzyme

(e.g., ligase, polymerase, etc.), unligated adapters, one or more primers, one or more probes, or any combination thereof, among others.

II. Methods of Library Characterization

This section provides an overview of exemplary methods of characterizing a library containing inserts attached to adapters. The method steps disclosed in this section and elsewhere in the present disclosure may be performed in any suitable combination, in any suitable order, and any suitable number of times.

Figure 2 shows a flowchart of selected aspects of an exemplary method 60 of characterizing library 40 of Figure 1 before the library is sequenced. Method 60 may, for example, be performed before sequencing to determine how much of the library to use in a sequencing protocol (e.g., to prevent underloading or overloading) and/or to determine whether or not the library is of sufficient quality for the sequencing protocol. Library characterization also or alternatively may be performed for any other purpose.

Method 60 may be used to perform a digital assay on the library members. The digital assay relies on the ability to detect the presence of a single library member in individual partitions of the library. In an exemplary digital assay, at least a portion of a library is separated into a set of partitions, which may be of equal volume. The library may be separated at limiting dilution, with some of the partitions containing no library members and others containing only one library member. If the library members are distributed randomly among the partitions, some partitions should contain no members, others only one member, and, if the number of partitions is large enough, still others should contain two members, three members, and even higher numbers of members. The probability of finding exactly 0, 1, 2, 3, or more library members in a partition, based on a given average concentration of members in the partitions, is described by Poisson statistics. Conversely, the concentration of the members in the partitions (and in the library) may be determined from the probability of finding a given number of library members in a partition.

Estimates of the probability of finding no library members and of finding one or more library members may be measured in the digital assay. Each partition can be tested to determine whether the partition is a positive partition that contains at least one library member, or is a negative partition that contains no library members. The probability of finding no library members in a partition can be approximated by the fraction of partitions tested that are negative (the "negative fraction"), and the probability of finding at least one library member by the fraction of partitions tested that are positive (the "positive fraction"). The positive fraction (or, equivalently, the negative fraction) then may be utilized in a Poisson equation to determine the concentration of library members in the partitions.

Digital amplification assays may rely on amplification of templates (e.g., templates provided by library members) in partitions to enable detection of a single library member. Amplification may, for example, be conducted via PCR, to achieve a digital PCR assay. Amplification of the library members can be detected optically from a luminescent reporter included in the reaction. In particular, the reporter can include a luminophore (e.g., a fluorophore) that emits light (luminesces) according to whether or not a library member has been amplified in a given partition. The luminophore may emit light in response to illumination with suitable excitation light.

A digital PCR assay can be multiplexed to permit detection of two or more different types of templates or targets (e.g., different types of library members, such as well-formed and malformed members, empty and filled/total members, etc.) within each partition. Amplification of the different types of library members can be distinguished by utilizing target-specific reporters (e.g., probes) that are optically distinguishable. For example, the reporters may include distinct luminophores producing distinguishable luminescence that can be detected with different detection regimes, such as different excitation and/or detection wavelengths or wavebands and/or different detection times after excitation, among others. In some cases, different target-specific reporters can be distinguished based on intensity differences measured in the same detection channel.

In method 60, a library for characterization may be obtained, indicated at 62. The library may include members each having at least one adapter region or a pair of different adapter regions, which may be constant regions. Library members also may include inserts and/or a variable region, with at least some of the inserts each being attached at one end to a first adapter region and at the other end to a second adapter region. The library may be obtained by constructing the library or may be received from a third party. The library may be constructed, at least in part, by attaching adapters to fragments (e.g., a diverse collection of fragments), such as in the presence of a ligase enzyme. After attachment of adapters, the library may be pre-amplified any suitable amount before the library is partitioned, to increase the quantity of library material available for quantification, quality analysis, and/or sequencing. Alternatively, after attachment of adapters, the library may be partitioned without prior amplification.

In some cases, first and second adapter regions that opposingly flank inserts of the library may be provided by a compound adapter that is attached to both ends of the inserts. The compound adapter may have a double-stranded region and a pair of single-stranded regions, with the single-stranded regions each provided by a different strand and extending from the same end of the double-stranded region. One strand of the compound adapter may provide the first adapter region and a second strand of the compound adapter may provide the second adapter region of library members.

In other cases, the library may be constructed by contacting fragments with a first adapter and a second adapter. The first and second adapters may be discrete from each other and not substantially base-paired to each other.

In still other cases, the library may be constructed by linking adapters to inserts by primer-based amplification. For example, a pair of tailed primers may be used to generate an insert from a template, with the primers providing the first and second adapters.

Library construction also may include any suitable supplementary reactions or processes, such as end-filling, nick repair, conversion to single-

stranded form, purification, size selection, or the like. Further aspects of library construction are described below in Section III.

At least a portion of the library may be divided into partitions, indicated at 64. Partitions of the library may be distributed to a plurality of reaction sites.

- 5 The reaction sites may be movable or fixed relative to one another. The reaction sites may be formed by discrete fluid volumes isolated from one another by one or more walls and/or by a separating fluid (e.g., a continuous phase of an emulsion). Alternatively, the reaction sites may be provided by a continuous surface (such as reaction sites arrayed on the surface of a chip) or
- 10 beads, among others. The partitions may be distributed at a limiting dilution of members of the library, meaning that a plurality of the reaction sites do not receive a library member and/or such that a plurality of the reaction sites receive only one library member.

- In some embodiments, at least part of the library may be partitioned
- 15 into fluid volumes that serve as reaction sites. The fluid volumes may be isolated from one another by a fluid phase, such as a continuous phase of an emulsion, by a solid phase, such as at least one wall of a container, or a combination thereof, among others. In some embodiments, the fluid volumes may be droplets disposed in a continuous phase, such that the droplets and
- 20 the continuous phase collectively form an emulsion.

- The fluid volumes may be formed by any suitable procedure, in any suitable manner, and with any suitable properties. For example, the fluid volumes may be formed with a fluid dispenser, such as a pipet, with a droplet generator, by forceful mixing (e.g., shaking, stirring, sonication, etc.), and/or
- 25 the like. Accordingly, the fluid volumes may be formed serially, in parallel, or in batch. The fluid volumes may be of substantially uniform volume or may have different volumes. Exemplary fluid volumes having the same volume are monodisperse droplets. Exemplary volumes for the fluid volumes include an average volume of less than about 100, 10 or 1 μL , less than about 100, 10,
- 30 or 1 nL, or less than about 100, 10, or 1 pL, among others.

The fluid volumes, when formed, may be competent for performance of one or more reactions in the fluid volumes, such as amplification of library

members (and, optionally, associated probe degradation). Alternatively, one or more reagents may be added to the fluid volumes after they are formed to render them competent for reaction. The reagents may be added by any suitable mechanism, such as a fluid dispenser, fusion of droplets, or the like.

5 Library members may be amplified, indicated at 66, at the reaction sites (e.g., in the fluid volumes). For example, library members may be amplified with a primer for each adapter region flanking the insert. In some cases, a pair of primers may be used, with one of the primers binding to the first adapter region attached at one end of the inserts and the other primer
10 binding to the second adapter region attached at the other end of the inserts. In other examples, a single primer may be suitable for amplification, if the first and second adapter regions share a sequence that allows the same primer to bind to both adapter regions. In any event, amplification may be substantially restricted to constructs that contain a pair of primer binding sites arranged for
15 convergent extension into the insert from binding sites for a pair of primers (see Section IV). For example, if amplification of well-formed member 48 (see Fig. 1) is performed with a forward primer that binds to "A" and a reverse primer that binds to "B," each of malformed members 50-54 also can be amplified. However, other constructs, such as a construct having an insert
20 flanked by a direct (not inverted) repeat of "A" or "B" generally would not amplify with these primers. Additional or other primers may be utilized if these other constructs are to be amplified and detected.

 Amplification may be performed by any suitable reactions. For example, amplification may be performed by a polymerase chain reaction.
25 Alternatively, or in addition, amplification may be performed by a ligase chain reaction. In any event, the reaction sites (e.g., fluid volumes) may be thermally cycled to promote amplification.

 Amplification may be performed in the presence of one or more labeled reporters. For example, each reaction site (e.g., fluid volume) may include a
30 first labeled probe capable of binding the first adapter region (e.g., adapter A; see Fig. 1) and a second labeled probe capable of binding the second adapter region (e.g., adapter B). Each labeled probe may include a luminophore. For

example, the probe may include an energy transfer pair, namely, an energy donor (generally a luminophore) and an energy acceptor. The energy acceptor may be another luminophore or a quencher. The probe may produce a stronger signal when in a degraded form. In other cases, the reporter(s)
5 may include an intercalating luminophore and/or a probe that binds selectively to a junction sequence formed by direct attachment of different adapter regions to each other.

Amplification data may be collected from the reporters, including data collected from an adapter region probe, indicated at 68. The data may be
10 collected from individual reaction sites (e.g., fluid volumes such as droplets). The amplification data may indicate whether a library member (e.g., a particular type of library member) capable of binding the probe is present (and amplified) at a given reaction site. For example, the adapter region probe may be bound selectively by empty library members having no insert. As another
15 example, the adapter region probe may be bound selectively by amplified library members containing adapter A (or adapter B) (see Fig. 1). The data may, for example, be collected by detecting light from individual reaction sites.

