(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 December 2014 (31.12.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/210353 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2014/044398

(22) International Filing Date:
26 June 2014 (26.06.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/840,403	27 June 2013 (27.06.2013)	US
61/844,804	10 July 2013 (10.07.2013)	US
13/966,150	13 August 2013 (13.08.2013)	US
PCT/US2013/054797	13 August 2013 (13.08.2013)	US
61/896,060	26 October 2013 (26.10.2013)	US
61/909,974	27 November 2013 (27.11.2013)	US
61/937,344	7 February 2014 (07.02.2014)	US
61/940,318	14 February 2014 (14.02.2014)	US
61/991,018	9 May 2014 (09.05.2014)	US

(71) Applicant: **10X TECHNOLOGIES, INC.** [US/US]; 7068 Koll Center Parkway, Suite 401, Pleasanton, CA 94566 (US).

(72) Inventors: **HINDSON, Benjamin**; 1789 Terra Court, Pleasanton, CA 94566 (US). **HINDSON, Christopher**; 4868 Bernal Avenue, Apt. B, Pleasanton, CA 94566 (US). **SCHNALL-LEVIN, Michael**; 4238 Rickeys Way, Unit T, Palo Alto, CA 94306 (US). **NESS, Kevin**; 4306 Campinia Place, Pleasanton, CA 94566 (US). **JAROSZ, Mirna**; 272 Palo Alto Avenue, Mountain View, CA 94041 (US). **MASQUELIER, Donald**; 440 Fiesta Court, Tracy, CA 95376 (US). **SAXONOV, Serge**; 5799 Balmoral Drive, Oakland, CA 94619 (US). **MERRILL, Landon**; 7068 Koll Center Parkway, Suite 401, Pleasanton, CA 94566 (US). **PRICE, Andrew**; 7068 Koll Center Parkway, Suite 401, Pleasanton, CA 94566 (US). **HARDENBOL, Paul**; 2709 21st Street, San Francisco, CA 94110 (US). **LI, Yuan**; 3465 Dublin Boulevard, Unit 429, Dublin, CA 94568 (US).

(74) Agents: **COLE, Adam, J.** 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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2014/210353 A2

(54) Title: COMPOSITIONS AND METHODS FOR SAMPLE PROCESSING

(57) Abstract: This disclosure provides methods and compositions for sample processing, particularly for sequencing applications. Included within this disclosure are bead compositions, such as diverse libraries of beads attached to large numbers of oligonucleotides containing barcodes. Often, the beads provides herein are degradable. For example, they may contain disulfide bonds that are susceptible to reducing agents. The methods provided herein include methods of making libraries of barcoded beads as well as methods of combining the beads with a sample, such as by using a microfluidic device.

COMPOSITIONS AND METHODS FOR SAMPLE PROCESSING

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 61/840,403 filed on June 27, 2013; U.S. Provisional Patent Application No. 61/844,804 filed on July 10, 2013; U.S. Patent Application No. 13/966,150 filed on August 13, 2013; PCT International Patent Application No. PCT/US13/54797 filed on August 13, 2013; U.S. Provisional Patent Application No. 61/896,060 filed on October 26, 2013; U.S. Provisional Patent Application No. 61/909,974 filed on November 27, 2013; U.S. Provisional Patent Application No. 61/937,344 filed on February 7, 2014; U.S. Provisional Patent Application No. 61/940,318 filed on February 14, 2014; and U.S. Provisional Patent Application No. 61/991,018, filed on May 9, 2014, which applications are incorporated herein by reference in their entireties for all purposes. Moreover, this application is related to U.S. Provisional Patent Application No. 61/683,192 filed on August 14, 2012, which application is incorporated herein by reference in its entirety for all purposes.

BACKGROUND

[0002] Genomic sequencing can be used to obtain information in a wide variety of biomedical contexts, including diagnostics, prognostics, biotechnology, and forensic biology. Sequencing may involve basic methods including Maxam-Gilbert sequencing and chain-termination methods, or de novo sequencing methods including shotgun sequencing and bridge PCR, or next-generation methods including polony sequencing, 454 pyrosequencing, Illumina sequencing, SOLiD sequencing, Ion Torrent semiconductor sequencing, HeliScope single molecule sequencing, SMRT® sequencing, and others. For most sequencing applications, a sample such as a nucleic acid sample is processed prior to introduction to a sequencing machine. A sample may be processed, for example, by amplification or by attaching a unique identifier. Often unique identifiers are used to identify the origin of a particular sample.

SUMMARY

[0003] The present disclosure generally provides methods, compositions, devices, and kits for the generation of beads with covalently attached polynucleotides. Such beads may be used for any suitable application.

[0004] An aspect of the disclosure provides a method of barcoding sample materials. A first partition comprising a plurality of nucleic acid barcode molecules associated therewith may be provided and the nucleic acid barcode molecules can comprise the same nucleic acid barcode sequence. The first partition may be co-partitioned with components of a sample material into a second partition and the barcode molecules can then be released from the first partition into the second partition. The released barcode molecules can be attached to one or more of the components of the sample material or fragments thereof within the second partition. In some cases, the first partition may comprise at least 1,000 barcode molecules, at least 10,000 barcode molecules, at least 100,000 barcode molecules, or at least 1,000,000 barcode molecules associated therewith having the same barcode sequence. Moreover, in some examples,

the first partition may be a bead, a microcapsule, or a droplet. In some cases, the first partition may comprise a bead (e.g., a gel bead) and the barcode molecules may be releasably coupled to the bead. Moreover, the second partition may comprise a droplet and/or may comprise no more than one first partition.

[0005] In some cases, the co-partitioning of the first partition and the components of the sample material into the second partition may comprise combining a first aqueous fluid comprising beads with a second aqueous fluid comprising the sample components in a droplet within an immiscible fluid. Moreover, the barcode molecules may be released from the first partition by degrading the first partition. In cases where the first partition is a bead, the barcode molecules may be released in the second partition by degrading the bead and/or cleaving a chemical linkage between the barcode molecules and the bead. In some cases, at least one of crosslinking of the bead and a linkage between the bead and the barcode molecules may comprise a disulfide linkage. In such cases, the barcode molecules may be released from the bead by exposing the bead to a reducing agent (e.g., dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP)).

[0006] The sample materials may comprise one or more template nucleic acid molecules and the barcode molecules may be attached to one or more fragments of the template nucleic acid molecules. In some cases, the barcode molecules may comprise a primer sequence complementary to at least a portion of the template nucleic acid molecules and the barcode molecules may be attached to the template nucleic acid molecule or fragments thereof by extending the barcode molecules to replicate at least a portion of the template nucleic acid molecules. Moreover, the sample materials may comprise the contents of a single cell, such as, for example, a cancer cell or a bacterial cell (e.g., a bacterial cell isolated from a human microbiome sample).

[0007] Furthermore, a plurality of first partitions comprising a plurality of different nucleic acid barcode sequences may be provided. Each of the first partitions can include a plurality of at least 1000 nucleic acid barcode molecules having the same nucleic acid barcode sequence associated therewith. The first partitions may be co-partitioned with components of the sample material into a plurality of second partitions. The nucleic acid barcode molecules from the first partitions may then be released into the second partitions. The released nucleic acid barcode molecules can then be attached to the components of the sample material or fragments thereof within the second partitions. In some cases, the plurality of different nucleic acid barcode sequences may comprise at least about 1,000 different barcode sequences, at least about 10,000 different barcode sequences, at least about 100,000 different barcode sequences, or at least about 500,000 different barcode sequences. Additionally, in some examples, a subset of the second partitions may comprise the same nucleic acid barcode sequence. For example, at least about 1%, at least about 2%, or at least about 5% of the second partitions may comprise the same nucleic acid barcode sequence. In addition, in some cases, at least 50% of the second partitions, at least 70% of the second partitions, or at least 90% of the second partitions may contain no more than one first partition. In some cases, at least 50% of the second partitions, at least 70% of the second partitions, or at least 90% of the second partitions may contain exactly one first partition.

[0008] Fragments of the components of the sample material may include one or more fragments of one or more template nucleic acid sequences. The fragments of the template nucleic acid sequences may be sequenced and characterized based at least in part upon a nucleic acid barcode sequence attached thereto. In some cases, the fragments of the template nucleic acid sequences may be characterized by mapping a fragment of an individual template nucleic acid sequence of the template nucleic acid sequences to an individual template nucleic acid sequence of the template nucleic acid sequences or a genome from which the individual template nucleic acid sequence was derived. In some cases, the fragments of the template nucleic acid sequence may be characterized by at least identifying an individual nucleic acid barcode sequence of the different nucleic acid barcode sequences and identifying a sequence of an individual fragment of the fragments of the template nucleic acid sequences attached to the individual nucleic acid barcode sequence.

[0009] An additional aspect of the disclosure provides a method of barcoding sample materials. A plurality of first partitions may be provided that comprise a plurality of different nucleic acid barcode sequences. Each of the first partitions may comprise a plurality of nucleic acid barcode molecules having the same nucleic acid barcode sequence associated therewith. The first partitions may by co-partitioned with components of a sample material into a plurality of second partitions. The barcode molecules can be released from the first partitions into the second partitions. The released barcode molecules can then be attached to the components of the sample material within the second partitions.

[0010] A further aspect of the disclosure provides a method of barcoding sample materials. An activatable nucleic acid barcode sequence may be provided and partitioned with one or more components of a sample material into a first partition. The activatable nucleic acid barcode sequence may be activated to produce an active nucleic acid barcode sequence in the first partition. The active nucleic acid barcode sequence can be attached to the one or more components of the sample material. In some cases, the activatable nucleic acid barcode sequence may be activated by releasing the activatable nucleic acid barcode sequence from a second partition within the first partition. In some cases, the activatable nucleic acid barcode sequence may be activated by removing a removable protecting group from the activatable nucleic acid barcode sequence.

[0011] An additional aspect of the disclosure provides a composition comprising a first partition that comprises one or more sample components and a second partition that is contained within the first partition. The second partition can have a plurality of oligonucleotides releasably associated therewith and the oligonucleotides may comprise a common barcode sequence. In some cases, the first partition may comprise an aqueous droplet in an emulsion and/or the second partition may comprise a microcapsule or bead. In some cases, the second partition may comprise a degradable bead that can be a photodegradable bead, a chemically degradable bead, and/or a thermally degradable bead. The degradable bead may comprise a chemically cleavable cross-linking such as, for example, disulfide cross-linking. Moreover, in some cases, the oligonucleotides may be releasably associated with the second partition by a cleavable linkage. The cleavable linkage may comprise, for example, a chemically cleavable linkage, a photocleavable linkage, and/or a thermally cleavable linkage. In some cases, the

cleavable linkage is a disulfide linkage. Furthermore, the sample components may comprise, for example, nucleic acids (e.g., genomic nucleic acid such as genomic DNA) or fragments thereof. The nucleic acids can comprise nucleic acid fragments that can have a length of between about 1 kb and about 100 kb, a length of between about 5 kb and about 50 kb, or a length of between about 10 kb and about 30 kb.

[0012] In some cases, the composition comprises a plurality of first partitions and a plurality of different second partitions. Each of the different second partitions can be disposed within a separate first partition and may comprise a plurality of oligonucleotides releasably associated therewith. The oligonucleotides associated with each second partition can comprise a common barcode sequence and the oligonucleotides associated with different second partitions can comprise different barcode sequences. In some cases, the different second partitions may comprise at least 1,000 different second partitions, at least 10,000 different second partitions, at least 100,000 different second partitions, or at least 500,000 different second partitions.

[0013] An additional aspect of the disclosure provides a method that comprises combining a sample of nucleic acids with a library of barcoded beads to form a mixture. The mixture can be partitioned into a plurality of partitions such that at least a subset of the partitions comprises at most one barcoded bead. Within the partitions, barcodes can be released from the barcoded beads. In some cases, the barcodes may be pre-synthesized with known sequences and/or may comprise a plurality of random N-mers. The random N-mers may be hybridized to the sample of nucleic acids in order to perform, for example, a nucleic acid amplification reaction within the partitions. In some cases, the barcoded beads may be capable of being dissolved by a reducing agent and may comprise disulfide bonds. Moreover, in some cases, the sample nucleic acids may be genomic DNA that may or may not be fragmented prior to being combined with the barcoded beads. In some cases, barcodes may be released from the barcoded beads by the action of a reducing agent. In some cases, the barcoded beads may comprise a matrix that is crosslinked with disulfide bonds and barcodes may be released from the barcoded beads by the action of a reducing agent that dissolves the barcoded beads. In some cases, barcodes may be released from the barcoded beads by heating the partitions.

[0014] In some cases, the sample of nucleic acids may be combined with the library of barcoded beads and/or the mixture of the two may be partitioned into a plurality of partitions using a microfluidic device. In some examples, the partitions may be aqueous droplets within a water-in-oil emulsion. Partitioning of the mixture into aqueous droplets within a water-in-oil emulsion may be completed using a microfluidic device.

[0015] A microfluidic device may be a droplet generator and, in some cases, may comprise a first input channel and a second input channel that meet at a junction that is fluidly connected to an output channel. The sample of nucleic acids can be introduced into the first input channel and the library of barcoded beads can be introduced to the second input channel to generate the mixture of the sample nucleic acids and the library of barcoded beads in the output channel. In some cases, a reducing agent may also be introduced to either or both of the first input channel and second input channel. Moreover, the first input channel and the second input channel may form a substantially perpendicular angle between one another.

[0016] In some cases, the output channel may be fluidly connected to a third input channel at a junction. Oil can be introduced into the third input channel such that aqueous droplets within a water-in-oil emulsion and that comprise barcoded beads are formed. The droplets may comprise on average, for example, at most ten barcoded beads, at most seven barcoded beads, at most five barcoded beads, at most three barcoded beads, at most two barcoded beads, or at most one barcoded bead. Moreover, the microfluidic device may comprise a fourth input channel that intersects the third input channel and the output channel at a junction. In some cases, oil may also be provided to the fourth input channel. In some cases, the microfluidic device may include an additional input channel that intersects the first input channel, the second input channel, or the junction of the first input channel and the second input channel. In some cases, a reducing agent may be introduced into the additional input channel.

[0017] An additional aspect of the disclosure provides a composition comprising a bead that is covalently linked to a plurality of oligonucleotides that comprise an identical barcode sequence and a variable domain. In some cases, the oligonucleotides may also comprise a primer binding site and/or a universal primer. Additionally, the identical barcode sequence may be between about 6 nucleotides and about 20 nucleotides in length. Moreover, the oligonucleotides may be covalently linked to the bead by disulfide linkages and/or the bead may comprise a cystamine or a modified cystamine. In some cases, the bead may be capable of being substantially dissolved by a reducing agent. Furthermore, in some cases, the bead may comprise at least about 1,000,000 oligonucleotides comprising an identical barcode sequence. In some cases, at least about 30% of the oligonucleotides may comprise variable domains with different sequences. In some cases, the variable domain may be a random N-mer. In some cases, the bead may be covalently linked to the oligonucleotides through a cleavable linkage such as, for example, a chemically cleavable linkage, a photocleavable linkage, and a thermally cleavable linkage.

[0018] A further aspect of the disclosure provides a composition comprising a bead that may comprise a plurality of more than 1,000,000 oligonucleotides, where each of the oligonucleotides comprises a constant region and a variable region. The bead can be capable of being substantially dissolved with a reducing agent. In some cases, each of the oligonucleotides may comprise an identical constant region. In some cases, at least 25% of the oligonucleotides may have an identical constant region. In some cases, the constant region may be a barcode sequence. In some cases, at least 25% of the oligonucleotides may have a variable region comprising a different sequence. A further aspect of the disclosure provides a library comprising at least about 1,000,000 beads that each comprise a plurality of more than 1,000,000 oligonucleotides that comprise a constant region and a variable region. In some cases, at least about 25% of the beads comprise oligonucleotides with different nucleotide sequences.

[0019] An additional aspect of the disclosure provides a composition comprising a plurality of beads where each of the beads comprises a plurality of oligonucleotides releasably coupled thereto. The oligonucleotides associated with an individual bead may comprise a common barcode domain and a variable domain. The common barcode domain can be different between two or more of the beads. In some cases, the beads may comprise at least about 10,000 different barcode domains coupled to different

beads. In some cases, each of the beads may comprise at least about 1,000,000 oligonucleotides releasably coupled thereto.

[0020] A further aspect of the disclosure provides a method of generating functionalized beads. A plurality of polymers or monomers may be mixed with one or more oligonucleotides. The polymers or monomers can be crosslinked such that disulfide bonds form between the polymers or monomers, thereby forming hardened beads. Moreover, covalent linkages can be caused to form between the oligonucleotides and the polymers or monomers. In some cases, the polymers or monomers may comprise acrylamide. In some cases, the polymers and monomers may be crosslinked to form hardened beads and covalent linkages can be caused to form between the oligonucleotides and the polymers or monomers either contemporaneously or sequentially. Moreover, in some cases, the oligonucleotides may comprise a primer (e.g., a universal primer, a sequencing primer) that may be linked to an acrydite moiety.

[0021] Additionally, one or more additional oligonucleotides may be attached to the oligonucleotides. The additional oligonucleotides may be a barcode sequence and, thus, upon attachment to the oligonucleotides, barcoded beads can be formed. In some cases, the barcode sequence may be between about 6 nucleotides and about 20 nucleotides in length.

[0022] In some cases, functionalized beads may be combined with a plurality of first additional oligonucleotides to create a mixture. The mixture may be partitioned into a plurality of partitions such that, on average, each partition comprises no more than one of the first additional oligonucleotides. In some cases, the partitions may be aqueous droplets within a water-in-oil emulsion and/or may be generated by a microfluidic device. In some cases, the partitions are generated by a bulk emulsification process. Moreover, the first additional oligonucleotides can be amplified within the partitions to produce beads comprising amplified first oligonucleotides. In some cases, a capture primer may be used during amplification and the capture primer may be attached to a capture moiety such as, for example, biotin, streptavidin or glutathione-S-transferase (GST). Following amplification, the contents of the partitions can be pooled into a common vessel. The beads comprising amplified first oligonucleotides can be separated from the contents of the partitions. In some cases, a probe may be hybridized to the amplified first oligonucleotides. The probe may comprise a capture moiety.

[0023] Furthermore, one or more second additional oligonucleotides can be attached to the amplified first oligonucleotides. In some cases, the second additional oligonucleotides may comprise a random N-mer sequence and/or a pseudo random N-mer sequence. In some cases, the second additional oligonucleotides may comprise a primer binding site that can comprise a universal sequence portion. In some cases, the primer binding site may comprise uracil containing nucleotide. Moreover, the universal sequence portion can be compatible with a sequencing device and/or may comprise a subsection of uracil containing nucleotides.

[0024] An additional aspect of the disclosure provides a method of preparing a barcode library. A plurality of separate first bead populations can be provided and a first oligonucleotide comprising a first barcode sequence segment can be attached to the separate first bead populations, such that each separate first bead population comprises a different first barcode sequence segment attached thereto. The separate

bead populations can then be pooled to provide a first pooled bead population. The first pooled bead population can then be separated into a plurality of second bead populations. A second oligonucleotide comprising a second barcode sequence segment may be attached to the first oligonucleotide attached to the second bead populations, such that each of the separate second bead populations comprises a different second barcode sequence segment. The separate second bead populations can then be pooled to provide a second pooled bead population that comprises a barcode library.