Signals may be created that are representative of light detected from each reaction site. The signals may represent data collected in one or more
20 different channels (e.g., in different wavebands (color regimes)) from different luminophores representing amplification of different adapters. Alternatively, or in addition, the signals may represent an aspect of light, such as the intensity of the light, detected in the same channel (e.g., in the same waveband for two different adapter region probes). Further aspects of using the same detection
25 channel for detection of signals from at least a pair of different reporters or probes is described in U.S. Patent Application Serial No. 13/548,062, filed July 12, 2012, which is incorporated herein by reference.

Detection may be performed at any suitable time(s). Exemplary times include at the end of an assay (an endpoint assay), when reactions have run
30 to completion and the data no longer are changing, or at some earlier time (a kinetic assay).

At least one characteristic of the library may be determined, indicated at 70, based on the amplification data collected. The characteristic may be an absolute or relative level (e.g., concentration) of library members (e.g., a type of library member). For example, the level may describe library members
5 containing a copy of each different adapter region, with the adapter regions having a defined orientation relative to one another and/or relative to the insert, namely, the relative orientation that permits productive amplification with the forward and reverse, adapter-specific primer(s) utilized at 66. The level determined may include or substantially exclude empty library members
10 having the first adapter region attached to the second adapter region without an intervening fragment. In other examples, the characteristic may be a quality metric of the library, which may, for example, be a measure of library complexity or a measure of the proportion of well-formed members in the library.

15 In some cases, the amplification data collected at 68 may be processed to determine whether individual reaction sites test positive or negative for amplification of a template or a particular type of template. Each of a plurality of reaction sites may be designated as being amplification-positive or amplification-negative for a first adapter region, for a second
20 adapter region, for a junction produced by direct attachment of the first and second adapter regions to each other, for amplified nucleic acid (e.g., when the reporter is an intercalating luminophore), or any combination thereof, among others. A reaction site may be designated as positive or negative for each of the adapters by comparing, to one or more thresholds or ranges, an
25 adapter signal strength (from an adapter-specific probe) for each adapter, from individual reaction sites. A relative or absolute level of each different adapter region may be determined based on the number and/or fraction of positive or negative reaction sites. The calculation may be based on library members containing both adapters having a Poisson distribution among the
30 reaction sites. The level of library members determined may account for occurrence of an amplifiable first adapter region and an amplifiable second

adapter region at the same reaction site by chance, without being linked to one another (see Example 2).

The library characteristic determined may be used to guide library processing, such as to determine whether or not the library is of sufficient quality for sequencing. For example, if the level of empty library members and/or other malformed members is too high, the apparent complexity of well-formed member is too low, or the like, a replacement library may need to be constructed. Alternatively, if the library is of sufficient quality, an amount of the library may be selected for sequencing, indicated at 72, based on the characteristic determined at 70. The amount of library selected may be contacted with a solid support(s) (e.g., a continuous surface or beads) in preparation for clonal amplification of library members on the solid support(s). Individual members of the library may be bound to distinct supports, such as distinct beads (e.g., as in pyrosequencing), or to spaced regions of the same support surface (e.g., as in bridge amplification prior to sequencing by synthesis (Illumina)). For example, the supports/support surface may be pre-attached to many copies of a capture sequence/primer that is complementary to an adapter sequence at the end of each well-formed library member. The supports may be disposed in droplets of an emulsion (e.g., as in pyrosequencing), with each droplet containing, on average, only about one or less well-formed member of the library. The library members may be clonally amplified on the solid support(s) and then may be sequenced by any suitable technology, such as pyrosequencing (Roche Diagnostics), sequencing by oligonucleotide ligation and detection (Life Technologies), sequencing by synthesis (Illumina), or the like.

Knowledge of the library concentration for well-formed members minimizes overloading or under-loading the solid support(s) with library members. With overloading, two or more well-formed library members may be attached to the same support or surface region, resulting later in superimposed sequencing reads, which reduces the amount of useful sequencing information obtained from the run. With under-loading, no library

members may be attached to the support or surface region, which lies fallow during the sequencing run.

Figure 3 schematically depicts exemplary execution of selected aspects of the method 60 of Figure 2, and exemplifies amplification data that
5 can be collected in a library characterization from different reaction sites.

Library 40 of Figure 1 is depicted at the left of Figure 3, with the variable insert indicated by "V." Portions of the library may be distributed at limiting dilution to a plurality of reaction sites, such as droplets 73. Here, only six exemplary droplets 73 are shown. However, the library may be distributed
10 to hundreds, thousands, or more reaction sites according to the statistical accuracy desired.

Each reaction site may receive none or one or more library members. For example, four of the droplets have received either one copy of a well-formed member 48, one copy of a malformed member 50 with two A
15 adapters, one copy of a malformed member 52 with two B adapters, or one copy of an empty member 54. Since the library has been partitioned at limiting dilution, a subset of the reaction sites, such as empty droplet 74, receive no members of the library. Another subset of the reaction sites may receive two or more library members. Here, for example, droplet 75 contains two library
20 members, namely, library member 50 and library member 52.

The reaction sites may contain reagents for amplification of the library members. For example, each reaction site may contain a forward primer 76 and a reverse primer 77 for amplification of well-formed library members 48 (and malformed members 50-54). The reaction site also may include all of the
25 other reagents necessary to promote amplification of the library members and amplification detection, such as dNTPs, a labeled probe for each adapter (e.g., see Fig. 7), an amplification enzyme (e.g., a polymerase, such as a heat-stable polymerase), buffer, salt, etc.

Amplification of the library members may be detected to collect
30 amplification data from adapter-specific probes at the reaction sites. The data may be processed to identify each reaction site as double negative, A-only positive, B-only positive, or AB double positive. The different types of double

positives show in Figure 3 may or may not be distinguishable from each other according to signal strength. Further aspects of collecting and processing amplification data are described below in Example 2.

Further aspects of libraries and further aspects of digital assays, such as generating emulsions and droplets, performing nucleic acid amplification at reaction sites, and collecting and processing amplification data, are described in the materials listed above under Cross-References, which are incorporated herein by reference, particularly U.S. Patent Application Publication No. 2010/0173394 A1, published July 8, 2010; U.S. Patent Application Publication No. 2011/0217712 A1, published September 8, 2011; U.S. Patent Application Publication No. 2012/0152369 A1, published June 21, 2012; U.S. Patent Application Serial No. 13/251,016, filed September 30, 2011; and U.S. Patent Application Serial No. 13/548,062, filed July 12, 2012.

III. Library Construction

This section describes selected, exemplary aspects of library construction; see Figures 4-6.

Figure 4 shows an exemplary reaction strategy 80 for constructing members of the library of Figure 1. A collection of adapter-less fragments 82 may be contacted with unlinked adapters 84, 86 and a ligase enzyme. The unlinked adapters become attached adapters that provide adapter regions 42, 44 that flank inserts 46 formed from fragments 82, to create well-formed library members 48, among others.

Figure 5 shows another exemplary reaction strategy 100 for constructing members of the library of Figure 1. Here, rather than using different unlinked adapters 84, 86 (see Fig. 4), the same compound adapter 102 is attached to both ends of modified fragments 104 having an extension 106. The extension may, for example, be a single nucleotide (e.g., an "A") added to both ends of fragments 82 by an adenylation reaction.

Compound adapter 102 has a general Y shape formed by a pair of oligonucleotides 108, 110. The oligonucleotides are complementary to each other at only one end of the adapter, to create a double-stranded region 112 ("C") and a pair of single-stranded regions corresponding to adapters A and B

(84, 86 of Fig. 4). Double-stranded region 112 may end with an overhang 114 of one of more nucleotides, such as a single nucleotide (e.g., a "T") complementary to extension 106. With this structure, the compound adapter can ligate efficiently to modified fragments 104 but not to one another.

5 Ligation may produce library member precursors 116 having forked ends. The precursors may be resolved into well-formed library members 48 by amplification with adapter-specific primers (or by denaturation alone). Double-stranded region 112 may be present in the library members as an inverted repeat, as shown, forming a common sequence of each adapter and
10 optionally providing one or more binding sites for one or more sequencing primers (and/or one or more probes).

 Figure 6 shows yet another exemplary approach for constructing members of the library of Figure 1. A polynucleotide 120 carrying a target region 122 may serve as a template for target-specific amplification with a pair
15 of primers 124, 126 that function as a forward primer and a reverse primer, respectively. Each primer may be described as a tailed primer or fusion primer capable of binding at a border of target 122. The primer also may provide an adapter region 128 or 130 that is linked to target region 122 by extension of
20 each primer during target amplification to produce well-formed library members 48. A library generated via tailed primers may have a substantially constant insert sequence and may be sequenced to identify variants in the target sequence, such as rare mutations.

IV. Primers and Probes

 This section describes exemplary primer and probe configurations for
25 library characterization according to Figures 2 and 3; see Figure 7.

 Figure 7 shows a schematic representation of exemplary amplification primers and probes for use in a digital amplification assay to quantify library members containing both of the different adapter regions of Figure 1. The basic idea is to bind the adapter regions of the library members with two
30 probes, with amplification spanning each insert. By using probes for both adapters A and B (see Fig. 1), one can measure the concentrations of well-formed library members and malformed library members, like those of Figure

1 or other combinations. This design ensures that the presence of different adapters in the same fluid volume can be tested directly for linkage. (Co-occupancy may also occur by chance without adapter linkage; see Figure 3 and Example 2.) With probes to each different adapter region, one can ensure
5 identification of well-formed library members (and ill-formed libraries).

Each adapter region 42, 44 and inserts 46 may be amplified from well-formed library members 48 with one or more adapter region primers that provide a forward primer ("FP") and a reverse primer ("RP") each extendable toward insert 46 and the other adapter region. In particular, the primers may
10 be positioned such that amplification products span the entire insert 46 and at least a flanking section of each adapter. The primers shown here also may be capable of amplifying each of the malformed library members of Figure 1. For example, primer "FP" can bind in an amplification-competent arrangement to both ends of malformed product 50, and primer "RP" can bind similarly to both
15 ends of malformed product 52.

Amplification products generated with the primers may be detected and distinguished with adapter-specific probes 140, 142 containing oligonucleotides 144, 146 that recognize adapter regions 42, 44, respectively. Each adapter probe may bind to the same strand of the products (e.g., both
20 binding to the sense strand or both binding to the antisense strand) or may bind to different strands (e.g., probe 140 binding to the sense strand and probe 142 to the antisense strand, or vice versa). Each adapter region probe may (or may not) bind closer to an adjacent insert/adapter junction than the corresponding adapter region primer.

25 Probes 140, 142 each may include a distinct luminophore (e.g., a fluorophore), labeled here as "L₁" and "L₂." The probe also may include a quencher, labeled here as "Q." The quencher may reduce detectable light emission from the luminophore in a proximity-dependent manner. Degradation of the probe, which may occur as a result of amplification, unlinks the
30 luminophore from the quencher, providing an increase in the signal detected. L₁ and L₂ may emit light of different wavelengths relative to each other, to provide optically distinguishable emission from the respective luminophores.

Alternatively, light emission may be detected in the same channel if the luminophores produce signals of distinguishable strength, as described in U.S. Patent Application Serial No. 13/548,062, filed July 12, 2012, which is incorporated herein by reference. In any event, detection of a probe may include detecting light emitted by the intact probe and/or a degraded form thereof (e.g., the released luminophore).

V. Examples

The following examples describe selected aspects and embodiments of library characterization by digital assay. These examples are intended for illustration only and should not limit the entire scope of the present disclosure.

Example 1. Exemplary Reporter Configurations for Detecting Empty Members

This example describes exemplary reporter configurations for detecting empty (and/or empty and filled) members of a library; see Figures 8 and 9.

Figure 8 shows a schematic representation of exemplary amplification primers (FP, RP) and probes 140, 160 bound to empty library member 54 and well-formed library member 48. Probe 140 binds to both types of library members and thus allows determination of the collective level of both types. In contrast, probe 160 includes an oligonucleotide 162 that binds selectively to empty members of the library having a junction sequence produced by direct attachment of adapter region 42 to adapter region 44. Accordingly, data collected from probe 160 can be used to determine a level of empty library members. In some cases, the level of empty library members may be compared to a level of amplified library members detected with probe 140.

Figure 9 shows a schematic representation similar to that of Figure 8, except that probe 140 is replaced by an intercalating reporter 180 that binds double-stranded nucleic acid. The reporter includes a luminophore ("L₃") that has altered luminescence (e.g., stronger light emission) when the reporter is bound to double-stranded nucleic acid. The reporter may, for example, be ethidium bromide, SYBR Green dye, SYBR Gold dye, Oxazole Yellow (YO) dye, Thiazole Orange (TO) dye, PicoGreen (PG) dye, or the like. An intercalating reporter may be useful to characterize insert sizes of the library,

since, in some configurations, the signal intensity produced by the intercalating reporter may be proportional to the size of amplicon produced at a reaction site. In some cases, the level of empty library members determined with data from probe 160 may be compared to the total level of amplifiable library members determined with data from reporter 180. In some cases, the level of amplifiable library members containing first and second adapter regions, as determined with probes that bind to each of these regions, may be compared to the total level determined with reporter 180.

Example 2. Exemplary Amplification Data and Calculations

This example describes exemplary amplification data and calculations for a library characterization; see Figures 10-12.

Figure 10 shows a two-dimensional plot of exemplary amplification data, in arbitrary intensity units, obtained from droplets with the method of Figures 2 and 3, using a library constructed according to Figure 5, and tested generally according to the primers and probes of Figure 7. Amplification of adapter sequences in the droplets was detected with the dye FAM as L_1 and the dye VIC as L_2 (see Fig. 7). Accordingly, the presence of an amplified adapter A sequence in a droplet generated a stronger FAM signal, the presence of an amplified adapter B sequence generated a stronger VIC signal, and the presence of both A and B sequences generated stronger FAM and VIC signals detected for the same droplet.

In the plot, data detected for each droplet is represented by a data point. Five clusters or populations of data points are circled and labeled for the plot according to the type(s) of adapter amplified in the corresponding droplet and/or, if both A and B are amplified, whether or not they flank an insert. The five types are as follows: (1) double-negatives (A^-B^-); (2) A-only positives (A^+B^-); (3) B-only positives (A^-B^+); (4) filled double-positives (A^+B^+ , with insert); and (5) empty double-positives (A^+B^+ , no insert) (see Fig. 3). In some cases, the double-positives may not be resolvable into a “filled” cluster and an “empty” cluster of data points.

The size and/or shape of the cluster formed by the double-positives, and particularly the filled double-positives, may be used as a quality metric for

the library. Here, the cluster extends with a positive slope along a diagonal on the plot. The cluster is spread out over a wide range of signal strength, suggesting substantial variability in the amplification efficiency of library members in the droplets. Amplification efficiency will vary among library members having a variable insert depending on length, GC content, etc. of the insert. Accordingly, the “effective” distribution of inserts can be detected qualitatively by the signal level of the double positives. In other words, in the present example, the heterogeneity of inserts in the library members translates into different signal strengths for double-positives on the plot, indicating that the complexity of the library is high. In contrast, positives generally form a tight cluster, if the same sequence is being amplified in each droplet. Accordingly, the range of signal strengths of the double-positives, the tightness with which they cluster, the shape of the cluster, or the like, may provide a quality metric indicating whether library construction was sufficiently successful. Library quality thus can be assessed in the digital assay disclosed herein without additional testing. Assay optimization may be based on the separability of different insert populations.

Figure 11 is a schematic representation of the data of Figure 10, with the filled and empty double-positives not resolved from each other. “N” is the number of fluid volumes (droplets) observed in each cluster that are double-negative (N_0), positive for A only (N_A), positive for B only (N_B), or positive for both A and B (N_{AB}^{obs}).

Even when there is no linkage between the A and B adapters, the number of fluid volumes expected to be positive for both A and B by chance (N_{AB}^{ch}) is given by the following equation:

$$N_{AB}^{ch} = \frac{N_A N_B}{N_0}$$

The number of volumes due to linkage of A and B (N_{AB}^{link}) can be calculated as the difference between the total observed number of counts for double positives and the number of double-positives expected by chance, as follows:

$$N_{AB}^{obs} - N_{AB}^{ch} = N_{AB}^{link}$$

The concentration of linked A and B adapters (i.e., the concentration of library members containing both adapters, (λ_{AB})) then can be calculated using the total number of volumes (N_{tot}) as follows:

$$\lambda_{AB} = \ln(N_{tot}) - \ln\left(N_0 + N_A + N_B + \frac{N_A N_B}{N_0}\right)$$

Figure 12 shows another plot of exemplary amplification data obtained with the methods of Figures 2 and 3 using a library constructed according to Figure 5 and tested with the primers and probes of Figure 7. Droplet clusters are labeled generally as in Figure 10.

Example 3. Selected Embodiments I

This example describes selected embodiments related to library analysis to quantify empty library members resulting from fusion of adapter sequences to each other in the absence of an intervening insert sequence during library construction, presented without limitation as a series of numbered paragraphs.

1. A method of quantifying empty members of a library, comprising:

(A) selecting a library of members formed by ligation of at least one type of adapter with a collection of fragments, each library member having either (a) an insert from the collection flanked at each end by an adapter sequence or (b) a pair of adapter sequences ligated to each other without an intervening insert from the collection, to generate an empty member of the library; (B) determining a total level of the library members in a library sample based on amplification; (C) determining a level of the empty members in a library sample based on amplification; and (D) comparing the level of the empty members to the total level of the library members.

2. The method of paragraph 1, further comprising a step of amplifying at least a region of the library members having an insert and at least a region of the library members lacking an insert using the same pair of primers.

3. The method of paragraph 2, wherein the at least one type of adapter includes a pair of adapter types, namely, a first adapter and a second adapter.

4. The method of paragraph 3, wherein each empty member includes a sequence from the first adapter and a different sequence from the second adapter.

5. The method of paragraph 3, wherein each library member
5 having an insert includes a sequence from the first adapter and a different sequence from the second adapter.

6. The method of paragraph 3, wherein one of the primers of the pair corresponds to the first adapter and the other primer of the pair corresponds to the second adapter.