[0025] In some cases, the first barcode sequence segments and the second barcode sequence segments may be independently selected from a first set of barcode sequence segments. Additionally, the first barcode sequence segments and the second barcode sequence segments may independently comprise at least 4 nucleotides in length, at least 6 nucleotides in length, or at least 10 nucleotides in length. In some cases, the first barcode sequence segments and the second barcode sequence segments may independently include from about 4 nucleotides in length to about 20 nucleotides in length. Moreover, in some cases, the first bead populations may comprise at least 100 different first barcode sequence segments or at least 1,000 different first barcode sequence segments. Furthermore, in some cases, at least 1,000,000 first oligonucleotide molecules may be attached to each bead in each of the separate first bead populations. In some cases, the second bead populations may comprise at least 100 different second barcode sequence segments or at least 1,000 different second barcode sequence segments. In some cases, at least 1,000,000 second oligonucleotide molecules may be attached to each bead in each of the second bead populations.

[0026] Further, in some cases, at least one of the first oligonucleotide and the second oligonucleotide may comprise a functional sequence such as, for example, a primer sequence, a primer annealing sequence, an attachment sequence, and a sequencing primer sequence. In some cases, at least one of the first oligonucleotide and the second oligonucleotide may comprise a sequence segment that comprises one or more of a uracil containing nucleotide and a non-native nucleotide.

[0027] In some cases, the first oligonucleotide may be attached to the separate first bead populations by providing a splint sequence that is in part complementary to at least a portion of the first oligonucleotide and in part complementary to at least a portion of an oligonucleotide attached to the separate first bead populations. In some cases, the first oligonucleotide may be attached to the separate first bead populations such that it is releasably attached to the separate first bead populations. For example, the first oligonucleotide may be attached to the separate first bead populations through a cleavable linkage. In some cases, the first oligonucleotide may be attached to the separate first bead populations either directly or indirectly.

[0028] Additionally, in some cases, the second oligonucleotide may be attached to the first oligonucleotide by ligation. In some cases, the second oligonucleotide may be attached to the first oligonucleotide by providing a splint sequence that is in part complementary to at least a portion of the first oligonucleotide and in part complementary to at least a portion of the second oligonucleotide. In some cases, the splint sequence may provides a first overhang sequence when hybridized to the first oligonucleotide, and the second barcode sequence segment may comprise a second overhang sequence complementary to the first overhang sequence. In some cases, the first overhang sequence and the second

overhang sequences may be from about 2 nucleotides in length to about 6 nucleotides in length. Furthermore, in some cases, the first overhang sequence may comprise a plurality of different overhang sequences, and the second oligonucleotides may comprise a plurality of different second overhang sequences complementary to the plurality of different first overhang sequences.

[0029] Moreover, the separate first bead populations may comprise degradable beads, such as, for example, chemically degradable beads, photodegradable beads, and/or thermally degradable beads. In some cases, the separate first bead populations may comprise beads that comprise chemically reducible cross-linkers such, as for example, chemically reducible cross-linkers that comprise disulfide linkages.

[0030] In some cases, a third oligonucleotide may be attached to the second oligonucleotide attached to the first oligonucleotide. The third oligonucleotide may comprise a functional sequence that may be a primer sequence (e.g., a universal primer sequence, a targeted primer sequence, or a random sequence) and/or may be a random N-mer sequence. In cases where the third oligonucleotide comprises a random N-mer sequence, the random N-mer sequence may be from about 5 nucleotides in length to about 25 nucleotides in length.

[0031] An additional aspect of the disclosure provides a method of preparing a barcode library. A first pooled bead population comprising a plurality of different first bead populations may be provided, where each different first bead population comprises a different first oligonucleotide attached thereto. Each different first oligonucleotide may comprise a different first barcode sequence segment. The first pooled bead population may be separated into a plurality of second bead populations. A second oligonucleotide comprising a second barcode sequence segment may be attached to the first oligonucleotide already attached to the second bead populations, where each second bead population comprises a different second barcode sequence segment. The second bead populations can be pooled to provide a second pooled bead population comprising a barcode library.

[0032] In some cases, the first oligonucleotide may be releasably attached to the beads in the first pooled bead population. In some cases, the first oligonucleotide may be attached to the beads in the first pooled bead population through a cleavable linkage. In some cases, the beads in the first pooled population may each comprise at least 1,000,000 first oligonucleotides attached thereto. In some cases, the first pooled bead population may comprise at least 10 different first bead populations, at least 100 different first bead populations, or at least 500 different first bead populations.

[0033] A further aspect of the disclosure provides a barcode library comprising a plurality of different oligonucleotides. Each different oligonucleotide may comprise a first barcode sequence segment selected from a first set of barcode sequence segments; a second barcode sequence segment selected from a second set of barcode sequence segments; and a linking sequence joining the first barcode sequence segment and the second barcode sequence segment. The linking sequence can be from about 2 nucleotides in length to about 6 nucleotides in length and may be selected from a set of linking sequences. In some cases, the set of linking sequences includes from about 2 different linking sequences to about 50 different linking sequences. In some cases, the first set of barcode sequence segments and the second set of barcode sequence segments are the same.

[0034] An additional aspect of the disclosure provides a method of amplifying a template nucleic acid sequence. A template nucleic acid sequence and a bead comprising a plurality of releasably attached oligonucleotides may be co-partitioned into a partition. The oligonucleotides may comprise a primer sequence complementary to one or more regions of the template nucleic acid sequence and may comprise a common sequence. The primer sequence can be annealed to the template nucleic acid sequence and the primer sequence can be extended to produce one or more first copies of at least a portion of the template nucleic acid sequence, where the one or more first copies comprising the primer sequence and the common sequence.

[0035] In some cases, the primer sequence may comprise a variable primer sequence (e.g., a random N-mer) and/or may comprise a targeted primer sequence. In some cases, the partition may comprise a droplet in an emulsion. Prior to annealing the primer sequence to the template nucleic acid sequence, the oligonucleotides may be released from the bead into the partition. In some examples, a polymerase enzyme (e.g., an exonuclease deficient polymerase enzyme) may be provided in the partition. Moreover, extension of the primer sequence may comprise extending the primer sequence using a strand displacing polymerase enzyme (e.g., a thermostable strand displacing polymerase enzyme having, for example, substantially no exonuclease activity). Furthermore, the oligonucleotides may be exonuclease resistant. For example, the oligonucleotides may comprise one or more phosphorothioate linkages. In some cases, the phosphorothioate linkages may comprise a phosphorothioate linkage at a terminal internucleotide linkage in the oligonucleotides.

[0036] Additionally, one or more variable primer sequences may be annealed to the first copies and extended to produce one or more second copies from the first copies, such that the second copies comprise the one or more variable primer sequences and the common sequence. In some cases, the second copies may comprise a sequence complementary to at least a portion of an individual first copy of the first copies and a sequence complementary to an individual variable sequence of the one or more variable primer sequences. In some cases, the second copies may preferentially form a hairpin molecule under annealing conditions. Moreover, in some cases, the oligonucleotides may comprise a sequence segment that is not copied during the extension of the variable primer sequences. The sequence segment that is not copied may comprise, for example, one or more uracil containing nucleotides. In addition, any steps of the method may be repeated to produce amplified nucleic acids.

[0037] A further aspect of the disclosure provides a method of amplifying a plurality of different nucleic acids. Different nucleic acids may be partitioned into separate first partitions, where each first partition comprises a second partition having a plurality of oligonucleotides releasably associated therewith. The plurality of oligonucleotides associated with a given second partition may comprise a variable primer sequence and a barcode sequence, with the oligonucleotides associated with different second partitions comprising different barcode sequences. The oligonucleotides associated with the plurality of second partitions can be released into the first partitions. The variable primer sequences in the first partitions can be released to nucleic acids within the first partitions and extended to produce one or more copies of at least a portion of the nucleic acids within the first partitions, such that the copies comprise the

oligonucleotides and associated barcode sequences released into the first partitions. In some cases, the first partitions may comprise droplets in an emulsion and the second partitions may comprise beads. In some cases, each bead may comprise more than 100,000 oligonucleotides associated therewith or more than 1,000,000 oligonucleotides associated therewith. In some cases, the second partitions may comprise at least 1,000 different barcode sequences, at least 10,000 different barcode sequences, or at least 100,000 different barcode sequences.

[0038] An additional aspect of the disclosure provides a method of whole genome amplification. A random primer may be hybridized to a genomic nucleic acid. The random primer may be attached to a universal nucleic acid sequence and a nucleic acid barcode sequence, where the universal nucleic acid sequence may comprise one or more uracil containing nucleotides. The random primer may be extended to form an amplified product and the amplified product may be exposed to conditions suitable to cause the amplified product to undergo an intramolecular hybridization reaction that forms a partial hairpin molecule. In some cases, the random primer may be a random N-mer sequence. In some cases, the universal nucleic acid sequence may comprise a segment of at least 10 nucleotides that do not comprise uracil. Moreover, the method may be performed in the presence of an oligonucleotide blocker. The oligonucleotide blocker may be capable of hybridizing to at least a portion of the universal nucleic acid sequence and/or may comprise a C3 spacer (/3SpC3/), a Dideoxy-C (/3ddC/), or a 3' phosphate.

[0039] An additional aspect of the disclosure provides a method of amplifying nucleic acids. A genomic component may be fragmented into a plurality of first fragments. The first fragments may be co-partitioned with a plurality of oligonucleotides into a plurality of partitions. The oligonucleotides in each of the partitions may comprise a primer sequence and a common sequence. The primer sequences in each partition may be annealed to a plurality of different regions of the first fragments within each partition and the primer sequences extended along the first fragments to produce amplified first fragments within each partition. In some cases, the amplified first fragments within the partitions may comprise at least 1X coverage of the genomic component, at least 2X coverage of the genomic component, or at least 10X coverage of the genomic component. In some cases, the genomic component may comprise a chromosome. In some cases, the genomic component may comprise a whole genome of an organism.

[0040] A further aspect of the disclosure provides a method of characterizing a nucleic acid segment. A nucleic acid segment may be co-partitioned with a bead comprising a comprising a plurality of oligonucleotides that comprise a common nucleic acid barcode sequence into a partition. The oligonucleotides may be attached to fragments of the nucleic acid segment or to copies of portions of the nucleic acid segment, such that the common nucleic acid barcode sequence is attached to the fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment. The fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment and attached common nucleic acid barcode sequence can be sequenced and the fragments of the nucleic acid segment or the copies of the nucleic acid segment can be characterized as being linked within the nucleic acid segment based at least in part, upon a their attachment to the common nucleic acid barcode sequence. The nucleic acid segment and the bead, for example, may be co-partitioned into a droplet in an emulsion or may be

co-partitioned into a microcapsule. In some cases, the fragments of the nucleic acid segment may comprise overlapping fragments of the nucleic acid segment. In some cases, the fragments of the nucleic acid segment may comprise greater than 2X coverage of the nucleic acid segment or greater than 10X coverage of the nucleic acid segment.

[0041] Moreover, in some cases, the oligonucleotides may be releasably attached to the bead. For example, the oligonucleotides may be releasable from the bead upon the application of a stimulus (e.g., a thermal stimulus, a photo stimulus, a chemical stimulus, etc.) to the bead. In some cases, the application of the stimulus may result in the cleavage of a linkage between the oligonucleotides and the bead and/or may result in the degradation of the bead, such that the oligonucleotides are released from the bead. Furthermore, the bead may comprise at least about 10,000 oligonucleotides attached thereto, at least about 100,000 oligonucleotides attached thereto, at least about 1,000,000 oligonucleotides attached thereto, at least about 10,000,000 oligonucleotides attached thereto, or at least about 100,000,000 oligonucleotides attached thereto. Additionally, in some cases, the oligonucleotides may comprise one or more functional sequences, such as, for example, a primer sequence, a primer annealing sequence, or an immobilization sequence. In some cases, the fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment and attached common nucleic acid barcode sequence may be sequenced via a sequencing by synthesis process.

[0042] Further, in some cases, the oligonucleotides may comprise a primer sequence capable of annealing with a portion of the nucleic acid segment or a complement thereof. The primer sequence can be extended to replicate at least a portion of the nucleic acid segment or complement thereof, to produce a copy of a portion of the nucleic acid segment or complement thereof that comprises the common nucleic acid barcode sequence. In some cases, the oligonucleotides may comprise at least a first sequencing primer sequence.

[0043] In some cases, a plurality of nucleic acid segments may be co-partitioned with a plurality of different beads into a plurality of separate partitions, such that each partition of a plurality of different partitions of the separate partitions contains a single bead. Each bead may comprise a plurality of oligonucleotides that comprise a common barcode sequence attached thereto, where the different beads comprises a plurality of different barcode sequences. Barcode sequences in each partition may be attached to fragments of the nucleic acid segments or to copies of portions of the nucleic acid segments within the separate partitions. The fragments or copies can then be pooled from the separate partitions and the fragments or copies and any associated barcode sequences may be sequenced to provide sequenced fragments or sequenced copies. The sequenced fragments or sequenced copies may be characterized as deriving from a common nucleic acid segment, based in part upon the sequenced fragments or sequenced copies comprising a common barcode sequence. In some cases, the nucleic acid segments may comprise fragments of at least a portion of a genome. In such cases, sequences may be assembled from the sequenced fragments or sequenced copies to provide a contiguous sequence of the at least a portion of the genome. Assembly of the sequences from the sequenced fragments or sequenced copies may be based in part upon each of a nucleotide sequence of the sequenced fragmentsor sequenced

copies and the sequenced fragments or sequenced copies comprising a common barcode sequence. Moreover, in some cases, the fragments of the nucleic acid segments or the copies of the portions of the nucleic acid segments may be characterized based in part upon each of a nucleotide sequence of the fragments of the nucleic acid segments or the copies of the portions of the nucleic acid segments and the sequenced fragments or sequenced copies comprising a common barcode sequence.

[0044] In some cases, the different beads may comprise at least 1,000 different barcode sequences, at least 10,000 different barcode sequences, at least 100,000 different barcode sequences, or at least 1,000,000 different barcode sequences. In some cases, two or more partitions of the separate partitions may comprise beads that comprise the same barcode sequence. In some cases, at least 1% of the separate partitions comprise beads having the same barcode sequence.

[0045] An additional aspect of the disclosure provides a method of characterizing a target nucleic acid. First fragments of a target nucleic acid may be partitioned into a plurality of droplets, where each droplet comprises a bead having a plurality of oligonucleotides attached thereto. The oligonucleotides attached to a given bead can comprise a common barcode sequence. The common barcode sequence can be attached to second fragments of the first fragments and the droplets can be pooled. The second fragments and attached barcode sequences can sequenced and the second fragments can be mapped to one or more of the first fragments based, at least in part, upon the second fragments comprising a common barcode sequence.

[0046] An additional aspect of the disclosure provides a method of sequencing nucleic acids. A plurality of target nucleic acid sequences may be provided and separated into a plurality of separate partitions. Each partition of the separate partitions may comprise one or more target nucleic acid sequences and a bead comprising a plurality of oligonucleotides attached thereto. The oligonucleotides attached to a given bead may comprise a common barcode sequence. The oligonucleotides may be attached to fragments of the one or more target nucleic acid sequences or to copies of portions of the one or more target nucleic acid sequences within a partition, thereby attaching the common barcode sequence to the fragments of the one or more target nucleic acid sequences or the copies of the portions of the one or more target nucleic acid sequences. The separate partitions can be pooled and the fragments of the one or more target nucleic acid sequences or the copies of the portions of the one or more target nucleic acid sequences and attached barcode sequences can be sequenced to provide barcoded fragment sequences or barcoded copy sequences. In some cases, the barcoded fragment sequences or barcoded copy sequences can be assembled into one or more contiguous nucleic acid sequences based, in part, upon a barcode portion of the barcoded fragment sequences or barcoded copy sequences.

[0047] An additional aspect of the disclosure provides a method of characterizing a nucleic acid segment. A nucleic acid segment may be co-partitioned with a bead comprising a plurality of oligonucleotides that comprise a common nucleic acid barcode sequence, into a first droplet. The oligonucleotides may be attached to fragments of the nucleic acid segment or to copies of portions of the nucleic acid segment, thereby attaching the common nucleic acid barcode sequence to the fragments of the nucleic acid segment or to the copies of the portions of the nucleic acid segment. The fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment and attached common nucleic acid barcode

sequence can be sequenced to provide a plurality of barcoded fragment sequences or barcoded copy sequences. The barcoded fragment sequences or barcoded copy sequences can be assembled into one or more contiguous nucleic acid sequences based at least in part on the common nucleic acid barcode sequence. In some cases, the barcoded fragment sequences or barcoded copy sequences may be assembled based in part upon a nucleic acid sequence of non-barcode potion of the barcoded fragment sequences or barcoded copy sequences.

[0048] An additional aspect of the disclosure provides a method of sequencing nucleic acids. A plurality of target nucleic acid sequences may be provided and the target nucleic acid sequences separated into a plurality of separate partitions. Each partition of the separate partitions may comprise one or more target nucleic acid sequences and a plurality of oligonucleotides. The oligonucleotides in a given partition may comprise a common barcode sequence and the plurality of separate partitions may comprise at least 10,000 different barcode sequences. The common barcode sequence in each partition may be attached to fragments of the one or more target nucleic acid sequences or to copies of portions of the one or more target nucleic acid sequences within the partition. The separate partitions can be pooled and the fragments of the one or more target nucleic acid sequences or the copies of the portions of the one or more target nucleic acid sequences and attached barcode sequences can be sequenced. In some cases, the separate partitions may comprise at least 100,000 different barcode sequences.

INCORPORATION BY REFERENCE

[0049] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference in their entireties for all purposes and 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

[0050] **Fig 1A** is a flow diagram for making barcoded beads.

[0051] **Fig 1B** is a flow diagram for processing a sample for sequencing.

[0052] **Fig 2** is a flow diagram for making beads.

[0053] **Fig 3A** is a flow diagram for adding barcodes to beads by limiting dilution.

[0054] **Fig 3B** is a flow diagram for adding additional sequences to oligonucleotides attached to beads.

[0055] **Figs 4A-N** are diagrams for attaching sequences to beads. "g/w" means gel-in-water; "g/w/o" means gel-in-water-in-oil;

[0056] **Fig 5** provides an illustration of a gel bead attached to an oligonucleotide **5A**, an image of a microfluidic chip used to make Gel Beads in Emulsions (GEM) **5B**, as well as images of GEMs **5C, D, E**.

[0057] **Fig 6** provides bright-field (**A, C, E**) and fluorescent (**B, D, F**) images of beads with attached oligonucleotides.

[0058] **Figs 7A-C** provide fluorescent images of beads attached to DNA.

[0059] **Figs 8A-F** provide images of barcode-enriched populations of beads.

[0060] **Figs 9A-D** provide images of the dissolution of beads by heating.

[0061] **Fig 10A** provides a schematic of a functionalized bead. **Figs 10B-G** provide images of beads dissolved with a reducing agent.

[0062] **Fig 11A** provides a schematic of a functionalized bead. **Figs 11B-D** provide graphic depictions of the presence of barcode oligonucleotides and primer-dimer pairs when beads are prepared using different conditions.

[0063] **Fig 12** is a graphic depiction of content attached to beads.

[0064] **Fig 13A** is a flow diagram illustrating the addition of barcodes to beads using partitions.

[0065] **Fig 13B** is a flow diagram illustrating the addition of additional sequences to beads.

[0066] **Fig 13C** is a diagram illustrating the use of a combinatorial approach in microwell plates to make barcoded beads.

[0067] **Figs 14A-C** are diagrams of oligonucleotides containing universal sequences (R1, P5) and uracil containing nucleotides.

[0068] **Figs 15A-G** are diagrams of steps used in the partial hairpin amplification for sequencing (PHASE) process.

[0069] **Fig 16A** is a graphic depiction of including uracil containing nucleotides in the universal portion of the primer.

[0070] **Fig 16B** is a graphic depiction of controlling amplification product length by including acyNTPs in the reaction mixture.

[0071] **Fig 17** is a graphic depiction of reducing start site bias by adding a blocker oligonucleotide.

[0072] **Fig 18** is a flow diagram of a digital processor and its related components.

[0073] **Fig 19** is a table providing example sequences for Illumina sequencers.

[0074] **Fig 20** is a table providing a list of example capture moiety concentrations used to label beads.

[0075] **Fig 21** is a table providing a list of sequencing metrics obtained using primers comprising thymine containing nucleotides.

[0076] **Fig 22** is a table providing a list of sequencing metrics obtained using primers comprising uracil containing nucleotides.

[0077] **Figs 23A-D** are schematics illustrating the use of an example ligation-based combinatorial approach to make barcoded beads.

[0078] **Figs 24A-B** are schematics illustrating an example use of spacer bases in a ligation-based combinatorial approach to make barcoded beads.

[0079] **Figs 25A-C** are schematics illustrating the use of an example ligation-based combinatorial approach to make barcoded beads.

[0080] **Fig 26** is a schematic illustrating example nucleic acids used in an example ligation-based combinatorial approach to make barcoded beads.

[0081] **Fig 27** is a schematic illustrating an example ligation-based combinatorial approach to make barcoded beads.

[0082] **Figs 28A-B** are schematic representations of example targeted barcode constructs suitable for strand-specific amplification.

[0083] **Figs 29A-C** are structural depictions of example monomers and cross-linkers that can be polymerized to generate beads.

[0084] **Figs 30A-C** are structural depictions of an example method that can be used to generate beads.

[0085] **Fig 31** is a schematic depiction of example beads comprising functional groups that can be used to attach species to the beads.

[0086] **Fig 32** provides structural depictions of example initiators that may be used during a polymerization reaction.

[0087] **Fig 33A** is a schematic depiction of barcode primers. **Figs 33B-E** are graphic depictions of data corresponding to example amplification reaction experiments described in Example 16.

[0088] **Figs 34A-C** are schematics of example hairpin constructs.

[0089] **Figs 35A-B** are schematics of example methods for functionalizing beads.

[0090] **Fig 36** is a photograph of a gel obtained during a gel electrophoresis experiment described in Example 17.

[0091] **Fig 37A** is a schematic depiction of oligonucleotides described in Example 18. **Fig 37B** is a photograph of a gel obtained during a gel electrophoresis experiment described in Example 18. **Fig 37C** is a micrograph of beads obtained during a fluorescence microscopy experiment described in Example 18.

[0092] **Fig 38** provides a schematic illustration of an exemplary nucleic acid barcoding and amplification process.

[0093] **Fig 39** provides a schematic illustration of an exemplary application of the methods described herein to nucleic acid sequencing and assembly.

[0094] **Fig 40** presents examples of alternative processing steps following barcoding and amplification of nucleic acids, as described herein.

DETAILED DESCRIPTION

I. General Overview

[0095] This disclosure provides methods, systems and compositions useful in the processing of sample materials through the controlled delivery of reagents to subsets of sample components, followed by analysis of those sample components employing, in part, the delivered reagents. In many cases, the methods and compositions are employed for sample processing, particularly for nucleic acid analysis applications, generally, and nucleic acid sequencing applications, in particular. Included within this disclosure are bead compositions that include diverse sets of reagents, such as diverse libraries of beads attached to large numbers of oligonucleotides containing barcode sequences, and methods of making and using the same.

[0096] Methods of making beads can generally include, *e.g.* combining bead precursors (such as monomers or polymers), primers, and cross-linkers in an aqueous solution, combining said aqueous solution with an oil phase, sometimes using a microfluidic device or droplet generator, and causing water-in-oil droplets to form. In some cases, a catalyst, such as an accelerator and/or an initiator, may be added before or after droplet formation. In some cases, initiation may be achieved by the addition of energy,

such, as for example via the addition of heat or light (e.g., UV light). A polymerization reaction in the droplet can occur to generate a bead, in some cases covalently linked to one or more copies of an oligonucleotide (*e.g.*, primer). Additional sequences can be attached to the functionalized beads using a variety of methods. In some cases, the functionalized beads are combined with a template oligonucleotide (*e.g.*, containing a barcode) and partitioned such that on average one or fewer template oligonucleotides occupy the same partition as a functionalized bead. While the partitions may be any of a variety of different types of partitions, e.g., wells, microwells, tubes, vials, microcapsules, etc., in preferred aspects, the partitions may be droplets (*e.g.*, aqueous droplets) within an emulsion. The oligonucleotide (*e.g.*, barcode) sequences can be attached to the beads within the partition by a reaction such as a primer extension reaction, ligation reaction, or other methods. For example, in some cases, beads functionalized with primers are combined with template barcode oligonucleotides that comprise a binding site for the primer, enabling the primer to be extended on the bead. After multiple rounds of amplification, copies of the single barcode sequence are attached to the multiple primers attached to the bead. After attachment of the barcode sequences to the beads, the emulsion can be broken and the barcoded beads (or beads linked to another type of amplified product) can be separated from beads without amplified barcodes. Additional sequences, such as a random sequence (*e.g.*, a random N-mer) or a targeted sequence, can then be added to the bead-bound barcode sequences, using, for example, primer extension methods or other amplification reactions. This process can generate a large and diverse library of barcoded beads.

[0097] **Fig 1A** illustrates an example method for generating a barcoded bead. First, gel precursors (*e.g.*, linear polymers and/or monomers), cross-linkers, and primers may be combined in an aqueous solution, **101.** Next, in a microfluidic device, the aqueous solution can then be combined with an oil phase, **102.** Combining the oil phase and aqueous solution can cause water-in-oil droplets to form, **103.** Within water-in-oil droplets, polymerization of the gel precursors occurs to form beads comprising multiple copies of a primer, **104.** Following generation of a primer-containing bead, the emulsion may be broken, **105** and the beads recovered. The recovered beads may be separated from unreacted components, via, for example, washing and introduced to any suitable solvent (e.g., an aqueous solvent, a non-aqueous solvent). In some cases, the primer-containing beads may then be combined (*e.g.*, via limiting dilution methods) with template barcode sequences in droplets of another emulsion, such that each droplet comprises on average at least one bead and on average one or less molecules of a template barcode sequence. The template barcode sequence may be clonally amplified, using the primer attached to the bead, resulting in attachment to the bead of multiple copies of a barcode sequence complementary to the template, **106.** The barcoded beads may then be pooled into a population of beads either containing barcodes or not containing barcodes, **107**. The barcoded beads may then be isolated by, for example, an enrichment step. The barcode molecules may also be provided with additional functional sequence components for exploitation in subsequent processing. For example, primer sequences may be incorporated into the same oligonucleotides that include the barcode sequence segments, to enable the use of the barcode containing oligonucleotides to function as extension primers for duplicating sample nucleic acids, or as priming sites

for subsequent sequencing or amplification reactions. In one example, random N-mer sequences may then be added to the barcoded beads, **108**, via primer extension or other amplification reaction and a diverse library of barcoded beads, **110**, may thereby be obtained, where such random n-mer sequences can provide a universal primer sequence. Likewise, functional sequences may include immobilization sequences for immobilizing barcode containing sequences onto surfaces, e.g., for sequencing applications. For ease of discussion, a number of specific functional sequences are described below, such as P5, P7, R1, R2, sample indexes, random Nmers, etc., and partial sequences for these, as well as complements of any of the foregoing. However, it will be appreciated that these descriptions are for purposes of discussion, and any of the various functional sequences included within the barcode containing oligonucleotides may be substituted for these specific sequences, including without limitation, different attachment sequences, different sequencing primer regions, different n-mer regions (targeted and random), as well as sequences having different functions, e.g., secondary structure forming, e.g., hairpins or other structures, probe sequences, e.g., to allow interrogation of the presence or absence of the oligonucleotides or to allow pull down of resulting amplicons, or any of a variety of other functional sequences.

[0098] Also included within this disclosure are methods of sample preparation for nucleic acid analysis, and particularly for sequencing applications. Sample preparation can generally include, *e.g.* obtaining a sample comprising sample nucleic acid from a source, optionally further processing the sample, combining the sample nucleic acid with barcoded beads, and forming emulsions containing fluidic droplets comprising the sample nucleic acid and the barcoded beads. Droplets may be generated, for example, with the aid of a microfluidic device and/or via any suitable emulsification method. The fluidic droplets can also comprise agents capable of dissolving, degrading, or otherwise disrupting the barcoded beads, and/or disrupting the linkage to attached sequences, thereby releasing the attached barcode sequences from the bead. The barcode sequences may be released either by degrading the bead, detaching the oligonucleotides from the bead such as by a cleavage reaction, or a combination of both. By amplifying (*e.g.*, via amplification methods described herein) the sample nucleic acid in the fluidic droplets, for example, the free barcode sequences can be attached to the sample nucleic acid. The emulsion comprising the fluidic droplets can then be broken and, if desired, additional sequences (*e.g.*, sequences that aid in particular sequencing methods, additional barcode sequences, etc.) can then be added to the barcoded sample nucleic acid using, for example, additional amplification methods. Sequencing can then be performed on the barcoded, amplified sample nucleic acid and one or more sequencing algorithms applied to interpret the sequencing data. As used herein, the sample nucleic acids may include any of a wide variety of nucleic acids, including, e.g., DNA and RNA, and specifically including for example, genomic DNA, cDNA, mRNA total RNA, and cDNA created from a mRNA or total RNA transcript.

[0099] **Fig 1B** illustrates an example method for barcoding and subsequently sequencing a sample nucleic acid. First, a sample comprising nucleic acid may be obtained from a source, **111,** and a set of barcoded beads may be obtained, e.g., as described herein, **112.** The beads are preferably linked to oligonucleotides containing one or more barcode sequences, as well as a primer, such as a random N-mer

or other primer. Preferably, the barcode sequences are releasable from the barcoded beads, e.g., through cleavage of a linkage between the barcode and the bead or through degradation of the underlying bead to release the barcode, or a combination of the two. For example, in certain preferred aspects, the barcoded beads can be degraded or dissolved by an agent, such as a reducing agent to release the barcode sequences. In this example, the sample comprising nucleic acid, **113,** barcoded beads, **114,** and e.g., a reducing agent, **116,** are combined and subject to partitioning. By way of example, such partitioning may involve introducing the components to a droplet generation system, such as a microfluidic device, **115**. With the aid of the microfluidic device **115**, a water-in-oil emulsion **117** may be formed, wherein the emulsion contains aqueous droplets that contain sample nucleic acid, reducing agent, and barcoded beads, **117.** The reducing agent may dissolve or degrade the barcoded beads, thereby releasing the oligonucleotides with the barcodes and random N-mers from the beads within the droplets, **118**. The random N-mers may then prime different regions of the sample nucleic acid, resulting in amplified copies of the sample after amplification, wherein each copy is tagged with a barcode sequence, **119.** Preferably, each droplet contains a set of oligonucleotides that contain identical barcode sequences and different random N-mer sequences. Subsequently, the emulsion is broken, **120** and additional sequences (*e.g.*, sequences that aid in particular sequencing methods, additional barcodes, etc.) may be added, **122**, via, for example, amplification methods (*e.g.*, PCR). Sequencing may then be performed, **123,** and an algorithm applied to interpret the sequencing data, **124.** Sequencing algorithms are generally capable, for example, of performing analysis of barcodes to align sequencing reads and/or identify the sample from which a particular sequence read belongs.

[00100] The methods and compositions of this disclosure may be used with any suitable digital processor. The digital processor may be programmed, for example, to operate any component of a device and/or execute methods described herein. In some embodiments, bead formation may be executed with the aid of a digital processor in communication with a droplet generator. The digital processor may control the speed at which droplets are formed or control the total number of droplets that are generated. In some embodiments, attaching barcode sequences to sample nucleic acid may be completed with the aid of a microfluidic device and a digital processor in communication with the microfluidic device. In some cases, the digital processor may control the amount of sample and/or beads provided to the channels of the microfluidic device, the flow rates of materials within the channels, and the rate at which droplets comprising barcode sequences and sample nucleic acid are generated.

[00101] The methods and compositions of this disclosure may be useful for a variety of different molecular biology applications including, but not limited to, nucleic acid sequencing, protein sequencing, nucleic acid quantification, sequencing optimization, detecting gene expression, quantifying gene expression, epigenetic applications, and single-cell analysis of genomic or expressed markers. Moreover, the methods and compositions of this disclosure have numerous medical applications including identification, detection, diagnosis, treatment, staging of, or risk prediction of various genetic and non-genetic diseases and disorders including cancer.

II. Beads or Particles

[00102] The methods, compositions, devices, and kits of this disclosure may be used with any suitable bead or particle, including gel beads and other types of beads. Beads may serve as a carrier for reagents that are to be delivered in accordance with the methods described herein. In particular, these beads may provide a surface to which reagents are releasably attached, or a volume in which reagents are entrained or otherwise releasably partitioned. These reagents may then be delivered in accordance with a desired method, for example, in the controlled delivery of reagents into discrete partitions. A wide variety of different reagents or reagent types may be associated with the beads, where one may desire to deliver such reagents to a partition. Non-limiting examples of such reagents include, e.g., enzymes, polypeptides, antibodies or antibody fragments, labeling reagents, e.g., dyes, fluorophores, chromophores, etc., nucleic acids, polynucleotides, oligonucleotides, and any combination of two or more of the foregoing. In some cases, the beads may provide a surface upon which to synthesize or attach oligonucleotide sequences. Various entities including oligonucleotides, barcode sequences, primers, crosslinkers and the like may be associated with the outer surface of a bead. In the case of porous beads, an entity may be associated with both the outer and inner surfaces of a bead. The entities may be attached directly to the surface of a bead (*e.g.*, via a covalent bond, ionic bond, van der Waals interactions, etc.), may be attached to other oligonucleotide sequences attached to the surface of a bead (*e.g.* adaptor or primers), may be diffused throughout the interior of a bead and/or may be combined with a bead in a partition (*e.g.* fluidic droplet). In preferred embodiments, the oligonucleotides are covalently attached to sites within the polymeric matrix of the bead and are therefore present within the interior and exterior of the bead. In some cases, an entity such as a cell or nucleic acid is encapsulated within a bead. Other entities including amplification reagents (*e.g.*, PCR reagents, primers) may also be diffused throughout the bead or chemically-linked within the interior (*e.g.*, via pores, covalent attachment to polymeric matrix) of a bead.

[00103] Beads may serve to localize entities or samples. In some embodiments, entities (*e.g.* oligonucleotides, barcode sequences, primers, crosslinkers, adaptors and the like) may be associated with the outer and/or an inner surface of the bead. In some cases, entities may be located throughout the bead. In some cases, the entities may be associated with the entire surface of a bead or with at least half the surface of the bead.

[00104] Beads may serve as a support on which to synthesize oligonucleotide sequences. In some embodiments, synthesis of an oligonucleotide may comprise a ligation step. In some cases, synthesis of an oligonucleotide may comprise ligating two smaller oligonucleotides together. In some cases, a primer extension or other amplification reaction may be used to synthesize an oligonucleotide on a bead via a primer attached to the bead. In such cases, a primer attached to the bead may hybridize to a primer binding site of an oligonucleotide that also contains a template nucleotide sequence. The primer can then be extended by an primer extension reaction or other amplification reaction, and an oligonucleotide complementary to the template oligonucleotide can thereby be attached to the bead. In some cases, a set of identical oligonucleotides associated with a bead may be ligated to a set of diverse oligonucleotides, such that each identical oligonucleotide is attached to a different member of the diverse set of

oligonucleotides. In other cases, a set of diverse oligonucleotides associated with a bead may be ligated to a set of identical oligonucleotides.

Bead Characteristics

[00105] The methods, compositions, devices, and kits of this disclosure may be used with any suitable bead. In some embodiments, a bead may be porous, non-porous, solid, semi-solid, semi-fluidic, or fluidic. In some embodiments, a bead may be dissolvable, disruptable, or degradable. In some cases, a bead may not be degradable. In some embodiments, the bead may be a gel bead. A gel bead may be a hydrogel bead. A gel bead may be formed from molecular precursors, such as a polymeric or monomeric species. A semi-solid bead may be a liposomal bead. Solid beads may comprise metals including iron oxide, gold, and silver. In some cases, the beads are silica beads. In some cases, the beads are rigid. In some cases, the beads may be flexible.

[00106] In some embodiments, the bead may contain molecular precursors (*e.g.*, monomers or polymers), which may form a polymer network via polymerization of the precursors. In some cases, a precursor may be an already polymerized species capable of undergoing further polymerization via, for example, a chemical cross-linkage. In some cases, a precursor comprises one or more of an acrylamide or a methacrylamide monomer, oligomer, or polymer. In some cases, the bead may comprise prepolymers, which are oligomers capable of further polymerization. For example, polyurethane beads may be prepared using prepolymers. In some cases, the bead may contain individual polymers that may be further polymerized together. In some cases, beads may be generated via polymerization of different precursors, such that they comprise mixed polymers, co-polymers, and/or block co-polymers.

[00107] A bead may comprise natural and/or synthetic materials, including natural and synthetic polymers. Examples of natural polymers include proteins and sugars such as deoxyribonucleic acid, rubber, cellulose, starch (*e.g.* amylose, amylopectin), proteins, enzymes, polysaccharides, silks, polyhydroxyalkanoates, chitosan, dextran, collagen, carrageenan, ispaghula, acacia, agar, gelatin, shellac, sterculia gum, xanthan gum, Corn sugar gum, guar gum, gum karaya, agarose, alginic acid, alginate, or natural polymers thereof. Examples of synthetic polymers include acrylics, nylons, silicones, spandex, viscose rayon, polycarboxylic acids, polyvinyl acetate, polyacrylamide, polyacrylate, polyethylene glycol, polyurethanes, polylactic acid, silica, polystyrene, polyacrylonitrile, polybutadiene, polycarbonate, polyethylene, polyethylene terephthalate, poly(chlorotrifluoroethylene), poly(ethylene oxide), poly(ethylene terephthalate), polyethylene, polyisobutylene, poly(methyl methacrylate), poly(oxymethylene), polyformaldehyde, polypropylene, polystyrene, poly(tetrafluoroethylene), poly(vinyl acetate), poly(vinyl alcohol), poly(vinyl chloride), poly(vinylidene dichloride), poly(vinylidene difluoride), poly(vinyl fluoride) and combinations (*e.g.*, co-polymers) thereof. Beads may also be formed from materials other than polymers, including lipids, micelles, ceramics, glass-ceramics, material composites, metals, other inorganic materials, and others.