10 7. The method of paragraph 1, wherein the step of determining a total level of the library members includes a step of detecting light from a first reporter, and wherein the step of determining a level of the empty members includes a step of detecting light from a second reporter that is different from the first reporter.

15 8. The method of paragraph 7, wherein the first reporter includes an intercalating dye.

9. The method of paragraph 7, wherein the first reporter includes a fluorophore attached to an oligonucleotide, and wherein the oligonucleotide corresponds to an adapter sequence.

20 10. The method of paragraph 7, wherein the second reporter includes a fluorophore attached to an oligonucleotide, and wherein the oligonucleotide binds selectively to a junction sequence formed by ligation of a pair of adapter sequences to one another.

11. The method of paragraph 1, further comprising a step of
25 amplifying at least a region of empty members using a primer that binds selectively to a junction sequence formed by ligation of a pair of adapter sequences to one another.

12. The method of paragraph 1, wherein the steps of determining a total level of library members and a level of empty members are performed by
30 kinetic analysis of amplification.

13. The method of paragraph 12, wherein the kinetic analysis includes real-time PCR.

14. The method of paragraph 1, wherein the steps of determining a total level of the library members and a level of the empty library members are performed by endpoint analysis of amplification in a digital amplification assay.

5 15. The method of paragraph 1, wherein the amplification is performed in droplets of at least one emulsion.

16. The method of paragraph 15, wherein at least one of the droplets contains no library members and at least one of the droplets contains only one library member.

10 17. The method of paragraph 16, wherein the droplets contain an average of less than about two library members per droplet.

18. The method of paragraph 16, wherein the droplets contain an average of less than about one library member per droplet.

15 19. The method of paragraph 15, wherein each step of determining includes a step of detecting light from individual droplets.

20. The method of paragraph 1, wherein each step of determining includes a step of detecting light emitted from a fluorophore, and wherein the fluorophore used in determining the total level is different than the fluorophore used in determining the level of empty members.

20 21. The method of paragraph 1, wherein the steps of determining both include a step of collecting amplification data from the same continuous aqueous phase or from the same droplets.

22. The method of paragraph 1, wherein the total level and the level of empty members are concentrations.

25 23. The method of paragraph 1, wherein the step of comparing the level of empty members to the total level includes a step of determining a ratio of levels.

24. A method of quantifying empty ligation products in a library, comprising: (A) selecting a library of ligation products formed by (a) ligation of
30 at least one type of adapter to a collection of fragments, to generate ligation products having inserts from the collection flanked at each end by an adapter, and (b) ligation to each other without an intervening insert, to generate empty

ligation products; (B) determining a total concentration of the ligation products in a library sample based on amplification of the ligation products; (C) determining a concentration of the empty ligation products in a library sample based on amplification of the empty ligation products; and (D) comparing the
5 concentration of the empty ligation products to the total concentration of ligation products.

25. The method of paragraph 24, wherein the step of comparing includes a step of determining a ratio of the concentrations.

Example 4. *Selected Embodiments II*

10 This example describes selected embodiments related to library characterization by digital assay, presented without limitation as a series of numbered paragraphs.

1. A method of library analysis, comprising: (A) obtaining a library including inserts attached at each end to a first adapter or a second adapter;
15 (B) distributing portions of the library at limiting dilution to a plurality of reaction sites; and (C) performing a digital amplification assay at individual reaction sites to determine a level of library members that contain both the first adapter and the second adapter.

2. A method of library analysis, comprising: (A) obtaining a library
20 including inserts attached at each end to a first adapter or a second adapter; (B) distributing portions of the library at limiting dilution to a plurality of reaction sites; (C) amplifying inserts at the reaction sites with one or more primers that bind to the first adapter and the second adapter; (D) collecting amplification data indicating a presence or absence of an amplified first
25 adapter sequence and a presence or absence of an amplified second adapter sequence at individual reaction sites; and (E) determining from the amplification data a level of library members that contain both the first adapter and the second adapter in a defined relative orientation.

3. A method of library analysis, comprising: (A) obtaining a library
30 including fragments flanked at each end by a first adapter or a second adapter; (B) partitioning the library at limiting dilution into fluid volumes; (C) amplifying fragments and at least part of each adapter in the fluid volumes

with one or more primers that bind to the first adapter and the second adapter;
(D) collecting amplification data indicating a presence or absence of an
amplified first adapter sequence and a presence or absence of an amplified
second adapter sequence in the same individual fluid volumes; and (E)
5 determining from the amplification data a level of library members that contain
both the first adapter and the second adapter.

4. The method of any of paragraphs 1 to 3, wherein the library is
obtained by constructing a library.

5. The method of paragraph 4, wherein the step of constructing a
10 library includes a step of attaching one or more adapter sequences to
fragments in the presence of a ligase enzyme.

6. The method of paragraph 5, wherein the step of attaching one or
more adapter sequences includes a step of contacting the fragments with a
compound adapter that provides both the first adapter and the second
15 adapter.

7. The method of paragraph 6, wherein the compound adapter has
a double-stranded region and a pair of single-stranded regions, and wherein
the single-stranded regions are each provided by a different strand and
extend from the same end of the double-stranded region.

8. The method of paragraph 6 or 7, wherein a first strand of the
20 compound adapter provides the first adapter and a second strand of the
compound adapter provides the second adapter.

9. The method of paragraph 5, wherein the step of attaching one or
more adapter sequences includes a step of contacting fragments/inserts with
25 a first adapter and a second adapter that are discrete from each other and not
substantially base-paired to each other.

10. The method of paragraph 4, wherein the step of constructing a
library includes a step of linking an adapter-specific sequence to
fragments/inserts by DNA synthesis.

11. The method of paragraph 10, wherein the step of constructing a
30 library includes (a) a step of contacting fragments/inserts with a pair of tailed
primers that bind to the fragments/inserts and provide the first and second

adapters, and (b) a step of copying the fragments/inserts by extending the tailed primers.

12. The method of any of paragraphs 1 to 3, wherein the library is received from a third party.

5 13. The method of any of paragraphs 1 to 12, further comprising a step of amplifying the library after the step of obtaining and before the step of distributing/partitioning.

10 14. The method of any of paragraphs 1 to 13, wherein the step of distributing/partitioning includes a step of generating droplets from a bulk volume containing at least a portion of the library.

15 15. The method of paragraph 14, wherein the step of generating droplets includes a step of forming an emulsion including the droplets disposed in a continuous phase.

15 16. The method of any of paragraphs 1 to 13, wherein the step of distributing/partitioning includes a step of disposing fluid volumes in a holder having an array of predefined sites each configured to receive a single fluid volume.

17. The method of paragraph 16, wherein each site is a well of the holder.

20 18. The method of paragraph 16 or 17, further comprising a step of contacting the fluid volumes with a same continuous liquid phase after the step of disposing the fluid volumes in the holder, wherein the fluid volumes are optionally immiscible with the liquid phase.

25 19. The method of any of paragraphs 1 to 18, wherein the step of distributing/partitioning the library is performed at a limiting dilution such that a plurality of the reaction sites/fluid volumes do not contain a member of the library.

20. The method of any of paragraphs 1 to 19, wherein the step of amplifying includes a step of performing a polymerase chain reaction.

30 21. The method of any of paragraphs 1 to 20, wherein the step of amplifying includes a step of thermally cycling the reaction sites/fluid volumes.

22. The method of any of paragraphs 1 to 21, wherein the step of amplifying includes a step of performing a ligase chain reaction.

23. The method of any of paragraphs 1 to 22, wherein the step of amplifying is performed with a first primer that binds to the first adapter and a
5 second primer that binds to the second adapter.

24. The method of any of paragraphs 1 to 22, wherein the step of amplifying is performed with a primer that binds to the first adapter and also binds to the second adapter.

25. The method of any of paragraphs 21 to 24, wherein the step of
10 amplifying is performed in the presence of a first labeled probe corresponding to the first adapter and a second labeled probe corresponding to the second adapter.

26. The method of paragraph 25, wherein each of the first labeled probe and the second labeled probe includes a fluorophore.

15 27. The method of paragraph 26, wherein each of the first labeled probe and the second labeled probe is a FRET probe.

28. The method of any of paragraphs 1 to 27, wherein the step of collecting amplification data includes a step of detecting fluorescence from reaction sites/fluid volumes.

20 29. The method of any of paragraphs 1 to 28, wherein the step of determining includes a step of designating each of a plurality of reaction sites/fluid volumes as positive or negative for the first adapter and as positive or negative for the second adapter.

30. The method of paragraph 29, wherein the step of designating
25 includes a step of comparing signal strengths for the first adapter and the second adapter from individual reaction sites/fluid volumes to one or more thresholds.

31. The method of any of paragraphs 1 to 30, wherein the level
30 determined includes a correction for occurrence of the first adapter and the second adapter in the same reaction site/fluid volume without being linked to one another.

32. The method of any of paragraphs 1 to 31, wherein the level determined is a concentration of the library members.

33. The method of any of paragraphs 1 to 32, wherein the library also includes empty members having the first adapter attached to the second
5 adapter without an intervening insert/fragment.

34. The method of paragraph 33, wherein the level includes empty members.

35. The method of paragraph 33, wherein the level substantially excludes empty members, further comprising an optional step of determining
10 a level of empty members in the library from the amplification data.