[00108] In some cases, a chemical cross-linker may be a precursor used to cross-link monomers during polymerization of the monomers and/or may be used to functionalize a bead with a species. In some

cases, polymers may be further polymerized with a cross-linker species or other type of monomer to generate a further polymeric network. Non-limiting examples of chemical cross-linkers (also referred to as a "crosslinker" or a "crosslinker agent" herein) include cystamine, gluteraldehyde, dimethyl suberimidate, N-Hydroxysuccinimide crosslinker BS3, formaldehyde, carbodiimide (EDC), SMCC, Sulfo-SMCC, vinylsilance, N,N'diallyltartardiamide (DATD), N,N'-Bis(acryloyl)cystamine (BAC), or homologs thereof. In some cases, the crosslinker used in the present disclosure contains cystamine.

[00109] Crosslinking may be permanent or reversible, depending upon the particular crosslinker used. Reversible crosslinking may allow for the polymer to linearize or dissociate under appropriate conditions. In some cases, reversible cross-linking may also allow for reversible attachment of a material bound to the surface of a bead. In some cases, a cross-linker may form disulfide linkages. In some cases, the chemical cross-linker forming disulfide linkages may be cystamine or a modified cystamine. In some embodiments, disulfide linkages may be formed between molecular precursor units *(e.g.* monomers, oligomers, or linear polymers). In some embodiments, disulfide linkages may be may be formed between molecular precursor units *(e.g.* monomers, oligomers, or linear polymers) or precursors incorporated into a bead and oligonucleotides.

[00110] Cystamine (including modified cystamines), for example, is an organic agent comprising a disulfide bond that may be used as a crosslinker agent between individual monomeric or polymeric precursors of a bead. Polyacrylamide may be polymerized in the presence of cystamine or a species comprising cystamine (e.g., a modified cystamine) to generate polyacrylamide gel beads comprising disulfide linkages (e.g., chemically degradable beads comprising chemically-reducible cross-linkers). The disulfide linkages may permit the bead to be degraded (or dissolved) upon exposure of the bead to a reducing agent.

[00111] In at least one alternative example, chitosan, a linear polysaccharide polymer, may be crosslinked with glutaraldehyde via hydrophilic chains to form a bead. Crosslinking of chitosan polymers may be achieved by chemical reactions that are initiated by heat, pressure, change in pH, and/or radiation.

[00112] In some embodiments, the bead may comprise covalent or ionic bonds between polymeric precursors (*e.g.* monomers, oligomers, linear polymers), oligonucleotides, primers, and other entities. In some cases, the covalent bonds comprise carbon-carbon bonds or thioether bonds.

[00113] In some cases, a bead may comprise an acrydite moiety, which in certain aspects may be used to attach one or more species (*e.g.*, barcode sequence, primer, other oligonucleotide) to the bead. In some cases, an acrydite moiety can refer to an acrydite analogue generated from the reaction of acrydite with one or more species, such as, for example, the reaction of acrydite with other monomers and cross-linkers during a polymerization reaction. Acrydite moieties may be modified to form chemical bonds with a species to be attached, such as an oligonucleotide (*e.g.*, barcode sequence, primer, other oligonucleotide). For example, acrydite moieties may be modified with thiol groups capable of forming a, disulfide bond or may be modified with groups already comprising a disulfide bond. The thiol or disulfide (via disulfide exchange) may be used as an anchor point for a species to be attached or another part of the acrydite moiety may be used for attachment. In some cases, attachment is reversible, such that when the disulfide

bond is broken (*e.g.*, in the presence of a reducing agent), the agent is released from the bead. In other cases, an acrydite moiety comprises a reactive hydroxyl group that may be used for attachment.

[00114] Functionalization of beads for attachment of other species, e.g., nucleic acids, may be achieved through a wide range of different approaches, including activation of chemical groups within a polymer, incorporation of active or activatable functional groups in the polymer structure, or attachment at the pre-polymer or monomer stage in bead production.

[00115] For example, in some examples, precursors (e.g., monomers, cross-linkers) that are polymerized to form a bead may comprise acrydite moieties, such that when a bead is generated, the bead also comprises acrydite moieties. Often, the acrydite moieties are attached to an oligonucleotide sequence, such as a primer (e.g., a primer for one or more of amplifying target nucleic acids and/or sequencing target nucleic acids barcode sequence, binding sequence, or the like)) that is desired to be incorporated into the bead. In some cases, the primer comprises a P5 sequence. For example, acrylamide precursors (e.g., cross-linkers, monomers) may comprise acrydite moieties such that when they are polymerized to form a bead, the bead also comprises acrydite moieties.

[00116] In some cases, precursors such as monomers and cross-linkers may comprise, for example, a single oligonucleotide (e.g., such as a primer or other sequence) or other species. **Fig 29A** depicts an example monomer comprising an acrydite moiety and single P5 sequence linked to the acrydite moiety via a disulfide bond. In some cases, precursors such as monomers and cross-linkers may comprise multiple oligonucleotides, other sequences, or other species. **Fig 29B** depicts an example monomer comprising multiple acrydite moieties each linked to a P5 primer via a disulfide bond. Moreover, **Fig 29C** depicts an example cross-linker comprising multiple acrydite moieties each linked to a P5 species via a disulfide bond. The inclusion of multiple acrydite moieties or other linker species in each precursor may improve loading of a linked species (e.g., an oligonucleotide) into beads generated from the precursors because each precursor can comprise multiple copies of a species to be loaded.

[00117] In some cases, precursors comprising a functional group that is reactive or capable of being activated such that it becomes reactive can be polymerized with other precursors to generate gel beads comprising the activated or activatable functional group. The functional group may then be used to attach additional species (e.g., disulfide linkers, primers, other oligonucleotides, etc.) to the gel beads. For example, some precursors comprising a carboxylic acid (COOH) group can co-polymerize with other precursors to form a gel bead that also comprises a COOH functional group, as shown in **Fig 31**. In some cases, acrylic acid (a species comprising free COOH groups), acrylamide, and bis(acryloyl)cystamine can be co-polymerized together to generate a gel bead comprising free COOH groups. The COOH groups of the gel bead can be activated (e.g., via 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and N-Hydroxysuccinimide (NHS) or 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM) as shown in **Fig 31**) such that they are reactive (e.g., reactive to amine functional groups where EDC/NHS or DMTMM are used for activation). The activated COOH groups can then react with an appropriate species (e.g., a species comprising an amine functional group where the carboxylic acid

groups are activated to be reactive with an amine functional group) comprising a moiety to be linked to the bead.

[00118] An example species comprising an amine group linked to a P5 primer via a disulfide bond (e.g., H_2N-C_6-S-S-C_6-P_5) is shown in **Fig 31**. COOH functional groups of a gel bead can be activated with EDC/NHS or DMTMM to generate an amine reactive species at one or more of the COOH sites. The amine group of the species H_2N-C_6-S-S-C_6-P_5 moiety can then react with the activated carboxylic acid such that the moiety and attached P5 oligonucleotide becomes covalently linked to the bead as shown in **Fig 31**. Unreacted COOH species can be converted to other species such that they are blocked.

[00119] Beads comprising disulfide linkages in their polymeric network may be functionalized with additional species via reduction of some of the disulfide linkages to free thiols. The disulfide linkages may be reduced via, for example, the action of a reducing agent (e.g., DTT, TCEP, etc.) to generate free thiol groups, without dissolution of the bead. Free thiols of the beads can then react with free thiols of a species or a species comprising another disulfide bond (e.g., via thiol-disulfide exchange)) such that the species can be linked to the beads (e.g., via a generated disulfide bond). In some cases, though, free thiols of the beads may react with any other suitable group. For example, free thiols of the beads may react with species comprising an acrydite moiety. The free thiol groups of the beads can react with the acrydite via Michael addition chemistry, such that the species comprising the acrydite is linked to the bead. In some cases, uncontrolled reactions can be prevented by inclusion of a thiol capping agent such as, for example, N-ethylmalieamide or iodoacetate.

[00120] Activation of disulfide linkages within a bead can be controlled such that only a small number of disulfide linkages are activated. Control may be exerted, for example, by controlling the concentration of a reducing agent used to generate free thiol groups and/or concentration of reagents used to form disulfide bonds in bead polymerization. In some cases, a low concentration (e.g., molecules of reducing agent:gel bead ratios of less than about 10000, 100000, 1000000, 10000000, 100000000, 1000000000, 10000000000, or 100000000000) of reducing agent may be used for reduction. Controlling the number of disulfide linkages that are reduced to free thiols may be useful in ensuring bead structural integrity during functionalization. In some cases, optically-active agents, such as fluorescent dyes may be may be coupled to beads via free thiol groups of the beads and used to quantify the number of free thiols present in a bead and/or track a bead.

[00121] An example scheme for functionalizing gel beads comprising disulfide linkages is shown in **Fig 35A**. As shown, beads **3501** (e.g., gel beads) comprising disulfide linkages can be generated using, for example, any of the methods described herein. Upon action of a reducing agent **3502** (e.g., DTT, TCEP, or any other reducing agent described herein) at a concentration not suitable for bead degradation, some of the gel bead **3501** disulfide linkages can be reduced to free thiols to generate beads **3503** comprising free thiol groups. Upon removal of the reducing agent (e.g., via washing) **3504**, beads **3503** can be reacted with an acrydite-S-S-species moiety **3505** comprising a species to be loaded (e.g., P5 oligonucleotide shown, but the species may be another type of polynucleotide such as, for, example, an oligonucleotide comprising P5, a barcode sequence, R1, and a random N-mer) linked to the acrydite via a

disulfide bond. Moiety **3505** can couple with the gel beads **3503** via Michael addition chemistry to generate beads **3506** comprising moiety **3505**. The generated beads **3506** can then be purified (e.g., via washing) by removing unwanted (e.g., non-attached) species.

[00122] Another example scheme for functionalizing gel beads comprising disulfide linkages is shown in **Fig 35B**. As shown, beads **3501** (e.g., gel beads) comprising disulfide linkages can be generated using, for example, any of the methods described herein. Upon action of a reducing agent **3502** (e.g., DTT, TCEP, or any other reducing agent described herein) at a concentration not suitable for bead degradation, some of the gel beads **3501** disulfide linkages can be reduced to free thiols to generate beads **3503** comprising free thiol groups. Upon removal of the reducing agent (e.g., via washing) **3504**, beads **3503** can be reacted with 2,2'-Dithiopyridine **3507** to generate gel beads **3509** linked to a pyridine moiety via a disulfide bond. As an alternative to 2,2'-Dithiopyridine, other similar species, such as 4,4'-Dithiopyridine or 5,5'-dithiobis-(2-nitrobenzoic acid) (e.g., DTNB or Ellman's Reagent) may be used. 2,2'-Dithiopyridine **3507** can couple with the gel beads **3503** via disulfide exchange to generate beads **3509** comprising a pyridine moiety linked to the beads **3509** via a disulfide bond. Gel beads **3509** can then be separated from unreacted species (e.g., via washing).

[00123] The purified gel beads **3509** can then be reacted with a moiety **3508** comprising a species of interest (e.g., a P5 oligonucleotide as shown) to be coupled to the gel beads and a free thiol group. In some cases, moiety **3508** may be generated from another species comprising a disulfide bond, such that when the disulfide bond is reduced (e.g., via the action of a reducing agent such as DTT, TCEP, etc.), moiety **3508** with a free thiol group is obtained. Moiety **3508** can participate in thiol-disulfide exchange with the pyridine group of beads **3509** to generate gel beads **3510** comprising moiety **3508**. The pyridine group is generally a good leaving group, which can permit effective thiol-disulfide exchange with the free thiol of moiety **3508**. The generated beads **3510** can then be purified (e.g., via washing) by removing unwanted species.

[00124] In some cases, addition of moieties to a gel bead after gel bead formation may be advantageous. For example, addition of a species after gel bead formation may avoid loss of the species during chain transfer termination that can occur during polymerization. Moreover, smaller precursors (e.g., monomers or cross linkers that do not comprise side chain groups and linked moieties) may be used for polymerization and can be minimally hindered from growing chain ends due to viscous effects. In some cases, functionalization after gel bead synthesis can minimize exposure of species (e.g., oligonucleotides) to be loaded with potentially damaging agents (e.g., free radicals) and/or chemical environments. In some cases, the generated gel may possess an upper critical solution temperature (UCST) that can permit temperature driven swelling and collapse of a bead. Such functionality may aid in species (e.g., a primer, a P5 primer) infiltration into the bead during subsequent functionalization of the bead with the species. Post-production functionalization may also be useful in controlling loading ratios of species in beads, such that, for example, the variability in loading ratio is minimized. Also, species loading may be performed in a batch process such that a plurality of beads can be functionalized with the species in a single batch.

[00125] In some cases, acrydite moieties linked to precursors, another species linked to a precursor, or a precursor itself comprise a labile bond, such as, for example, chemically, thermally, or photo-sensitive bonds e.g., disulfide bonds, UV sensitive bonds, or the like. Once acrydite moieties or other moieties comprising a labile bond are incorporated into a bead, the bead may also comprise the labile bond. The labile bond may be, for example, useful in reversibly linking (e.g., covalently linking) species (e.g., barcodes, primers, etc.) to a bead. In some cases, a thermally labile bond may include a nucleic acid hybridization based attachment, e.g., where an oligonucleotide is hybridized to a complementary sequence that is attached to the bead, such that thermal melting of the hybrid releases the oligonucleotide, e.g., a barcode containing sequence, from the bead or microcapsule. Moreover, the addition of multiple types of labile bonds to a gel bead may result in the generation of a bead capable of responding to varied stimuli. Each type of labile bond may be sensitive to an associated stimulus (e.g., chemical stimulus, light, temperature, etc.) such that release of species attached to a bead via each labile bond may be controlled by the application of the appropriate stimulus. Such functionality may be useful in controlled release of species from a gel bead. In some cases, another species comprising a labile bond may be linked to a gel bead after gel bead formation via, for example, an activated functional group of the gel bead as described above. As will be appreciated, barcodes that are releasably, cleavably or reversibly attached to the beads described herein include barcodes that are released or releasable through cleavage of a linkage between the barcode molecule and the bead, or that are released through degradation of the underlying bead itself, allowing the barcodes to be accessed or accessible by other reagents, or both. In general, the barcodes that are releasable as described herein, may generally be referred to as being activatable, in that they are available for reaction once released. Thus, for example, an activatable barcode may be activated by releasing the barcode from a bead (or other suitable type of partition described herein). As will be appreciated, other activatable configurations are also envisioned in the context of the described methods and systems. In particular, reagents may be provided releasably attached to beads, or otherwise disposed in partitions, with associated activatable groups, such that once delivered to the desired set of reagents, e.g., through co-partitioning, the activatable group may be reacted with the desired reagents. Such activatable groups include caging groups, removable blocking or protecting groups, e.g., photolabile groups, heat labile groups, or chemically removable groups.

[00126] In addition to thermally cleavable bonds, disulfide bonds and UV sensitive bonds, other non-limiting examples of labile bonds that may be coupled to a precursor or bead include an ester linkage (e.g., cleavable with an acid, a base, or hydroxylamine), a vicinal diol linkage (e.g., cleavable via sodium periodate), a Diels-Alder linkage (e.g., cleavable via heat), a sulfone linkage (e.g., cleavable via a base), a silyl ether linkage (e.g., cleavable via an acid), a glycosidic linkage (e.g., cleavable via an amylase), a peptide linkage (e.g., cleavable via a protease), or a phosphodiester linkage (e.g., cleavable via a nuclease (e.g., DNAase)).

[00127] A bead may be linked to a varied number of acrydite moieties. For example, a bead may comprise about 1, 10, 100, 1000, 10000, 100000, 1000000, 10000000, 100000000, 1000000000, or 10000000000 acrydite moieties linked to the beads. In other examples, a bead may comprise at least 1,

10, 100, 1000, 10000, 100000, 1000000, 10000000, 100000000, 1000000000, or 10000000000 acrydite moieties linked to the beads. For example, a bead may comprise about 1, 10, 100, 1000, 10000, 100000, 1000000, 10000000, 100000000, 1000000000, or 10000000000 oligonucleotides covalently linked to the beads, such as via an acrydite moiety. In other examples, a bead may comprise at least 1, 10, 100, 1000, 10000, 100000, 1000000, 10000000, 100000000, 1000000000, or 10000000000 oligonucleotides covalently linked to the beads, such as via an acrydite moiety.

[00128] Species that do not participate in polymerization may also be encapsulated in beads during bead generation (e.g., during polymerization of precursors). Such species may be entered into polymerization reaction mixtures such that generated beads comprise the species upon bead formation. In some cases, such species may be added to the gel beads after formation. Such species may include, for example, oligonucleotides, species necessary for a nucleic acid amplification reaction (e.g., primers, polymerases, dNTPs, co-factors (e.g., ionic co-factors)) including those described herein, species necessary for enzymatic reactions (e.g., enzymes, co-factors, substrates), or species necessary for a nucleic acid modification reaction such as polymerization, ligation, or digestion. Trapping of such species may be controlled by the polymer network density generated during polymerization of precursors, control of ionic charge within the gel bead (e.g., via ionic species linked to polymerized species), or by the release of other species. Encapsulated species may be released from a bead upon bead degradation and/or by application of a stimulus capable of releasing the species from the bead.

[00129] Beads may be of uniform size or heterogeneous size. In some cases, the diameter of a bead may be about 1μm, 5μm, 10μm, 20μm, 30μm, 40μm, 45μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 90μm, 100μm, 250μm, 500μm, or 1mm. In some cases, a bead may have a diameter of at least about 1μm, 5μm, 10μm, 20μm, 30μm, 40μm, 45μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 90μm, 100μm, 250μm, 500μm, 1mm, or more. In some cases, a bead may have a diameter of less than about 1μm, 5μm, 10μm, 20μm, 30μm, 40μm, 45μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 90μm, 100μm, 250μm, 500μm, or 1mm. In some cases, a bead may have a diameter in the range of about 40-75μm, 30-75μm, 20-75μm, 40-85μm, 40-95μm, 20-100μm, 10-100μm, 1-100μm, 20-250μm, or 20-500μm.

[00130] In certain preferred aspects, the beads are provided as a population of beads having a relatively monodisperse size distribution. As will be appreciated, in some applications, where it is desirable to provide relatively consistent amounts of reagents within partitions, maintaining relatively consistent bead characteristics, such as size, contributes to that overall consistency. In particular, the beads described herein may have size distributions that have a coefficient of variation in their cross-sectional dimensions of less than 50%, less than 40%, less than 30%, less than 20%, and in some cases less than 15%,less than 10%, or even less than 5%.

[00131] Beads may be of a regular shape or an irregular shape. Examples of bead shapes include spherical, non-spherical, oval, oblong, amorphous, circular, cylindrical, and homologs thereof.

Degradable Beads

[00132] In addition to, or as an alternative to the cleavable linkages between the beads and the associated molecules, e.g., barcode containing oligonucleotides, described above, the beads may be degradable, disruptable, or dissolvable spontaneously or upon exposure to one or more stimuli (*e.g.*, temperature changes, pH changes, exposure to particular chemical species or phase, exposure to light, reducing agent, etc.). In some cases, a bead may be dissolvable, such that material components of the beads are solubilized when exposed to a particular chemical species or an environmental changes, such as, for example, temperature, or pH. For example, a gel bead may be degraded or dissolved at elevated temperature and/or in basic conditions. In some cases, a bead may be thermally degradable such that when the bead is exposed to an appropriate change in temperature (e.g., heat), the bead degrades. Degradation or dissolution of a bead bound to a species (e.g., a nucleic acid species) may result in release of the species from the bead.

[00133] A degradable bead may comprise one or more species with a labile bond such that when the bead/species is exposed to the appropriate stimuli, the bond is broken and the bead degrades. The labile bond may be a chemical bond (*e.g.*, covalent bond, ionic bond) or may be another type of physical interaction (*e.g.*, van der Waals interactions, dipole-dipole interactions, etc.). In some cases, a crosslinker used to generate a bead may comprise a labile bond. Upon exposure to the appropriate conditions, the labile bond is broken and the bead is degraded. For example, a polyacrylamide gel bead may comprise cystamine crosslinkers. Upon exposure of the bead to a reducing agent, the disulfide bonds of the cystamine are broken and the bead is degraded.

[00134] A degradable bead may be useful in more quickly releasing an attached species (*e.g.*, an oligonucleotide, a barcode sequence) from the bead when the appropriate stimulus is applied to the bead. For example, for a species bound to an inner surface of a porous bead or in the case of an encapsulated species, the species may have greater mobility and accessibility to other species in solution upon degradation of the bead. In some cases, a species may also be attached to a degradable bead via a degradable linker (*e.g.*, disulfide linker). The degradable linker may respond to the same stimuli as the degradable bead or the two degradable species may respond to different stimuli. For example, a barcode sequence may be attached, via a disulfide bond, to a polyacrylamide bead comprising cystamine. Upon exposure of the barcoded-bead to a reducing agent, the bead degrades and the barcode sequence is released upon breakage of both the disulfide linkage between the barcode sequence and the bead and the disulfide linkages of the cystamine in the bead.

[00135] A degradable bead may be introduced into a partition, such as a droplet of an emulsion or a well, such that the bead degrades within the partition and any associated species are released within the droplet when the appropriate stimulus is applied. The free species may interact with other species. For example, a polyacrylamide bead comprising cystamine and linked, via a disulfide bond, to a barcode sequence, may be combined with a reducing agent within a droplet of a water-in-oil emulsion. Within the droplet, the reducing agent breaks the various disulfide bonds resulting in bead degradation and release of the barcode sequence into the aqueous, inner environment of the droplet. In another example, heating of a droplet

comprising a bead-bound barcode sequence in basic solution may also result in bead degradation and release of the attached barcode sequence into the aqueous, inner environment of the droplet.

[00136] As will be appreciated from the above disclosure, while referred to as degradation of a bead, in many instances as noted above, that degradation may refer to the disassociation of a bound or entrained species from a bead, both with and without structurally degrading the physical bead itself. For example, entrained species may be released from beads through osmotic pressure differences due to, for example, changing chemical environments. By way of example, alteration of bead pore sizes due to osmotic pressure differences can generally occur without structural degradation of the bead itself. In some cases, an increase in pore size due to osmotic swelling of a bead can permit the release of entrained species within the bead. In other cases, osmotic shrinking of a bead may cause a bead to better retain an entrained species due to pore size contraction.

[00137] As will be appreciated, where degradable beads are provided, it may be desirable to avoid exposing such beads to the stimulus or stimuli that cause such degradation prior to the desired time, in order to avoid premature bead degradation and issues that arise from such degradation, including for example poor flow characteristics, clumping and aggregation. By way of example, where beads comprise reducible cross-linking groups, such as disulfide groups, it will be desirable to avoid contacting such beads with reducing agents, e.g., DTT or other disulfide cleaving reagents. In such cases, treatments to the beads described herein will, in some cases be provided to be free of reducing agents, such as DTT. Because reducing agents are often provided in commercial enzyme preparations, it is often desirable to provide reducing agent free (or DTT free) enzyme preparations in treating the beads described herein. Examples of such enzymes include, e.g., polymerase enzyme preparations, ligase enzyme preparations, as well as many other enzyme preparations that may be used to treat the beads described herein. By "reducing agent free" or "DTT free" preparations means that the preparation will have less than 1/10th, less than $1/50^{th}$, and even less than $1/100^{th}$ of the lower ranges for such materials used in degrading the beads. For example, for DTT, the reducing agent free preparation will typically have less than 0.01 mM, 0.005 mM, 0.001 mM DTT, 0.0005 mM DTT, or even less than 0.0001 mM DTT or less. In many cases, the amount of DTT will be undetectable.

Methods for Degrading Beads

[00138] In some cases, a stimulus may be used to trigger degrading of the bead, which may result in the release of contents from the bead. Generally, a stimulus may cause degradation of the bead structure, such as degradation of the covalent bonds or other types of physical interaction. These stimuli may be useful in inducing a bead to degrade and/or to release its contents. Examples of stimuli that may be used include chemical stimuli, thermal stimuli, light stimuli and any combination thereof, as described more fully below.

[00139] Numerous chemical triggers may be used to trigger the degradation of beads. Examples of these chemical changes may include, but are not limited to pH-mediated changes to the integrity of a

component within the bead, degradation of a component of a bead via cleavage of cross-linked bonds, and depolymerization of a component of a bead.

[00140] In some embodiments, a bead may be formed from materials that comprise degradable chemical crosslinkers, such as BAC or cystamine. Degradation of such degradable crosslinkers may be accomplished through a number of mechanisms. In some examples, a bead may be contacted with a chemical degrading agent that may induce oxidation, reduction or other chemical changes. For example, a chemical degrading agent may be a reducing agent, such as dithiothreitol (DTT). Additional examples of reducing agents may include β-mercaptoethanol, (2*S*)-2-amino-1,4-dimercaptobutane (dithiobutylamine or DTBA), *tris*(2-carboxyethyl) phosphine (TCEP), or combinations thereof. A reducing agent may degrade the disulfide bonds formed between gel precursors forming the bead, and thus, degrade the bead. In other cases, a change in pH of a solution, such as an increase in pH, may trigger degradation of a bead. In other cases, exposure to an aqueous solution, such as water, may trigger hydrolytic degradation, and thus degrading the bead.

[00141] Beads may also be induced to release their contents upon the application of a thermal stimulus. A change in temperature can cause a variety of changes to a bead. For example, heat can cause a solid bead to liquefy. A change in heat may cause melting of a bead such that a portion of the bead degrades. In other cases, heat may increase the internal pressure of the bead components such that the bead ruptures or explodes. Heat may also act upon heat-sensitive polymers used as materials to construct beads.

[00142] The methods, compositions, devices, and kits of this disclosure may be used with any suitable agent to degrade beads. In some embodiments, changes in temperature or pH may be used to degrade thermo-sensitive or pH-sensitive bonds within beads. In some embodiments, chemical degrading agents may be used to degrade chemical bonds within beads by oxidation, reduction or other chemical changes. For example, a chemical degrading agent may be a reducing agent, such as DTT, wherein DTT may degrade the disulfide bonds formed between a crosslinker and gel precursors, thus degrading the bead. In some embodiments, a reducing agent may be added to degrade the bead, which may or may not cause the bead to release its contents. Examples of reducing agents may include dithiothreitol (DTT), β-mercaptoethanol, (2*S*)-2-amino-1,4-dimercaptobutane (dithiobutylamine or DTBA), *tris*(2-carboxyethyl) phosphine (TCEP), or combinations thereof. The reducing agent may be present at 0.1mM, 0.5mM, 1mM, 5mM, or 10mM. The reducing agent may be present at more than 0.1mM, 0.5mM, 1mM, 5mM, 10mM, or more. The reducing agent may be present at less than 0.1mM, 0.5mM, 1mM, 5mM, or 10mM.

Timing of Degrading Step

[00143] Beads may be degraded to release contents attached to and contained within the bead. This degrading step may occur simultaneously as the sample is combined with the bead. This degrading step may occur simultaneously when the sample is combined with the bead within a fluidic droplet that may be formed in a microfluidic device. This degrading step may occur after the sample is combined with the bead within a fluidic droplet that may be formed in a microfluidic device. As will be appreciated, in many applications, the degrading step may not occur.

[00144] The reducing agent may be combined with the sample and then with the bead. In some cases, the reducing agent may be introduced to a microfluidic device as the same time as the sample. In some cases, the reducing agent may be introduced to a microfluidic device after the sample is introduced. In some cases, the sample may be mixed with the reducing agent in a microfluidic device and then contacted with the gel bead in the microfluidic device. In some embodiments, the sample may be pre-mixed with the reducing agent and then added to the device and contacted with the gel bead.

[00145] A degradable bead may degrade instantaneously upon application of the appropriate stimuli. In other cases, degradation of the bead may occur over time. For example, a bead may degrade upon application of an appropriate stimulus instantaneously or within about 0, 0.01, 0.1, 0.5, 1, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 11, 12, 13, 14, 15 or 20 minutes. In other examples, a bead may degrade upon application of a proper stimulus instantaneously or within at most about 0, 0.01, 0.1, 0.5, 1, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 11, 12, 13, 14, 15 or 20 minutes.

[00146] Beads may also be degraded at different times, relative to combining with a sample. For example, the bead may be combined with the sample and subsequently degraded at a point later in time. The time between combining the sample with the bead and subsequently degrading the bead may be about 0.0001, 0.001, 0.01, 1, 10, 30, 60, 300, 600, 1800, 3600, 18000, 36000, 86400, 172800, 432000, or 864000 seconds. The time between combining the sample with the bead and subsequently degrading the bead may be more than about 0.0001, 0.001, 0.01, 1, 10, 30, 60, 300, 600, 1800, 3600, 18000, 36000, 86400, 172800, 432000, 864000 seconds or more. The time between combining the sample with the bead and subsequently degrading the bead may be less than about 0.0001, 0.001, 0.01, 1, 10, 30, 60, 300, 600, 1800, 3600, 18000, 36000, 86400, 172800, 432000, or 864000 seconds.

Preparing Beads Pre-functionalized with Oligonucleotides

[00147] The beads described herein may be produced using a variety of methods. In some cases, beads may be formed from a liquid containing molecular precursors (*e.g.* linear polymers, monomers, cross-linkers). The liquid is then subjected to a polymerization reaction, and thereby hardens or gels into a bead (or gel bead). The liquid may also contain entities such as oligonucleotides that become incorporated into the bead during polymerization. This incorporation may be via covalent or non-covalent association with the bead. For example, in some cases, the oligonucleotides may be entrained within a bead during formation. Alternatively, they may be coupled to the bead or the bead framework either during formation or following formation. Often, the oligonucleotides are connected to an acrydite moiety that becomes cross-linked to the bead during the polymerization process. In some cases, the oligonucleotides are attached to the acrydite moiety by a disulfide linkage. As a result, a composition comprising a bead-acrydite-S-S-oligonucleotide linkage is formed. **Fig 4A** is an exemplary diagram of a bead functionalized with an acrydite-linked primer.

[00148] In one exemplary process, functionalized beads may be generated by mixing a plurality of polymers and/or monomers with one or more oligonucleotides, such as, for example, one or more

oligonucleotides that comprises a primer (e.g., a universal primer, a sequencing primer). The polymers and/or monomers may comprise acrylamide and may be crosslinked such that disulfide bonds form between the polymers and/or monomers, resulting in the formation of hardened beads. The oligonucleotides may be covalently linked to the plurality of polymers and/or monomers during the formation of the hardened beads (e.g., contemporaneously) or may be covalently linked to the plurality of polymers and/or monomers after the formation of the hardened beads (e.g., sequentially). In some cases, the oligonucleotides may be linked to the beads via an acrydite moiety.

[00149] In most cases, a population of beads is pre-functionalized with the identical oligonucleotide such as a universal primer or primer binding site. In some cases, the beads in a population of beads are pre-functionalized with multiple different oligonucleotides. These oligonucleotides may optionally include any of a variety of different functional sequences, e.g., for use in subsequent processing or application of the beads. Functional sequences may include, e.g., primer sequences, such as targeted primer sequences, universal primer sequences, e.g., primer sequences that are sufficiently short to be able to hybridize to and prime extension from large numbers of different locations on a sample nucleic acid, or random primer sequences, attachment or immobilization sequences, ligation sequences, hairpin sequences, tagging sequences, e.g., barcodes or sample index sequences, or any of a variety of other nucleotide sequences.

[00150] By way of example, in some cases, the universal primer (*e.g.*, P5 or other suitable primer) may be used as a primer on each bead, to attach additional content (*e.g.*, barcodes, random N-mers, other functional sequences) to the bead. In some cases, the universal primer (e.g., P5) may also be compatible with a sequencing device, and may later enable attachment of a desired strand to a flow cell within the sequencing device. For example, such attachment or immobilization sequences may provide a complementary sequence to oligonucleotides that are tethered to the surface of a flow cell in a sequencing device, to allow immobilization of the sequences to that surface for sequencing. Alternatively, such attachments sequences may additionally be provided within, or added to the oligonucleotide sequences attached to the beads. In some cases, the beads and their attached species may be provided to be compatible with subsequent analytical process, such as sequencing devices or systems. In some cases, more than one primer may be attached to a bead and more than one primer may contain a universal sequence, in order to, for example, allow for differential processing of the oligonucleotide as well as any additional sequences that are coupled to that sequence, in different sequential or parallel processing steps, e.g., a first primer for amplification of a target sequence, with a second primer for sequencing the amplified product. For example, in some cases, the oligonucleotides attached to the beads will comprise a first primer sequence for conducting a first amplification or replication process, e.g., extending the primer along a target nucleic acid sequence, in order to generate an amplified barcoded target sequence(s). By also including a sequencing primer within the oligonucleotides, the resulting amplified target sequences will include such primers, and be readily transferred to a sequencing system. For example, in some cases, e.g., where one wishes to sequence the amplified targets using, e.g., an Illumina sequencing system, an R1 primer or primer binding site may also be attached to the bead.

[00151] Entities incorporated into the beads may include oligonucleotides having any of a variety of functional sequences as described above. For example, these oligonucleotides may include any one or more of P5, R1, and R2 sequences, non cleavable 5'acrydite-P5, a cleavable 5' acrydite-SS-P5, R1c, sequencing primer, read primer, universal primer, P5_U, a universal read primer, and/or binding sites for any of these primers. In some cases, a primer may contain one or more modified nucleotides nucleotide analogues, or nucleotide mimics. For example, in some cases, the oligonucleotides may include peptide nucleic acids (PNAs), locked nucleic acid (LNA) nucleotides, or the like. In some cases, these oligonucleotides may additionally or alternatively include nucleotides or analogues that may be processed differently, in order to allow differential processing at different steps of their application. For example, in some cases one or more of the functional sequences may include a nucleotide or analogue that is not processed by a particular polymerase enzyme, thus being uncopied in a process step utilizing that enzyme. For example, e.g., in some cases, one or more of the functional sequence components of the oligonucleotides will include, e.g., a uracil containing nucleotide, a nucleotide containing a non-native base, a blocker oligonucleotide, a blocked 3' end, 3'ddCTP. **Fig 19** provides additional examples. As will be appreciated, sequences of any of these entities may function as primers or primer binding sites depending on the particular application.

[00152] Polymerization may occur spontaneously. In some cases, polymerization may be initiated by an initiator and/or an accelerator, by electromagnetic radiation, by temperature changes (e.g., addition or removal of heat), by pH changes, by other methods, and combinations thereof. An initiator may refer to a species capable of initiating a polymerization reaction by activating (e.g., via the generation of free radicals) one or more precursors used in the polymerization reaction. An accelerator may refer to a species capable of accelerating the rate at which a polymerization reaction occurs. In some cases, an accelerator may speed up the activation of an initiator (e.g., via the generation of free radicals) used to then activate monomers (e.g., via the generation of free radicals) and, thus, initiate a polymerization reaction. In some cases, faster activation of an initiator can give rise to faster polymerization rates. In some cases, though, acceleration may also be achieved via non-chemical means such as thermal (e.g., addition and removal of heat) means, various types of radiative means (e.g., visible light, UV light, etc.), or any other suitable means. To create droplets containing molecular precursors, which may then polymerize to form hardened beads, an emulsion technique may be employed. For example, molecular precursors may be added to an aqueous solution. The aqueous solution may then be emulsified with an oil (*e.g.*, by agitation, microfluidic droplet generator, or other method). The molecular precursors may then be polymerized in the emulsified droplets to form the beads.

[00153] An emulsion may be prepared, for example, by any suitable method, including methods known in the art, such as bulk shaking, bulk agitation, flow focusing, and microsieve (See *e.g.*, Weizmann *et al.*, Nature Methods, 2006, 3(7):545-550; Weitz *et al.* U.S. Pub. No. 2012/0211084). In some cases, an emulsion may be prepared using a microfluidic device. In some cases, water-in-oil emulsions may be used. These emulsions may incorporate fluorosurfactants such as Krytox FSH with a PEG-containing compound such as bis krytox peg (BKP). In some cases, oil-in-water emulsions may be used. In some

cases, polydisperse emulsions may be formed. In some cases, monodisperse emulsions may be formed. In some cases, monodisperse emulsions may be formed in a microfluidic flow focusing device. (Gartecki *et al.*, Applied Physics Letters, 2004, 85(13):2649-2651).

[00154] In at least one example, a microfluidic device for making the beads may contain channel segments that intersect at a single cross intersection that combines two or more streams of immiscible fluids, such as an aqueous solution containing molecular precursors and an oil. Combining two immiscible fluids at a single cross intersection may cause fluidic droplets to form. The size of the fluidic droplets formed may depend upon the flow rate of the fluid streams entering the fluidic cross, the properties of the two fluids, and the size of the microfluidic channels. Initiating polymerization after formation of fluidic droplets exiting the fluidic cross may cause hardened beads to form from the fluidic droplets. Examples of microfluidic devices, channel networks and systems for generating droplets, both for bead formation and for partitioning beads into discrete droplets as discussed elsewhere herein, are described for example in U.S. Provisional Patent Application No. 61/977,804, filed April 4, 2014, and incorporated herein by reference in its entirety for all purposes.

[00155] To manipulate when individual molecular precursors, oligomers, or polymers begin to polymerize to form a hardened bead, an initiator and/or accelerator may be added at different points in the bead formation process. An accelerator may be an agent which may initiate the polymerization process (e.g., in some cases, via activation of a polymerization initiator) and thus may reduce the time for a bead to harden. In some cases, a single accelerator or a plurality of accelerators may be used for polymerization. Careful tuning of acceleration can be important in achieving suitable polymerization reactions. For example, if acceleration is too fast, weight and excessive chain transfer events may cause poor gel structure and low loading of any desired species. If acceleration is too slow, high molecular weight polymers can generate trapped activation sites (e.g., free radicals) due to polymer entanglement and high viscosities. High viscosities can impede diffusion of species intended for bead loading, resulting in low to no loading of the species. Tuning of accelerator action can be achieved, for example, by selecting an appropriate accelerator, an appropriate combination of accelerators, or by selecting the appropriate accelerator(s) and any stimulus (e.g., heat, electromagnetic radiation (e.g., light, UV light), another chemical species, etc.) capable of modulating accelerator action. Tuning of initiator action may also be achieved in analogous fashion.