36. The method of any of paragraphs 1 to 35, wherein the step of determining is based on a Poisson distribution for members of the library among the reaction sites/fluid volumes.

37. The method of any of paragraphs 1 to 36, wherein the
15 amplification data includes first signal data detected from a first fluorophore corresponding to the first adapter and second signal data detected from a second fluorophore corresponding to the second adapter, further comprising a step of plotting the first signal data against the second signal data to generate a plot.

20 38. The method of paragraph 36, further comprising a step of analyzing a distribution of points representing the library members in the plot.

39. The method of paragraph 38, wherein the step of analyzing a distribution of points includes a step of determining a range of the points, a linear correlation of the points, clustering of the points into one or more
25 groups, or a combination thereof.

40. The method of paragraph 38, further comprising a step of assigning a quality metric to the library based at least in part on one or more results from the step of analyzing a distribution of points.

41. The method of any of paragraphs 1 to 40, further comprising a
30 step of assigning a quality metric to the library based on the level of library members that contain both adapters relative to a level of library members that have only one of the adapters.

42. The method of any of paragraphs 1 to 41, further comprising a step of comparing the level of library members that contain both adapters to a level of library members that have only one of the adapters.

43. The method of any of paragraphs 1 to 42, further comprising a
5 step of sequencing members of the library.

44. The method of any of paragraphs 1 to 43, further comprising a step of selecting an amount of the library for sequencing based on the level determined.

The disclosure set forth above may encompass multiple distinct
10 inventions with independent utility. Although each of these inventions has been disclosed in its preferred form(s), the specific embodiments thereof as disclosed and illustrated herein are not to be considered in a limiting sense, because numerous variations are possible. The subject matter of the inventions includes all novel and nonobvious combinations and
15 subcombinations of the various elements, features, functions, and/or properties disclosed herein. The following claims particularly point out certain combinations and subcombinations regarded as novel and nonobvious. Inventions embodied in other combinations and subcombinations of features, functions, elements, and/or properties may be claimed in applications claiming
20 priority from this or a related application. Such claims, whether directed to a different invention or to the same invention, and whether broader, narrower, equal, or different in scope to the original claims, also are regarded as included within the subject matter of the inventions of the present disclosure. Further, ordinal indicators, such as first, second, or third, for identified
25 elements are used to distinguish between the elements, and do not indicate a particular position or order of such elements, unless otherwise specifically stated.

WE CLAIM:

1. A method of library characterization, comprising:
 - obtaining a nucleic acid library including members each having a first
5 adapter region and a second adapter region, wherein at least a subset of the
members have an insert disposed between the first and second adapter
regions;
 - dividing at least a portion of the library into partitions;
 - performing a digital assay on the partitions with an adapter region
10 probe to generate data indicating whether a library member is present in each
partition; and
 - determining a characteristic of the library based on the data.
2. The method of claim 1, wherein the step of obtaining a nucleic
15 acid library includes a step of attaching at least one type of adapter to
opposing ends of nucleic acid fragments corresponding to the insert.
3. The method of claim 1, wherein the partitions are uniform in
size, and wherein some of the partitions contain no library members and
20 others of the partitions contain only one library member.
4. The method of claim 1, wherein the step of dividing includes a
step of forming droplets containing members of the library.
- 25 5. The method of claim 1, wherein the step of performing a digital
assay includes a step of amplifying members of the library in the partitions
with one or more primers that bind to the first adapter region and the second
adapter region.
- 30 6. The method of claim 1, wherein the step of performing a digital
assay includes a step of detecting luminescence from a luminophore of the
probe.

7. The method of claim 6, wherein the luminescence is at least predominantly produced by a degraded form of the probe.

8. The method of claim 1, wherein the probe binds selectively to
5 empty members of the library having no insert between the adapter regions, relative to members of the library having an insert.

9. The method of claim 1, wherein the step of determining a characteristic includes a step of determining a level of empty members of the
10 library having no insert between the adapter regions.

10. The method of claim 9, wherein the step of determining a characteristic includes a step of determining a level of members of the library that include an insert between the adapter regions.

15

11. The method of claim 9, wherein the probe binds selectively to a junction sequence produced by direct attachment of the first and second adapter regions to one another without an insert.

12. The method of claim 1, wherein the step of performing a digital assay uses a first probe that binds to the first adapter region and a second probe that binds to the second adapter region.

13. The method of claim 12, wherein the step of determining a
25 characteristic includes a step of determining a level of library members containing both the first adapter region and the second adapter region.

14. The method of claim 13, wherein the level determined accounts for a presence of the first and second adapter regions in the same partitions
30 without being linked to one another.

15. The method of claim 13, wherein the step of determining a characteristic also includes a step of determining a level of library members containing the first adapter region and not the second adapter region and a level of library members containing the second adapter region and not the first adapter region.

16. The method of claim 13, wherein the level substantially excludes empty members, further comprising a step of determining a level of empty members from the data.

17. The method of claim 1, further comprising a step of selecting an amount of the library for use in a sequencing protocol based on the characteristic.

18. A method of library characterization, comprising:
obtaining a nucleic acid library including members each having a first constant region and a second constant region, wherein at least a subset of the members have a variable region disposed between the first and second constant regions;
forming droplets containing members of the library at limiting dilution;
amplifying members of the library in the droplets using a primer for each constant region;
collecting amplification data from a constant region probe in the droplets; and
determining a level of members of the library based on the amplification data.