[00156] An accelerator may be water-soluble, oil-soluble, or may be both water-soluble and oil-soluble. For example, an accelerator may be tetramethylethylenediamine (TMEDA or TEMED), dimethylethylenediamine, N,N, N,'N'- tetramethylmethanediamine, N,N' – dimorpholinomethane, or N,N,N',N'-Tetrakis(2-Hydroxypropyl)ethylenediamine. For example, an initiator may be ammonium persulfate (APS), calcium ions, or any of the compounds (I-IX) shown in **Fig 32**. The compounds (I-IX) shown in **Fig 32** can function as water-soluble azo-based initiators. Azo-based initiators may be used in the absence of TEMED and APS and can function as thermal based initiators. A thermal based initiator can activate species (e.g., via the generation of free radicals) thermally and, thus, the rate of initiator action can be tuned by temperature and/or the concentration of the initiator. A polymerization

accelerator or initiator may include functional groups including phosphonate, sulfonate, carboxylate, hydroxyl, albumin binding moieties, N-vinyl groups, and phospholipids. A polymerization accelerator or initiator may be a low molecular weight monomeric-compound. An accelerator or initiator may be a) added to the oil prior to droplet generation, b) added in the line after droplet generation, c) added to the outlet reservoir after droplet generation, or d) combinations thereof.

[00157] Polymerization may also be initiated by electromagnetic radiation. Certain types of monomers, oligomers, or polymers may contain light-sensitive properties. Thus, polymerization may be initiated by exposing such monomers, oligomers, or polymers to UV light, visible light, UV light combined with a sensitizer, visible light combined with a sensitizer, or combinations thereof. An example of a sensitizer may be riboflavin.

[00158] The time for a bead to completely polymerize or harden may vary depending on the size of the bead, whether an accelerator may be added, when an accelerator may be added, the type of initiator, when electromagnetic radiation may be applied, the temperature of solution, the polymer composition, the polymer concentration, and other relevant parameters. For example, polymerization may be complete after about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 minutes. Polymerization may be complete after more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 minutes or more. Polymerization may be complete in less than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 minutes.

[00159] Beads may be recovered from emulsions (*e.g.* gel-water-oil) by continuous phase exchange. Excess aqueous fluid may be added to the emulsion (*e.g.* gel-water-oil) and the hardened beads may be subjected to sedimentation, wherein the beads may be aggregated and the supernatant containing excess oil may be removed. This process of adding excess aqueous fluid followed by sedimentation and removal of excess oil may be repeated until beads are suspended in a given purity of aqueous buffer, with respect to the continuous phase oil. The purity of aqueous buffer may be about 80%, 90%, 95%, 96%, 97%, 98%, or 99% (v/v). The purity of aqueous buffer may be more than about 80%, 90%, 95%, 96%, 97%, 98%, 99% or more (v/v). The purity of aqueous buffer may be less than about 80%, 90%, 95%, 96%, 97%, 98%, or 99% (v/v). The sedimentation step may be repeated about 2, 3, 4, or 5 times. The sedimentation step may be repeated more than about 2, 3, 4, 5 times or more. The sedimentation step may be repeated less than about 2, 3, 4, or 5 times. In some cases, sedimentation and removal of the supernatant may also remove un-reacted starting materials.

[00160] Examples of droplet generators may include single flow focuser, parallel flow focuser, and microsieve membrane, such as those used by Nanomi B.V., and others. Preferably, a microfluidic device is used to generate the droplets.

[00161] An example emulsion based scheme for generating gel beads pre-functionalized with an acrydite moiety linked to a P5 primer via a disulfide bond is depicted in **Fig 30**. As shown in **Fig 30A**, acrylamide, bis(acryloyl)cystamine, acrydite-S-S-P5 moieties, and ammonium persulfate are combined into a droplets of an emulsion. TEMED can be added to the emulsion oil phase and can diffuse into the droplets to initiate the polymerization reaction. As shown in **Fig 30A**, TEMED action on ammonium persulfate

results in the generation of SO_4^- free radicals that can then activate the carbon-carbon double bond of the acrylamide via generation of a free radical at one of the carbons of the carbon-carbon double bond.

[00162] As shown in **Fig 30B,** activated acrylamide can react with non-activated acrylamide (again, at its carbon-carbon double bond) to begin polymerization. Each product generated can again be activated via the formation of a free radical resulting in polymer propagation. Moreover, both the bis(acryloyl)cystamine cross-linker and acrydite-S-S-P5 moieties comprise carbon-carbon double bonds that can react with activated species and the products themselves can then become activated. The inclusion of the bis(acryloyl)cystamine cross-linker into the polymerization reaction can result in cross-linking of polymer chains that are generated as shown in **Fig 30C**. Thus, a hydrogel polymer network comprising acrydite-S-S-P5 moieties linked to polymer backbones can be generated, as depicted in **Fig 30C**. The polymerization reaction can continue until it terminates. Upon reaction termination, continuous phase exchange or other suitable method can be used to break the emulsion and obtain gel beads comprising a cross-linked hydrogel (shown schematically in **Fig 30A**) coupled to the acrydite-S-S-P5 moieties.

Barcode and Random N-mers (introduction)

[00163] Certain applications, for example polynucleotide sequencing, may rely on unique identifiers ("barcodes") to identify a sequence and, for example, to assemble a larger sequence from sequenced fragments. Therefore, it may be desirable to add barcodes to polynucleotide fragments before sequencing. In the case of nucleic acid applications, such barcodes are typically comprised of a relatively short sequence of nucleotides attached to a sample sequence, where the barcode sequence is either known, or identifiable by its location or sequence elements. In some cases, a unique identifier may be useful for sample indexing. In some cases, though, barcodes may also be useful in other contexts. For example, a barcode may serve to track samples throughout processing (e.g., location of sample in a lab, location of sample in plurality of reaction vessels, etc.); provide manufacturing information; track barcode performance over time (e.g., from barcode manufacturing to use) and in the field; track barcode lot performance over time in the field; provide product information during sequencing and perhaps trigger automated protocols (e.g., automated protocols initiated and executed with the aid of a computer) when a barcode associated with the product is read during sequencing; track and troubleshoot problematic barcode sequences or product lots; serve as a molecular trigger in a reaction involving the barcode, and combinations thereof. In particularly preferred aspects, and as alluded to above, barcode sequence segments as described herein, can be used to provide linkage information as between two discrete determined nucleic acid sequences. This linkage information may include, for example, linkage to a common sample, a common reaction vessel, e.g., a well or partition, or even a common starting nucleic acid molecule. In particular, by attaching common barcodes to a specific sample component, or subset of sample components within a given reaction volume, one can attribute the resulting sequences bearing that barcode to that reaction volume. In turn, where the sample is allocated to that reaction volume based upon its sample of origin, the processing steps to which it is subsequently exposed, or on an individual

molecule basis, one can better identify the resulting sequences as having originated from that reaction volume.

[00164] Barcodes may be generated from a variety of different formats, including bulk synthesized polynucleotide barcodes, randomly synthesized barcode sequences, microarray based barcode synthesis, native nucleotides, partial complement with N-mer, random N-mer, pseudo random N-mer, or combinations thereof. Synthesis of barcodes is described herein, as well as in, for example, in U.S. Patent Application No. 14/175,973, filed February 7, 2014, the full disclosure of which is hereby incorporated herein by reference in its entirety for all purposes.

[00165] As described above, oligonucleotides incorporating barcode sequence segments, which function as a unique identifier, may also include additional sequence segments. Such additional sequence segments may include functional sequences, such as primer sequences, primer annealing site sequences, immobilization sequences, or other recognition or binding sequences useful for subsequent processing, e.g., a sequencing primer or primer binding site for use in sequencing of samples to which the barcode containing oligonucleotide is attached. Further, as used herein, the reference to specific functional sequences as being included within the barcode containing sequences also envisioned the inclusion of the complements to any such sequences, such that upon complementary replication will yield the specific described sequence.

[00166] In some examples, barcodes or partial barcodes may be generated from oligonucleotides obtained from or suitable for use in an oligonucleotide array, such as a microarray or bead array. In such cases, oligonucleotides of a microarray may be cleaved, (e.g., using cleavable linkages or moieties that anchor the oligonucleotides to the array (such as photoclevable, chemically cleavable, or otherwise cleavable linkages)) such that the free oligonucleotides are capable of serving as barcodes or partial barcodes. In some cases, barcodes or partial barcodes are obtained from arrays are of known sequence. The use of known sequences, including those obtained from an array, for example, may be beneficial in avoiding sequencing errors associated with barcodes of unknown sequence. A microarray may provide at least about 10,000,000, at least about 1,000,000, at least about 900,000, at least about 800,000, at least about 700,000, at least about 600,000, at least about 500,000, at least about 400,000, at least about 300,000, at least about 200,000, at least about 100,000, at least about 50,000, at least about 10,000, at least about 1,000, at least about 100, or at least about 10 different sequences that may be used as barcodes or partial barcodes.

[00167] The beads provided herein may be attached to oligonucleotide sequences that may behave as unique identifiers (*e.g.*, barcodes). Often, a population of beads provided herein contains a diverse library of barcodes, wherein each bead is attached to multiple copies of a single barcode sequence. In some cases, the barcode sequences are pre-synthesized and/or designed with known sequences. In some cases, each bead within the library is attached to a unique barcode sequence. In some cases, a plurality of beads will have the same barcode sequence attached to them. For example, in some cases about 1%, 2%, 3%, 4%, 5%, 10%, 20%, 25%, 30%, 50%, 75%, 80%, 90%, 95%, or 100% of the beads in a library are attached to a barcode sequence that is identical to a barcode sequence attached to a different bead in the

library. Sometimes, about 1%, 2%, 3%, 4%, 5%, 10%, 20%, 25%, or 30% of the beads are attached to the same barcode sequence.

[00168] The length of a barcode sequence may be any suitable length, depending on the application. In some cases, a barcode sequence may be about 2 to about 500 nucleotides in length, about 2 to about 100 nucleotides in length, about 2 to about 50 nucleotides in length, about 2 to about 20 nucleotides in length, about 6 to about 20 nucleotides in length, or about 4 to 16 nucleotides in length. In some cases, a barcode sequence is about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 85, 90, 95, 100, 150, 200, 250, 300, 400, or 500 nucleotides in length. In some cases, a barcode sequence is greater than about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 85, 90, 95, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 5000, or 10000 nucleotides in length. In some cases, a barcode sequence is less than about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 85, 90, 95, 100, 150, 200, 250, 300, 400, 500, 750, or 1000 nucleotides in length.

[00169] The barcodes may be loaded into beads so that one or more barcodes are introduced into a particular bead. In some cases, each bead may contain the same set of barcodes. In other cases, each bead may contain different sets of barcodes. In other cases, each bead may comprise a set of identical barcodes. In other cases, each bead may comprise a set of different barcodes.

[00170] The beads provided herein may be attached to oligonucleotide sequences that are random, pseudo-random, or targeted N-mers capable of priming a sample (*e.g.*, genomic sample) in a downstream process. In some cases, the same n-mer sequences will be present on the oligonucleotides attached to a single bead or bead population. This may be the case for targeted priming methods, e.g., where primers are selected to target certain sequence segments within a larger target sequence. In other cases, each bead within a population of beads herein is attached to a large and diverse number of N-mer sequences to, among other things, diversify the sampling of these primers against template molecules, as such random n-mer sequences will randomly prime against different portions of the sample nucleic acids.

[00171] The length of an N-mer may vary. In some cases, an N-mer (e.g., a random N-mer, a pseudo-random N-mer, or a targeted N-mer) may be between about 2 and about 100 nucleotides in length, between about 2 and about 50 nucleotides in length, between about 2 and about 20 nucleotides in length, between about 5 and about 25 nucleotides in length, or between about 5 and about 15 nucleotides in length. In some cases, an N-mer (e.g., a random N-mer, a pseudo-random N-mer, or a targeted N-mer) may be about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 85, 90, 95, 100, 150, 200, 250, 300, 400, or 500 nucleotides in length. In some cases, an N-mer (e.g., a random N-mer, a pseudo-random N-mer, or targeted a N-mer) may be greater than about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 85, 90, 95, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 5000, or 10000 nucleotides in length. In some cases, an N-mer (e.g., a random N-mer, a pseudo-random N-mer, or a targeted N-mer) may be less than about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 85, 90, 95, 100, 150, 200, 250, 300, 400, 500, 750, or 1000 nucleotides in length.

[00172] N-mers (including random N-mers) can be engineered for priming a specific sample type. For example, N-mers of different lengths may be generated for different types of sample nucleic acids or different regions of a sample nucleic acid, such that each N-mer length corresponds to each different type of sample nucleic acid or each different region of a sample nucleic acid. For example, an N-mer of one length may be generated for sample nucleic acid originating from the genome of one species (e.g., for example, a human genome) and an N-mer of another length may be generated for a sample nucleic acid originating from another species (e.g., for example, a yeast genome). In another example, an N-mer of one length may be generated for sample nucleic acid comprising a particular sequence region of a genome and an N-mer of another length may be generated for a sample nucleic acid comprising another sequence region of the genome. Moreover, in addition or as an alternative to N-mer length, the base composition of the N-mer (e.g., GC content of the N-mer) may also be engineered to correspond to a particular type or region of a sample nucleic acid. Base content may vary in a particular type of sample nucleic acid or in a particular region of a sample nucleic acid, for example, and, thus, N-mers of different base content may be useful for priming different sample types of nucleic acid or different regions of a sample nucleic acid.

[00173] Populations of beads described elsewhere herein can be generated with an N-mer engineered for a particular sample type or particular sample sequence region. In some cases, a mixed population of beads (e.g., a mixture of beads comprising an N-mer engineered for one sample type or sequence region and beads comprising another N-mer engineered for another sample type or sequence region) with respect to N-mer length and content may be generated. In some cases, a population of beads may be generated, where one or more of the beads can comprise a mixed population of N-mers engineered for a plurality of sample types or sequence regions.

[00174] As noted previously, in some cases, the N-mers, whether random or targeted, may comprise nucleotide analogues, mimics, or non-native nucleotides, in order to provide primers that have improved performance in subsequent processing steps. For example, in some cases, it may be desirable to provide N-mer primers that have different melting/annealing profiles when subjected to thermal cycling, e.g., during amplification, in order to enhance the relative priming efficiency of the n-mer sequence. In some cases, nucleotide analogues or non-native nucleotides may be incorporated into the N-mer primer sequences in order to alter the melting temperature profile of the primer sequence as compared to a corresponding primer that includes native nucleotides. In certain cases, the primer sequences, such as the N-mer sequences described herein, may include modified nucleotides or nucleotide analogues, e.g., LNA bases, at one or more positions within the sequence, in order to provide elevated temperature stability for the primers when hybridized to a template sequence, as well as provide generally enhanced duplex stability. In some cases, LNA nucleotides are used in place of the A or T bases in primer synthesis to replace those weaker binding bases with tighter binding LNA analogues. By providing enhanced hybridizing primer sequences, one may generate higher efficiency amplification processes using such primers, as well as be able to operate within different temperature regimes.

[00175] Other modifications may also be provided to the oligonucleotides described above. For example, in some cases, the oligonucleotides may be provided with protected termini or other regions, in order to

prevent or reduce any degradation of the oligonucleotides, e.g., through any present exonuclease activity. In one example, the oligonucleotides may be provided with one or more phosphorothioate nucleotide analogue at one or more positions within the oligonucleotide sequence, e.g., adjacent or proximal to the 3' and/or 5' terminal position. These phosphorothioate nucleotides typically provide a sulfur group in place of the non-linking oxygen in an internucleotide linkage within the oligonucleotide to reduce or eliminate nuclease activity on the oligonucleotides, including, e.g., 3'-5' and/or 5'-3' exonucleases. In general, phosphorothioate analogues are useful in imparting exo and/or endonuclease resistance to oligonucleotides that include them, including providing protection against, e.g., 3'-5' and/or 5'-3' exonuclease digestion of the oligonucleotides. Accordingly, in some aspects, these one or more phosphorothioate linkages will be in one or more of the last 5 to 10 internucleotide linkages at either the 3' or the 5' terminus of the oligonucleotides, and preferably include one or more of the last 3' or 5' terminal internucleotide linkage and second to last 5' terminal internucleotide linkage, in order to provide protection against 3'-5' or 5'-3' exonuclease activity. Other positions within the oligonucleotides may also be provided with phosphorothiate linkages as well. In addition to providing such protection on the oligonucleotides that comprise the barcode sequences (and any associated functional sequences), the above described modifications are also useful in the context of the blocker sequences described herein, e.g., incorporating phosphorothioate analogues within the blocker sequences, e.g., adjacent or proximal to the 3' and/or 5' terminal position as well as potentially other positions within the oligonucleotides.

Attaching Content to Pre-functionalized Beads

[00176] A variety of content may be attached to the beads described herein, including beads functionalized with oligonucleotides. Often, oligonucleotides are attached, particularly oligonucleotides with desired sequences (*e.g.*, barcodes, random N-mers). In many of the methods provided herein, the oligonucleotides are attached to the beads through a primer extension reaction. Beads pre-functionalized with primer can be contacted with oligonucleotide template. Amplification reactions may then be performed so that the primer is extended such that a copy of the complement of the oligonucleotide template is attached to the primer. Other methods of attachment are also possible such as ligation reactions.

[00177] In some cases, oligonucleotides with different sequences (or the same sequences) are attached to the beads in separate steps. For example, in some cases, barcodes with unique sequences are attached to beads such that each bead has multiple copies of a first barcode sequence on it. In a second step, the beads can be further functionalized with a second sequence. The combination of first and second sequences may serve as a unique barcode, or unique identifier, attached to a bead. The process may be continued to add additional sequences that behave as barcode sequences (in some cases, greater than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 barcode sequences are sequentially added to each bead). The beads may also be further functionalized random N-mers that can, for example, act as a random primer for downstream whole genome amplification reactions.

[00178] In some cases, after functionalization with a certain oligonucleotide sequence (e.g., barcode sequence), the beads may be pooled and then contacted with a large population of random Nmers that are then attached to the beads. In some cases, particularly when the beads are pooled prior to the attachment of the random Nmers, each bead has one barcode sequence attached to it, (often as multiple copies), but many different random Nmer sequences attached to it. **Fig 4** provides a step-by-step depiction of one example method, an example limiting dilution method, for attaching oligonucleotides, such as barcodes and Nmers, to beads.

[00179] Limiting dilution may be used to attach oligonucleotides to beads, such that the beads, on average, are attached to no more than one unique oligonucleotide sequence such as a barcode. Often, the beads in this process are already functionalized with a certain oligonucleotide, such as primers. For example, beads functionalized with primers (e.g., such as universal primers) and a plurality of template oligonucleotides may be combined, often at a high ratio of beads: template oligonucleotides, to generate a mixture of beads and template oligonucleotides. The mixture may then be partitioned into a plurality of partitions (e.g., aqueous droplets within a water-in-oil emulsion), such as by a bulk emulsification process, emulsions within plates, or by a microfluidic device, such as, for example, a microfluidic droplet generator. In some cases, the mixture can be partitioned into a plurality of partitions such that, on average, each partition comprises no more than one template oligonucleotide.

[00180] Moreover, the template oligonucleotides can be amplified (e.g., via primer extension reactions) within the partitions via the primers attached to the beads. Amplification can result in the generation of beads comprising amplified template oligonucleotides. Following amplification, the contents of the partitions may be pooled into a common vessel (e.g., a tube, a well, etc.). The beads comprising the amplified template oligonucleotides may then be separated from the other contents of the partitions (including beads that do not comprise amplified template oligonucleotides) by any suitable method including, for example, centrifugation and magnetic separation, with or without the aid of a capture moiety as described elsewhere herein.

[00181] Beads comprising amplified template oligonucleotides may be combined with additional template oligonucleotides to generate a bulk mixture comprising the beads and the additional template oligonucleotides. The additional template oligonucleotides may comprise a sequence that is at least partially complementary to the amplified template oligonucleotides on the beads, such that the additional template oligonucleotide hybridizes to the amplified template oligonucleotides. The amplified template oligonucleotides can then be extended via the hybridized additional template oligonucleotides in an amplification reaction, such that the complements of the additional template oligonucleotides are attached to the amplified template oligonucleotides. The cycle of binding additional template oligonucleotides to amplified oligonucleotides, followed by extension of the amplified oligonucleotides in an amplification reaction, can be repeated for any desired number of additional oligonucleotides that are to be added to the bead.

[00182] The oligonucleotides attached to the amplified template oligonucleotides may comprise, for example, one or more of a random N-mer sequence, a pseudo random N-mer sequence, or a primer

binding site (e.g.., a universal sequence portion, such as a universal sequence portion that is compatible with a sequencing device). Any of these sequences or any other sequence attached to a bead may comprise at least a subsection of uracil containing nucleotides, as described elsewhere herein.

[00183] An example of a limiting dilution method for attaching a barcode sequence and a random N-mer to beads is shown in **Fig 4**. As shown in **Fig 4A**, beads **401**, (e.g., disulfide cross-linked polyacrylamide gel beads) are pre-functionalized with a first primer **403**. The first primer **403** may be, for example, coupled to the beads via a disulfide linkage **402** with an acrydite moiety bound to the surface of the beads **401**. In some cases, though, first primer **403** may be coupled to a bead via an acrydite moiety, without a disulfide linkage **402**. The first primer **403** may be a universal primer for priming template sequences of oligonucleotides to be attached to the beads and/or may be a primer binding site (e.g., P5) for use in sequencing an oligonucleotide that comprises first primer **403**.

[00184] The first primer **403** functionalized beads **401** can then be mixed in an aqueous solution with template oligonucleotides (e.g., oligonucleotides comprising a first primer binding site **404** (e.g., P5c), a template barcode sequence **405**, and a template primer binding site **407** (e.g., R1c)) and reagents necessary for nucleic acid amplification (e.g., dNTPs, polymerase, co-factors, etc.) as shown in **Fig 4B**. The aqueous mixture may also comprise a capture primer **406** (e.g., sometimes referred to as a read primer) linked to a capture moiety (e.g., biotin), identical in sequence to the template primer binding site **407** of the template oligonucleotide.

[00185] The aqueous mixture is then emulsified in a water/oil emulsion to generate aqueous droplets (e.g., the droplets comprising one or more beads **401**, a template oligonucleotide, reagents necessary for nucleic amplification, and, if desired, any capture primers **406**) in a continuous oil phase. In general, the droplets comprise, on average, at most one template oligonucleotide per droplet. As shown in **Fig 4B** and **4C**, a first round of thermocycling of the droplets results in priming of the template oligonucleotides at primer binding site **404** by first primer **403** and extension of first primer **403** such that oligonucleotides complementary to the template oligonucleotide sequences are attached to the gel beads at first primer **403**. The complementary oligonucleotides comprises first primer **403**, a barcode sequence **408** (e.g., complementary to template barcode sequence **405**), and a capture primer binding site **415** complementary to both template primer binding site **407** and capture primer **406**. Capture primer binding site **415** may also be used as a read primer binding site (e.g., R1) during sequencing of the complementary oligonucleotide.

[00186] As shown in **Fig 4D,** capture primer **406** can bind to capture primer binding site **415** during the next round of thermocycling. Capture primer **406**, comprising a capture moiety (e.g., biotin) at its 5' end, can then be extended to generate additional template oligonucleotides (e.g., comprising sequences **404**, **405**, and **406**), as shown in **Fig 4E**. Thermocyling may continue for a desired number of cycles (e.g., at least about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 or more cycles) up until all first primer **403** sites of beads **401** are linked to a barcode sequence **408** and a capture primer binding site **415**. Because each droplet generally comprises one or zero template oligonucleotides to start, each droplet will generally comprise beads attached to multiple copies of a sequence complementary to the template oligonucleotide

or no copies of a sequence complementary to the template oligonucleotide. At the conclusion of thermocycling, the oligonucleotide products attached to the beads are hybridized to template oligonucleotides also comprising the capture moiety (e.g., biotin), as shown in **Fig 4E**.

[00187] The emulsion may then be broken via any suitable means and the released beads can be pooled into a common vessel. Using a capture bead (or other device, including capture devices described herein) **409** linked to a moiety (e.g., streptavidin) capable of binding with the capture moiety of capture primer **406**, positive beads (e.g., beads comprising sequences **403, 408,** and **415**) may be enriched from negative beads (e.g., beads not comprising sequences **403**, **408**, and **415**) by interaction of the capture bead with the capture moiety, as shown in **Fig 4F** and **Fig 4G**. In cases where capture beads are used, the beads may be magnetic, such that a magnet may be used for enrichment. As an alternative, centrifugation may be used for enrichment. Upon enrichment of the positive beads, the hybridized template oligonucleotides comprising the capture moiety and linked to the capture bead may be denatured from the bead-bound oligonucleotide via heat or chemical means, including chemical means described herein, as shown in **Fig 4H**. Denatured oligonucleotides (e.g., oligonucleotides comprising sequences **404**, **405** and **406**) may then be separated from the positive beads via the capture beads attached to the denatured oligonucleotides. As shown in **Fig 4H**, beads comprising sequences **403**, **408**, and **415** are obtained. As an alternative to capture beads, positive beads may also be sorted from positive beads via flow cytometry by including, for example, an optically active dye in partitions capable of binding to beads or species coupled to beads.

[00188] In bulk aqueous fluid, the beads comprising sequences **403**, **408**, and **415** can then be combined with template random sequences (e.g., random N-mers) **413** each linked to a sequence **412** complementary to capture primer binding site **415**, as shown in **Fig 4I**. As shown in **Fig 4J**, capture primer binding site **415** can prime oligonucleotides comprising template random sequences **413** at sequence **412** upon heating. Following priming, capture primer binding site **415** can be extended (e.g., via polymerase) to link capture primer binding site **415** with a random sequence **414** that is complementary to template random sequence **413**. Oligonucleotides comprising template random sequences **413** and sequence **412** can be denatured from the bead using heat or chemical means, including chemical means described herein. Centrifugation and washing of the beads, for example, may be used to separate the beads from denatured oligonucleotides. Following removal of the denatured oligonucleotides, beads comprising a barcode sequence **408** and a random sequence **414** are obtained, as shown in **Fig 4K , 4L,** and **4M**. Because the attachment of random sequence **414** was done in bulk, each bead that comprises multiple copies of a unique barcode sequence **408**, also comprises various random sequences **414**.

[00189] To release bead-bound oligonucleotides from the beads, stimuli described elsewhere herein, such as, for example, a reducing agent, may be used. As shown in **Fig 4N**, contact of a bead comprising disulfide bonds and linkages to oligonucleotides via disulfide bonds with a reducing agent degrades both the bead and the disulfide linkages freeing the oligonucleotide from the bead. Contact with a reducing agent may be completed, for example, in another partition (e.g., a droplet of another emulsion), such that, upon oligonucleotide release from the bead, each droplet generally comprises free oligonucleotides all

comprising the same barcode sequence **408**, yet various random sequences **414**. Via random sequence **414** acting as a random primer, free oligonucleotides may be used to barcode different regions of a sample nucleic acid also in the partition. Amplification or ligation schemes, including those described herein, may be used to complete attachment of barcodes to the sample nucleic acid.

[00190] With limiting dilution, the partitions (*e.g.*, droplets) may contain on average at most one oligonucleotide sequence per partition. This frequency of distribution at a given sequence-bead dilution follows Poisson distribution. Thus, in some cases, about 6%, 10%, 18%, 20%, 30%, 36%, 40%, or 50% of the droplets or partitions may comprise one or fewer oligonucleotide sequences. In some cases, more than about 6%, 10%, 18%, 20%, 30%, 36%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%,99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or more of the droplets may comprise one or fewer oligonucleotide sequences. In other cases, less than about 6%, 10%, 18%, 20%, 30%, 36%, 40%, or 50% of the droplets may comprise one or fewer oligonucleotide sequences.

[00191] In some cases, limiting dilution steps may be repeated, prior to the addition of a random N-mer sequence in order to increase the number of positive beads with copies of barcodes. For example, a limiting dilution could be prepared such that a desired fraction (e.g., 1/10 to 1/3) of emulsion droplets comprises a template for amplification. Positive beads could be generated via amplification of the template (as depicted in **Fig 4**) such that positives generally comprise no more primer for amplification (e.g., all P5 primer sites have been extended). The emulsion droplets can then be broken, and subsequently re-emulsified with fresh template at limiting dilution for a second round of amplification. Positive beads generated in the first round of amplification generally would not participate in further amplification because their priming sites would already be occupied. The process of amplification followed by re-emulsification can be repeated for a suitable number of steps, until the desired fraction of positive beads is obtained.

[00192] In some cases, negative beads obtained during sorting after a limiting dilution functionalization may be recovered and further processed to generate additional positive beads. For example, negative beads may be dispensed into wells of a plate (e.g., a 384 well plate) after recovery such that each well generally comprises 1 bead. In some cases, dispensing may be achieved with the aid of flow cytometry (e.g., a flow cytometer directs each negative bead into a well during sorting – an example flow cytometer being a BD FACS Jazz) or via a dispensing device, such as for example, a robotic dispensing device. Each well can also comprise a template barcode sequence and the process depicted in **Fig 4** repeated, except that each well partitions each bead, rather than a fluidic droplet. Because each well comprises template and a bead, each well can produce a positive bead. The beads can then be pooled from each well and additional sequences (e.g., a random N-mer sequence) can be added in bulk as described elsewhere herein.

[00193] The barcodes may be loaded into the beads at an expected or predicted ratio of barcodes per bead to be barcoded. In some cases, the barcodes are loaded such that a ratio of about 0.0001, 0.001, 0.1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 500000, 1000000,

5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 barcodes are loaded per bead. In some cases, the barcodes are loaded such that a ratio of more than 0.0001, 0.001, 0.1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 30000000, 40000000, 50000000, 60000000, 70000000, 80000000, 90000000, 100000000, 200000000, 300000000, 400000000, 500000000, 600000000, 700000000, 800000000, 900000000, 1000000000, 2000000000, 3000000000, 4000000000, 5000000000, 6000000000, 7000000000, 8000000000, 9000000000, 10000000000, 20000000000, 30000000000, 40000000000, 50000000000, 60000000000, 70000000000, 80000000000, 90000000000, 100000000000 or more barcodes are loaded per bead. In some cases, the barcodes are loaded such that a ratio of less than about 0.0001, 0.0002, 0.0003, 0.0004, 0.0005, 0.0006, 0.0007, 0.0008, 0.0009, 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 barcodes are loaded per bead.

[00194] Beads, including those described herein (e.g., substantially dissolvable beads, in some cases, substantially dissolvable by a reducing agent), may be covalently or non-covalently linked to a plurality of oligonucleotides, wherein at least a subset of the oligonucleotides comprises a constant region or domain (e.g., a barcode sequence, a barcode domain, a common barcode domain, or other sequence that is constant among the oligonucleotides of the subset) and a variable region or domain (e.g., a random sequence, a random N-mer, or other sequence that is variable among the oligonucleotides of the subset). In some cases, the oligonucleotides may be releasably coupled to a bead, as described elsewhere herein. Oligonucleotides may be covalently or non-covalently linked to a bead via any suitable linkage, including types of covalent and non-covalent linkages described elsewhere herein. In some cases, an oligonucleotide may be covalently linked to a bead via a cleavable linkage such as, for example, a chemically cleavable linkage (e.g., a disulfide linkage), a photocleavable linkage, or a thermally cleavable linkage. Beads may comprise more than about or at least about 1, 10, 50, 100, 500, 1000, 5000, 10000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, 5000000000, 10000000000, 50000000000, 100000000000, 500000000000, or 1000000000000 oligonucleotides comprising a constant region or domain and a variable region or domain.

[00195] In some cases, the oligonucleotides may each comprise an identical constant region or domain (e.g., an identical barcode sequence, identical barcode domain, a common domain, etc.). In some cases, the oligonucleotides may each comprise a variable domain with a different sequence. In some cases, the percentage of the oligonucleotides that comprise an identical constant region (or common domain) may be at least about 0.01%, 0.1%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%. In some cases, the percentage of the oligonucleotides that comprise a variable region with a different sequence may be at least about 0.01%, 0.1%, 1%, 5%,

10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%. In some cases, the percentage of beads in a plurality of beads that comprise oligonucleotides with different nucleotide sequences (including those comprising a variable and constant region or domain) is at least about 0.01%, 0.1%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%. In some cases, the oligonucleotides may also comprise one or more additional sequences, such as, for example a primer binding site (e.g., a sequencing primer binding site), a universal primer sequence (e.g., a primer sequence that would be expected to hybridize to and prime one or more loci on any nucleic acid fragment of a particular length, based upon the probability of such loci being present within a sequence of such length) or any other desired sequence including types of additional sequences described elsewhere herein.

[00196] As described elsewhere herein, a plurality of beads may be generated to form, for example, a bead library (e.g., a barcoded bead library). In some cases, the sequence of a common domain (e.g., a common barcode domain) or region may vary between at least a subset of individual beads of the plurality. For example, the sequence of a common domain or region between individual beads of a plurality of beads may be different between 2 or more, 10 or more, 50 or more, 100 or more, 500 or more, 1000 or more, 5000 or more, 10000 or more, 50000 or more, 100000 or more, 500000 or more, 1000000 or more, 5000000 or more, 10000000 or more, 50000000 or more, 100000000 or more, 500000000 or more, 1000000000 or more, 5000000000 or more, 10000000000 or more, 50000000000 or more, or 100000000000 or more beads of the plurality. In some cases, each bead of a plurality of beads may comprise a different common domain or region. In some cases, the percentage of individual beads of a plurality of beads that comprise a different common domain or region may be at least about 0.01%, 0.1%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%. In some cases, a plurality of beads may comprise at least about 2, 10, 50, 100, 500, 1000, 5000, 10000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, or more different common domains coupled to different beads in the plurality.

[00197] As an alternative to limiting dilution (e.g., via droplets of an emulsion), other partitioning methods may be used to attach oligonucleotides to beads. As shown in **Fig 13A**, the wells of a plate may be used. Beads comprising a primer (e.g., P5, primer linked to the bead via acrydite and, optionally, a disulfide bond) may be combined with a template oligonucleotide (e.g., a template oligonucleotide comprising a barcode sequence) and amplification reagents in the wells of a plate. Each well can comprise one or more copies of a unique template barcode sequence and one or more beads. Thermal cycling of the plate extends the primer, via hybridization of the template oligonucleotide to the primer, such that the bead comprises an oligonucleotide with a sequence complementary to the oligonucleotide template. Thermal cycling may continue for a desired number of cycles (e.g., at least about 1, 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 or more cycles) up until all primers have been extended.

[00198] Upon completion of thermal cycling, the beads may be pooled into a common vessel, washed (e.g., via centrifugation, magnetic separation, etc.), complementary strands denatured, washed again, and then subject to additional rounds of bulk processing if desired. For example, a random N-mer sequence

may be added to the bead-bound oligonucleotides using the primer extension method described above for limiting dilution and as shown in **Fig 13B** and **Fig 4I-M**.

[00199] As another alternative approach to limiting dilution, a combinatorial process involving partitioning in multiwell plates can be used to generate beads with oligonucleotide sequences as shown in **Fig 13C**. In such methods, the wells may contain pre-synthesized oligonucleotides such as oligonucleotide templates. The beads (e.g., beads with preincorporated oligonucleotides such as primers) may be divided into the individual wells of the multiwell plate. For example, a mixture of beads containing P5 oligonucleotides may be divided into individual wells of a multiwell plate (*e.g.*, 384 wells), wherein each well contains a unique oligonucleotide template (e.g., an oligonucleotide including a first partial barcode template or barcode template). A primer extension reaction may be performed within the individual wells using, for example, the oligonucleotides templates as the template and the primer attached to the beads as primers. Subsequently, all wells may be pooled together and the unreacted products may be removed.

[00200] The mixture of beads attached to the amplified product may be re-divided into wells of a second multiwell plate (*e.g.*, 384-well plate), wherein each well of the second multiwell plate contains another oligonucleotide sequence (e.g., including a second partial barcode sequence and/or a random N-mer). In some cases, the oligonucleotide sequence may be attached (e.g., via hybridization) to a blocker oligonucleotide. Within the wells of the second multiwell plate, a reaction such as a single-stranded ligation reaction may be performed to add additional sequences to each bead (e.g., via ligation of the primer extension products attached to the beads as in the first step with the oligonucleotide in the wells of the second step). In some cases, a partial barcode sequence linked to the bead in the first step is ligated to a second partial barcode sequence in the second step, to generate beads comprising full barcode sequences. In some cases, the beads comprising full barcode sequences also comprise random sequences (e.g., random N-mers) and/or blocking oligonucleotides. In some cases, a PCR reaction or primer extension reaction is performed to attach the additional sequence to the beads. Beads from the wells may be pooled together, and the unreacted products may be removed. In some cases, the process is repeated with additional multi-well plates. The process may be repeated over 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 50, 100, 500, 1000, 5000, or 10000 times.

[00201] In some combinatorial approaches, ligation methods may be used to assemble oligonucleotide sequences comprising barcode sequences on beads (e.g., degradable beads as described elsewhere herein). For example, separate populations of beads may be provided to which barcode containing oligonucleotides are to be attached. These populations may include anchor components (or linkage) for attaching nucleotides, such as activatable chemical groups (phosphoramidites, acrydite moieties, or other thermally, optically or chemically activatable groups), cleavable linkages, previously attached oligonucleotide molecules to which the barcode containing oligonucleotides may be ligated, hybridized, or otherwise attached, DNA binding proteins, charged groups for electrostatic attachment, or any of a variety of other attachment mechanisms.

[00202] A first oligonucleotide or oligonucleotide segment that includes a first barcode sequence segment, is attached to the separate populations, where different populations include different barcode sequence segments attached thereto. Each bead in each of the separate populations may be attached to at least 2, 10, 50, 100, 500, 1000, 5000, 10000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, or more first oligonucleotide molecules or oligonucleotide segment molecules. The first oligonucleotide or oligonucleotide segment may be releasably attached to the separate populations. In some cases, the first oligonucleotide or oligonucleotide segments may be attached directly to respective beads in the separate populations or may be indirectly attached (e.g., via an anchor component coupled to the beads, as described above) to respective beads in the separate populations.