19. The method of claim 18, wherein the level a level of empty members of the library.

20. The method of claim 18, wherein the level is a level of members of the library having both constant regions.

1/8

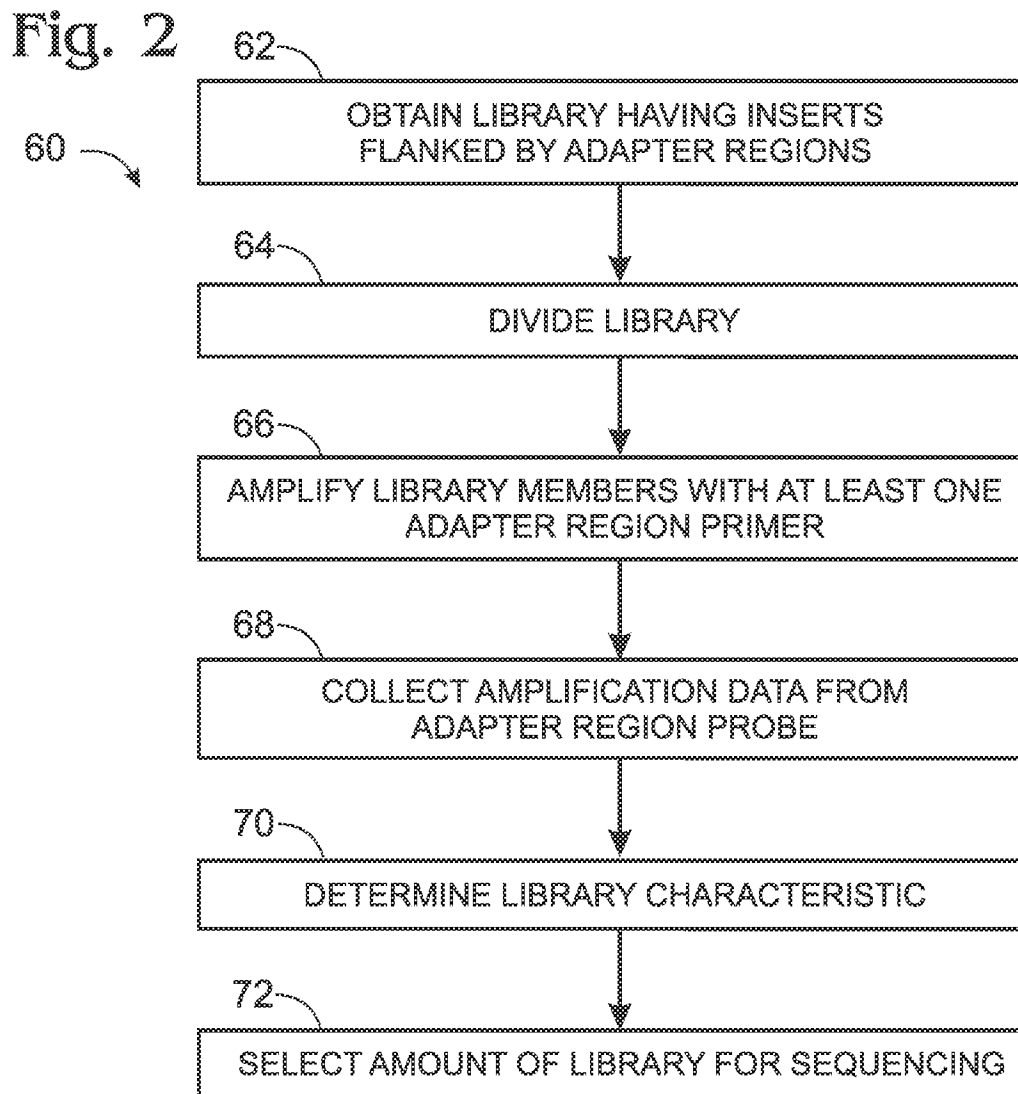
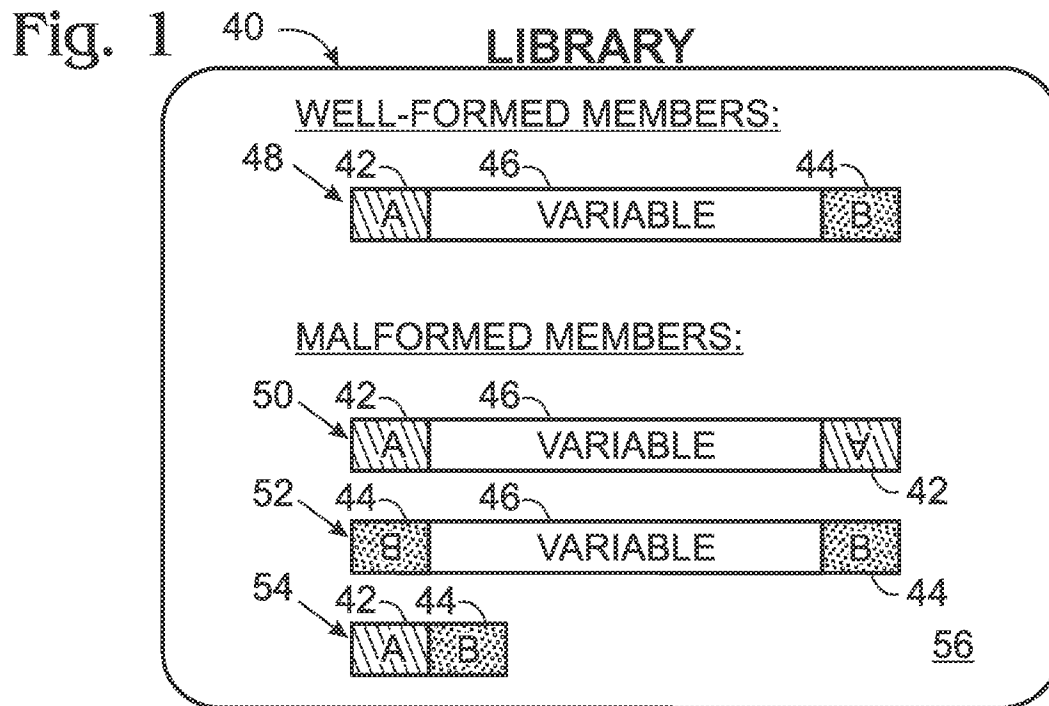
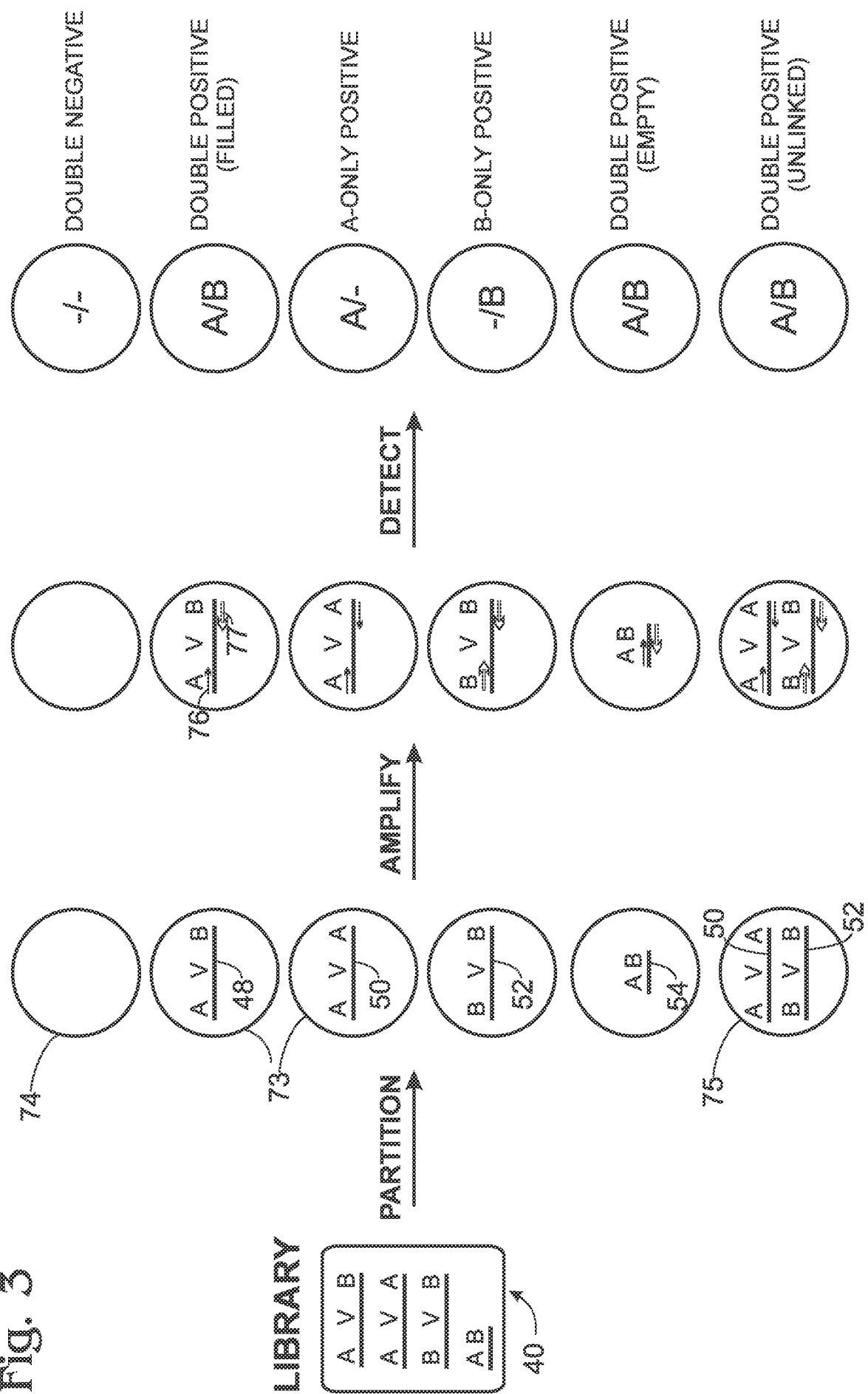


Fig. 3



3/8

Fig. 4

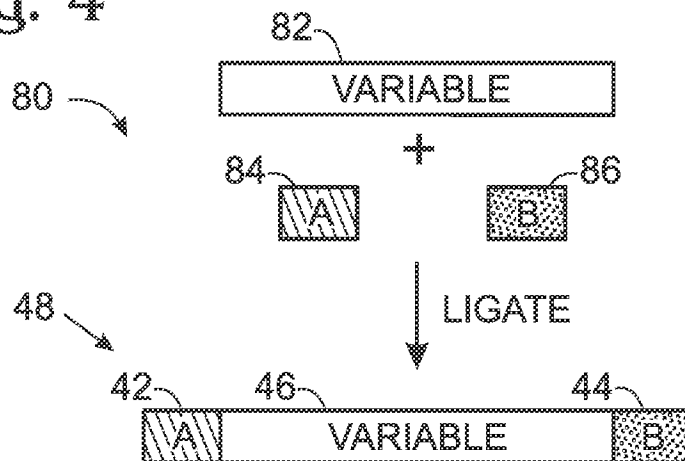
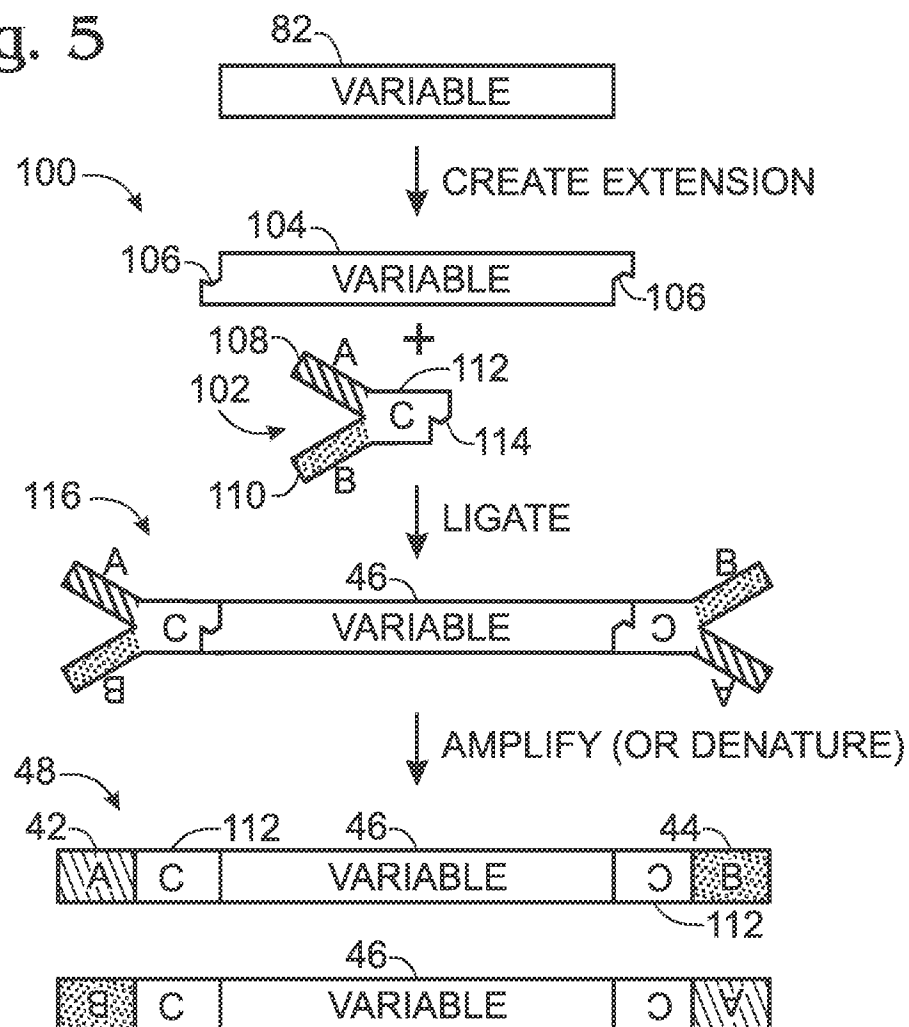
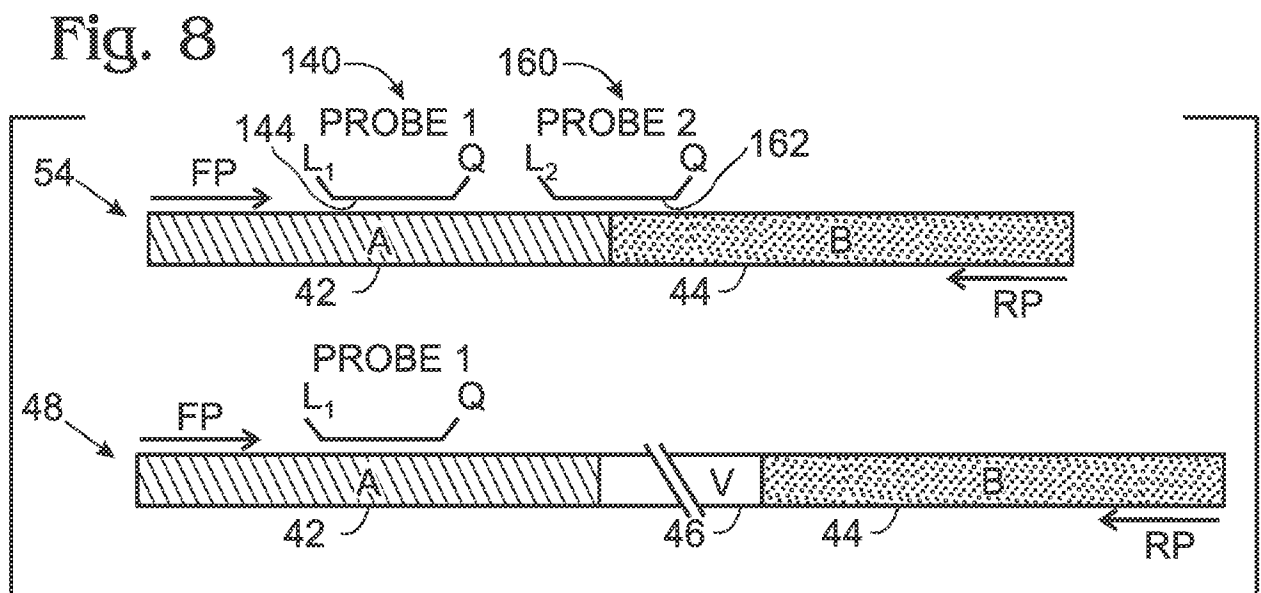
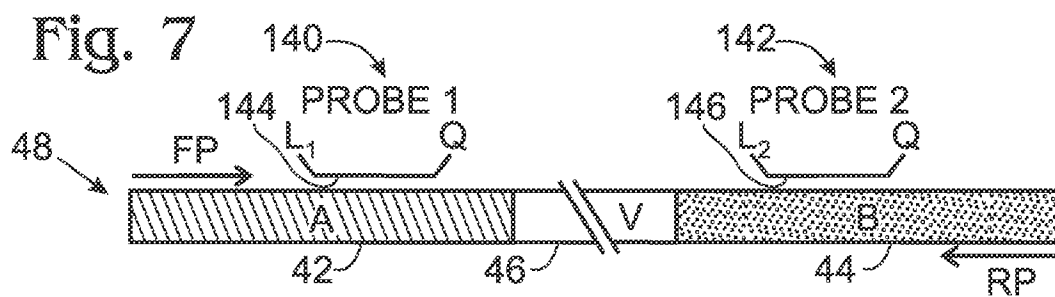
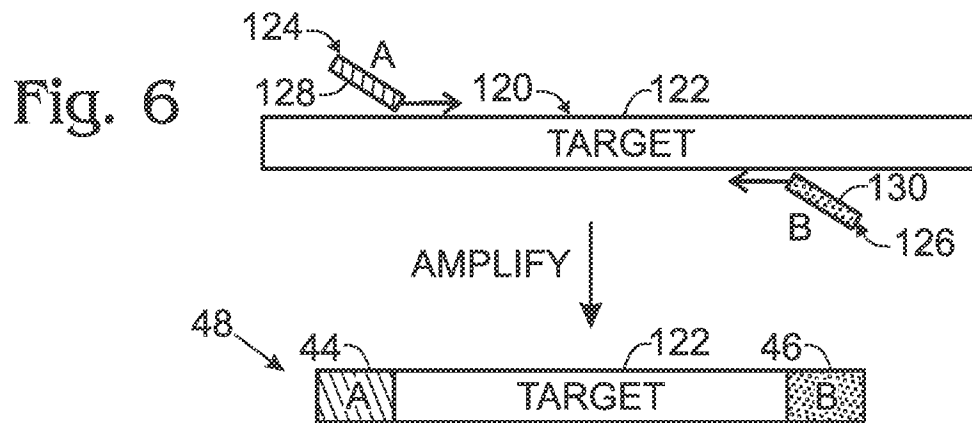


Fig. 5



4/8



5/8

Fig. 9

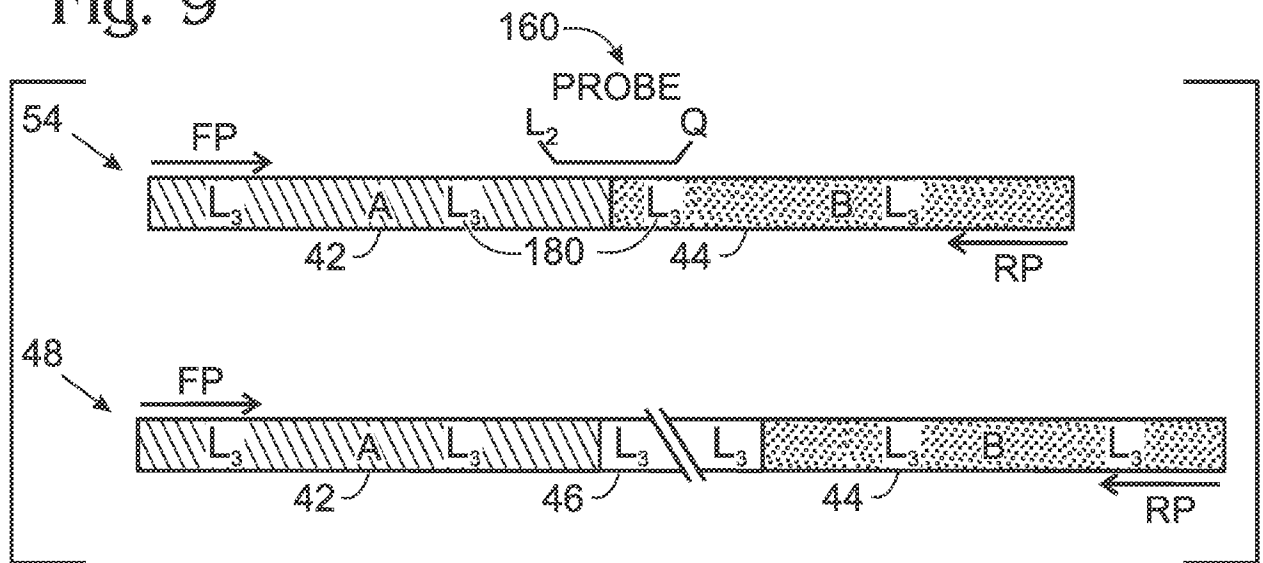
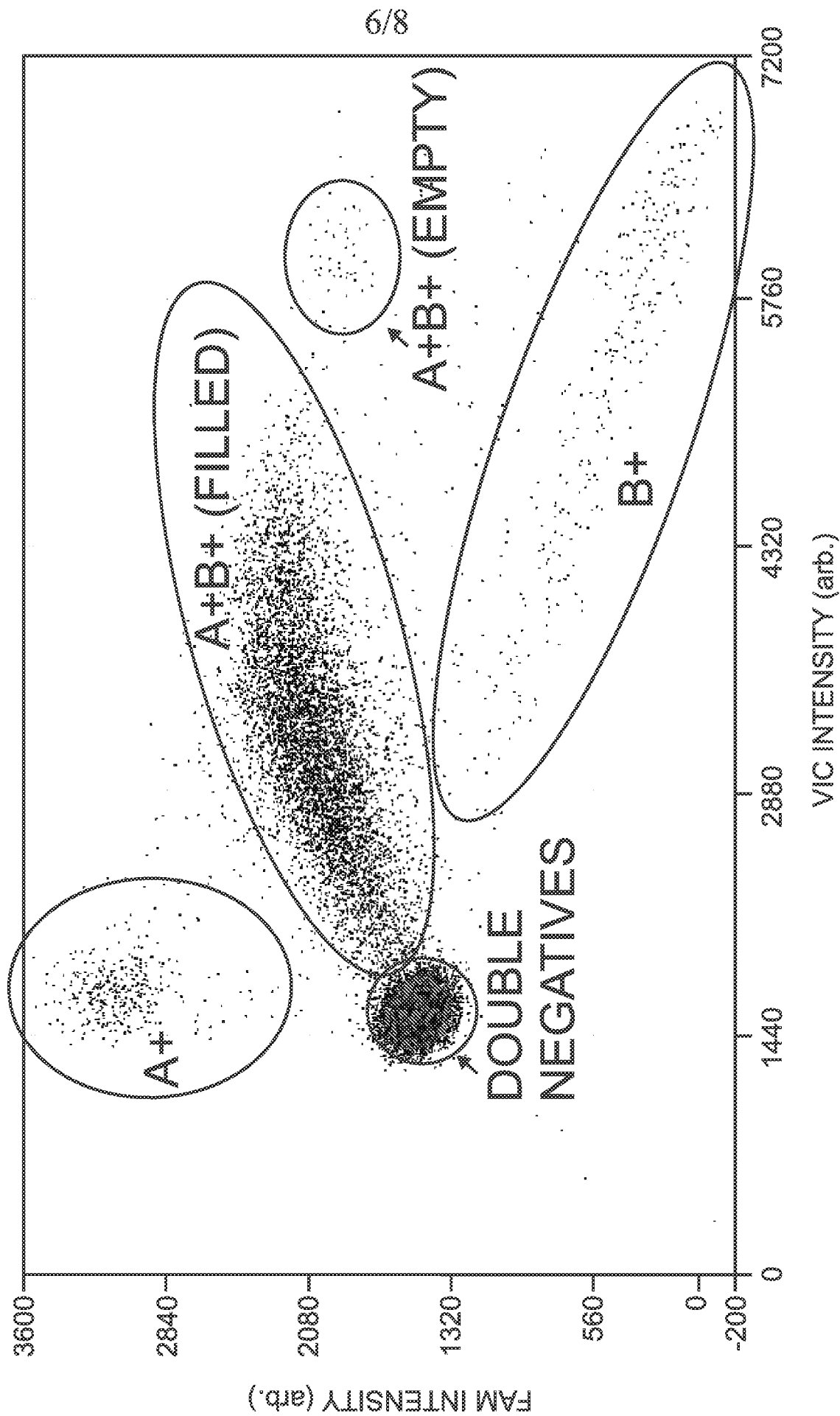


Fig. 10



7/8

Fig. 11

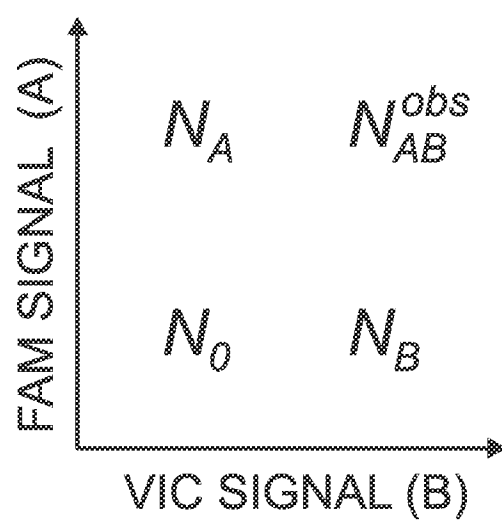
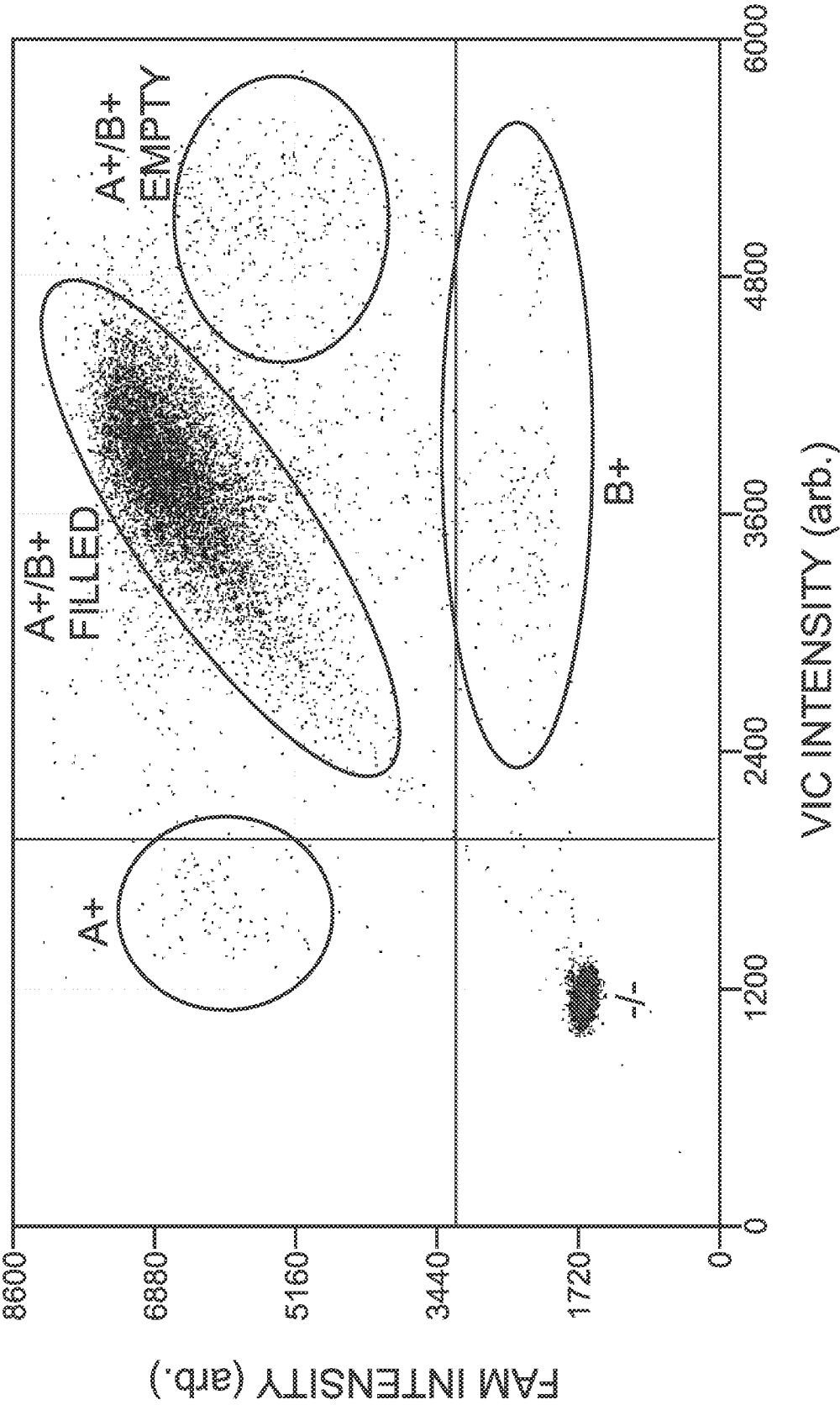


Fig. 12



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/48892

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12Q 1/68; C40B 20/04 (2012.01)

USPC - 435/6.1, 506/4

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

USPC - 435/6.1, 506/4

IPC(8) - C12Q 1/68; C40B 20/04 (2012.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

USPC - 435/6.12, 506/2

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Dialog Classic Files -- 2,6,35,64,144,155,315,371,344,347,348,349,654,652,340,345,351; Google Scholar; Google; Google Patents; Search terms -- nucleic acid;dna;library;library quality;characteristic;library;adapter;insert;partition;droplet;dilution;limiting dilution;distinct;discrete;probe;digital;junction; taqman;pyrosequencing;digital pcr;malform

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2010/0173394 A1 (Colston et al.) 8 July 2010 (08.07.2010); abstract; Para[0139];[1147];[0147];[0178];[0944];[0945];[0966];[0120];[0159];[0187];[1031];[1079]	1-20
Y	US 2004/0067493 A1 (Matsuzaki et al.) 8 April 2004 (08.04.2004); para[0086];[0038];[0040];[0098];[0126];[0071];[0007];[0022];[0097]; FIG. 5	1-20
A	Huang et al. "Rapid Screening of Complex DNA Samples by Single-Molecule Amplification and Sequencing" PLoS ONE 19 May 2011, 6(5):e19723; 19 May 2011; p3, col 1, para 2; col 2 para 1; p 4, col 1 para 1	8,15,16,19

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

29 September 2012 (29.09.2012)

Date of mailing of the international search report

22 OCT 2012

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774