[00203] In some cases, the first oligonucleotide may be attached to the separate populations with the aid of a splint (an example of a splint is shown as **2306** in **Fig. 23A**). A splint, as used herein, generally refers to a double-stranded nucleic acid, where one strand of the nucleic acid comprises an oligonucleotide to-be-attached to one or more receiving oligonucleotides and where the other strand of the nucleic acid comprises an oligonucleotide with a sequence that is in part complementary to at least a portion of the oligonucleotide to-be-attached and in part complementary to at least a portion of the one or more receiving oligonucleotides. In some cases, an oligonucleotide may be in part complementary to at least a portion of a receiving oligonucleotide via an overhang sequence as shown in **Fig 23A**). An overhang sequence can be of any suitable length, as described elsewhere herein.

[00204] For example, a splint may be configured such that it comprises the first oligonucleotide or oligonucleotide segment hybridized to an oligonucleotide that comprises a sequence that is in part complementary to at least a portion of the first oligonucleotide or oligonucleotide segment and a sequence (e.g., an overhang sequence) that is in part complementary to at least a portion of an oligonucleotide attached to the separate populations. The splint can hybridize to the oligonucleotide attached to the separate populations via its complementary sequence. Once hybridized, the first oligonucleotide or oligonucleotide segment of the splint can then be attached to the oligonucleotide attached to the separate populations via any suitable attachment mechanism, such as, for example, a ligation reaction.

[00205] Following attachment of the first oligonucleotide or oligonucleotide segment to the separate populations, the separate populations are then pooled to create a mixed pooled population, which is then separated into a plurality of separate populations of the mixed, pooled population. A second oligonucleotide or segment including a second barcode sequence segment is then attached to the first oligonucleotides on the beads in each separate mixed, pooled population, such that different mixed pooled bead populations have a different second barcode sequence segment attached to it. Each bead in the separate populations of the mixed, pooled population may be attached to at least 2, 10, 50, 100, 500, 1000, 5000, 10000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, or more second oligonucleotide molecules or oligonucleotide segment molecules.

[00206] In some cases, the second oligonucleotide may be attached to the first oligonucleotide with the aid of a splint. For example, the splint used to attach the first oligonucleotide or oligonucleotide segment

to the separate populations prior to generating the mixed pooled population may also comprise a sequence (e.g., an overhang sequence) that is in part complementary to at least a portion of the second oligonucleotide. The splint can hybridize to the second oligonucleotide via the complementary sequence. Once hybridized, the second oligonucleotide can then be attached to the first oligonucleotide via any suitable attachment mechanism, such as, for example, a ligation reaction. The splint strand complementary to both the first and second oligonucleotides can then be then denatured (or removed) with further processing. Alternatively, a separate splint comprising the second oligonucleotide may be provided to attach the second oligonucleotide to the first oligonucleotide in analogous fashion as described above for attaching the first oligonucleotide to an oligonucleotide attached to the separate populations with the aid of splint. Also, in some cases, the first barcode segment of the first oligonucleotide and second barcode segment of the second oligonucleotide may be joined via a linking sequence as described elsewhere herein.

[00207] The separate populations of the mixed, pooled population can then be pooled and the resulting pooled bead population then includes a diverse population of barcode sequences, or barcode library that is represented by the product of the number of different first barcode sequences and the number of different second barcode sequences. For example, where the first and second oligonucleotides include, e.g., all 256 4-mer barcode sequence segments, a complete barcode library may include 65,536 diverse 8 base barcode sequences.

[00208] The barcode sequence segments may be independently selected from a set of barcode sequence segments or the first and second barcode sequence segments may each be selected from separate sets of barcode sequence segments. Moreover, the barcode sequence segments may individually and independently comprise from 2 to 20 nucleotides in length, preferably from about 4 to about 20 nucleotides in length, more preferably from about 4 to about 16 nucleotides in length or from about 4 to about 10 nucleotides in length. In some cases, the barcode sequence segments may individually and independently comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more nucleotides in length. In particular, the barcode sequence segments may comprise 2-mers, 3-mers, 4-mers, 5-mers, 6-mers, 7-mers, 8-mers, 9-mers, 10-mers, 11-mers, 12-mers, 13-mers, 14-mers, 15-mers, 16-mers, 17-mers, 18-mers, 19-mers, 20-mers, or longer sequence segments.

[00209] Furthermore, the barcode sequence segments included within the first and second oligonucleotide sequences or sequence segments will typically represent at least 10 different barcode sequence segments, at least 50 different barcode sequence segments, at least 100 different barcode sequence segments, at least 500 different barcode sequence segments, at least 1,000 different barcode sequence segments, at least about 2,000 different barcode sequence segments, at least about 4,000 different barcode sequence segments, at least about 5,000 different barcode sequence segments, at least about 10,000 different barcode sequence segments, at least 50,000 different barcode sequence segments, at least 100,000 barcode sequence segments, at least 500,000 barcode sequence segments, at least 1,000,000 barcode sequence segments, or more. In accordance with the processes described above, these different oligonucleotides may be allocated amongst a similar or the same number of separate bead populations in

either the first or second oligonucleotide addition step, e.g., at least 10, 100, 500, 1000, 2000, 4000, 5000, 10000, 50000, 100000, 500000, 1000000, etc., different barcode sequence segments being separately added to at least 10, 100, 500, 1000, 2000, 4000, 5000, 10000, 50000, 100000, 500000, 1000000, etc., separate bead populations.

[00210] As a result, resulting barcode libraries may range in diversity of from at least about 100 different barcode sequence segments to at least about 1,000,000, 2,000,000, 5,000,000, 10,000,000 100,000,000 or more different barcode sequence segments as described elsewhere herein, being represented within the library.

[00211] As noted previously, either or both of the first and second oligonucleotide sequences or sequence segments, or subsequently added oligonucleotides (e.g., addition of a third oligonucleotide to the second oligonucleotide, addition of a fourth oligonucleotide to an added third oligonucleotide, etc.), may include additional sequences, e.g., complete or partial functional sequences (e.g., a primer sequence (e.g., a universal primer sequence, a targeted primer sequence, a random primer sequence), a primer annealing sequence, an attachment sequence, a sequencing primer sequence, a random N-mer, etc.), for use in subsequent processing. These sequences will, in many cases, be common among beads in the separate populations, subsets of populations, and/or common among all beads in the overall population. In some cases, the functional sequences may be variable as between different bead subpopulations, different beads, or even different molecules attached to a single bead. Moreover, either or both of the first and second oligonucleotide sequences or sequence segments may comprise a sequence segment that includes one or more of a uracil containing nucleotide and a non-native nucleotide, as described elsewhere herein. In addition, although described as oligonucleotides comprising barcode sequences, it will be appreciated that such references includes oligonucleotides that are comprised of two, three or more discrete barcode sequence segments that are separated by one or more bases within the oligonucleotide, e.g., a first barcode segment separated from a second barcode segment by 1, 2, 3, 4, 5, 6, or 10 or more bases in the oligonucleotide in which they are contained. Preferably, barcode sequence segments will be located adjacent to each other or within 6 bases, 4 bases, 3 bases or two bases of each other in the oligonucleotide sequence in which they are contained. Together, whether contiguous within an oligonucleotide sequence, or separated by one or more bases, such collective barcode sequence segments within a given oligonucleotide are referred to herein as a barcode sequence, barcode sequence segment, or barcode domain.

[00212] An example combinatorial method for generating beads with sequences comprising barcode sequences as well as specific types of functional sequences is shown in **Fig 23**. Although described in terms of certain specific sequence segments for purposes of illustration, it will be appreciated that a variety of different configurations may be incorporated into the barcode containing oligonucleotides attached to the beads described herein, including a variety of different functional sequence types, primer types, e.g., specific for different sequencing systems, and the like. As shown in **Fig 23A**, beads **2301** may be generated and covalently linked (e.g., via an acrydite moiety or other species) to a first oligonucleotide component to be used as an anchoring component and/or functional sequence or partial functional

sequence, e.g., partial P5 sequence **2302**. In each well of a plate (e.g., a 384-well plate) an oligonucleotide **2303**, comprising the remaining P5 sequence and a unique first partial barcode sequence (indicated by bases "DDDDDD" in oligonucleotide **2303**), can be hybridized to an oligonucleotide **2304** that comprises the complement of oligonucleotide **2303** and additional bases that overhang each end of oligonucleotide **2303**. Hybridized product (a "splint") **2306** can thus be generated. Each overhang of the splint can be blocked (indicated with an "X" in **Fig 23A**) with a blocking moiety to prevent side product formation. Non-limiting examples of blocking moieties include 3' Inverted dT, dideoxycytidine (ddC), and 3'C3 Spacer. Accordingly, in the example described, different splints can be generated, each with a unique first partial barcode sequence or its complement, e.g., 384 different splints, as described.

[00213] As shown in **Fig 23B,** beads **2301** can be added to each well of the plate and the splint **2306** in each well can hybridize with the corresponding anchor sequence, e.g., partial P5 sequence **2302,** of beads **2301**, via one of the overhangs of oligonucleotide **2304**. Limited stability of the overhang of oligonucleotide **2304** in hybridizing partial P5 sequence **2302** can permit dynamic sampling of splint **2306**, which can aid in ensuring that subsequent ligation of oligonucleotide **2303** to partial P5 sequence **2302** is efficient. A ligation enzyme (e.g., a ligase) can ligate partial P5 sequence **2302** to oligonucleotide **2303**. An example of a ligase would be T4 DNA ligase. Following ligation, the products can be pooled and the beads washed to remove unligated oligonucleotides.

[00214] As shown in **Fig 23C**, the washed products can then be redistributed into wells of another plate (e.g., a 384-well plate), with each well of the plate comprising an oligonucleotide **2305** that has a unique second partial barcode sequence (indicated by "DDDDDD" in oligonucleotide **2305**) and an adjacent short sequence (e.g., "CC" adjacent to the second partial barcode sequence and at the terminus of oligonucleotide **2305**) complementary to the remaining overhang of oligonucleotide **2304**. Oligonucleotide **2305** can also comprise additional sequences, such as R1 sequences and a random N-mer (indicated by "NNNNNNNNNN" in oligonucleotide **2305**). In some cases, oligonucleotide **2305** may comprise a uracil containing nucleotide. In some cases, any of the thymine containing nucleotides of oligonucleotide **2305** may be substituted with uracil containing nucleotides. In some cases, in order to improve the efficiency of ligation of the oligonucleotide comprising the second partial barcode sequence, e.g., sequence **2305**, to the first partial barcode sequence, e.g., sequence **2303**, a duplex strand, e.g., that is complementary to all or a portion of oligonucleotide **2305**, may be provided hybridized to some portion or all of oligonucleotide **2305**, while leaving the overhang bases available for hybridization to splint **2304**. As noted previously, splint **2304** and/or the duplex strand, may be provided blocked at one or both of their 3' and 5' ends to prevent formation of side products from or between one or both of the splint and the duplex strand. In preferred aspects, the duplex strand may be complementary to all or a portion of oligonucleotide **2305**. For example, where oligonucleotide **2305** includes a random n-mer, the duplex strand may be provided that does not hybridize to that portion of the oligonucleotide.

[00215] Via the adjacent short sequence, oligonucleotide **2305** can be hybridized with oligonucleotide **2304**, as shown in **Fig 23C**. Again, the limited stability of the overhang in hybridizing the short complementary sequence of oligonucleotide **2305** can permit dynamic sampling of oligonucleotide **2305**,

which can aid in ensuring that subsequent ligation of oligonucleotide **2305** to oligonucleotide **2303** is efficient. A ligation enzyme (e.g., a ligase) can then ligate oligonucleotide **2305** to oligonucleotide **2303**. Ligation of oligonucleotide **2305** to oligonucleotide **2303** can result in the generation of a full barcode sequence, via the joining of the first partial barcode sequence of oligonucleotide **2305** and the second partial barcode sequence of oligonucleotide **2303**. As shown in **Fig 23D**, the products can then be pooled, the oligonucleotide **2304** can be denatured from the products, and unbound oligonucleotides can then be washed away. Following washing, a diverse library of barcoded beads can be obtained, with each bead bound to, for example, an oligonucleotide comprising a P5 sequence, a full barcode sequence, an R1 sequence, and a random N-mer. In this example, 147, 456 unique barcode sequences can be obtained (e.g., 384 unique first partial barcode sequence x 384 unique second partial barcode sequences).

[00216] In some cases, the inclusion of overhang bases that aid in ligation of oligonucleotides as described above can result in products that all have the same base at a given position, including in between portions of a barcode sequence as shown in **Fig 24A**. Limited or no base diversity at a given sequence position across sequencing reads may result in failed sequencing runs, depending upon the particular sequencing method utilized. Accordingly, in a number of aspects, the overhang bases may be provided with some variability as between different splints, either in terms of base identity or position within the overall sequenced portion of the oligonucleotide. For example, in a first example, one or more spacer bases **2401** (e.g., "1" "2" in **Fig 24B** at **2401**) can be added to some oligonucleotides used to synthesize larger oligonucleotides on beads, such that oligonucleotide products differ slightly in length from one another, and thus position the overhang bases at different locations in different sequences. Complementary spacer bases may also be added to splints necessary for sequence component ligations. A slight difference in oligonucleotide length between products can result in base diversity at a given read position, as shown in **Fig 24B**.

[00217] In another example shown in **Fig 25**, splints comprising a random base overhang may be used to introduce base diversity at read positions complementary to splint overhangs. For example, a double-stranded splint **2501** may comprise a random base (e.g., "NN" in **Fig 25A**) overhang **2503** and a determined base (e.g., "CTCT" in **Fig 25A**) overhang **2506** on one strand and a first partial barcode sequence (e.g., "DDDDDD" in **Fig 25A**) on the other strand. Using an analogous ligation scheme as described above for the Example depicted in **Fig 23**, the determined overhang **2506** may be used to capture sequence **2502** (which may be attached to a bead as shown in **Fig 23**) via hybridization for subsequent ligation with the upper strand (as shown in **Fig 25A**) of splint **2501**. Although overhang **2506** is illustrated as a four base determined sequence overhang, it will be appreciated that this sequence may be longer in order to improve the efficiency of hybridization and ligation in the first ligation step. As such determined base overhang **2506** may include 4, 6, 8, 10 or more bases in length that are complementary to partial P5 sequence **2502**. Moreover, the random base overhang **2503** may be used to capture the remaining component (e.g., sequence **2504**) of the final desired sequence. Sequence **2504** may comprise a second partial barcode sequence ("DDDDDD" in sequence **2504** of **Fig 25C**), the

complement **2505** (e.g., "NN" at **2505** in **Fig 25C**) of the random base overhang **2503** at one end and a random N-mer **2507** at its other end (e.g., "NNNNNNNNNN" in sequence **2504** of **Fig 25C**).

[00218] Due to the randomness of the bases in random base overhang **2503**, bases incorporated into the ligation product at complement **2505** can vary, such that products comprise a variety of bases at the read positions of complement **2505**. As will be appreciated, in preferred aspects, the second partial barcode sequence portion to be ligated to the first partial barcode sequence will typically include a population of such second partial barcode sequences that includes all of the complements to the random overhang sequences, e.g., a given partial barcode sequence will be present with, e.g., 16 different overhang portions, in order to add the same second partial barcode sequence to each bead in a given well where multiple overhang sequences are represented. While only two bases are shown for random overhang **2503** and complement **2505** in **Fig 25**, the example is not meant to be limiting. Any suitable number of random bases in an overhang may be used. Further, while described as random overhang sequences, in some cases, these overhang sequences may be selected from a subset of overhang sequences. For example, in some cases, the overhangs will be selected from subsets of overhang sequences that include fewer than all possible overhang sequences of the length of the overhang, which may be more than one overhang sequence, and in some cases, more than 2, more than 4, more than 10, more than 20, more than 50, or even more overhang sequences.

[00219] In another example, a set of splints, each with a defined overhang selected from a set of overhang sequences of a given length, e.g., a set of at least 2, 4, 10, 20 or more overhang sequences may be used to introduce base diversity at read positions complementary to splint overhangs. Again, because these overhangs are used to ligate a second partial barcode sequence to the first barcode sequence, it will be desirable to have all possible overhang complements represented in the population of second partial barcode sequences. As such, in many cases, it will be preferred to keep the numbers of different overhang sequences lower, e.g., less than 50, less than 20, or in some cases, less than 10 or less than 5 different overhang sequences. In many cases, the number of different linking sequences in a barcode library will be between 2 and 4096 different linking sequences, with preferred libraries having between about 2 and about 50 different linking sequences. Likewise it will typically be desirable to keep these overhang sequences of a relatively short length, in order to avoid introducing non-relevant bases to the ultimate sequence reads. As such, these overhang sequences will typically be designed to introduce no more than 10, no more than 9, no more than 8, no more than 7, no more than 6, no more than 5, no more than 4, and in some cases, 3 or fewer nucleotides to the overall oligonucleotide construct. In some cases, the length of an overhang sequence may be from about 1 to about 10 nucleotides in length, from about 2 to about 8 nucleotides in length, from about 2 to about 6 nucleotides in length, or from about 2 to about 4 nucleotides in length. In general, each splint in the set can comprise an overhang with a different sequence from other splints in the set, such that the base at each position of the overhang is different from the base in the same base position in the other splints in the set. An example set of splints is depicted in **Fig 26**. The set comprises splint **2601** (comprising an overhang of "AC" **2602**), splint **2603** (comprising an overhang of "CT" **2604**), splint **2605** (comprising an overhang of "GA" **2606**), and splint **2607**

(comprising an overhang of "TG" **2608**). Each splint can also comprise an overhang **2609** (e.g., "CTCT" in each splint) and first partial barcode sequence ("DDDDDD"). As shown in **Fig 26**, each splint can comprise a different base in each position of its unique overhang (e.g., overhang **2602** in splint **2601**, overhang **2604** in splint **2603**, overhang **2606** in splint **2605**, and overhang **2608** in splint **2607**) such that no splint overhang comprises the same base in the same base position. Because each splint comprises a different base in each position of its unique overhang, products generated from each splint can also have a different base in each complementary position when compared to products generated from one of the other splints. Thus, base diversity at these positions can be achieved.

[00220] Such products can be generated by hybridizing the first component of the desired sequence (e.g., sequence **2502** in **Fig 25** comprising a first partial barcode sequence; the first component may also be attached to a bead) with the overhang common to each splint (e.g., overhang **2609** in **Fig 26**); ligating the first component of the sequence to the splint; hybridizing the second part of the desired sequence (e.g., a sequence similar to sequence **2504** in **Fig 25** comprising a second partial barcode sequence, except that the sequence comprises bases complementary to the unique overhang sequence at positions **2505** instead of random bases) to the unique overhang of the splint; and ligating the second component of the desired sequence to the splint. The unligated portion of the splint (e.g., bottom sequence comprising the overhangs as shown in **Fig 26**) can then be denatured, the products washed, etc. as described previously to obtain final products. As will be appreciated, and as noted previously, these overhang sequences may provide 1, 2, 3, 4, 5 or 6 or more bases between different partial barcode sequences (or barcode sequence segments), such that they provide a linking sequence between barcode sequence segments, with the characteristics described above. Such a linking sequence may be of varied length, such as for example, from about 2 to about 10 nucleotides in length, from about 2 to about 8 nucleotides in length, from about 2 to about 6 nucleotides in length, from about 2 to about 5 nucleotides in length, or from about 2 to about 4 nucleotides in length.

[00221] An example workflow using the set of splints depicted in **Fig 26** is shown in **Fig 27**. For each splint in the set, the splint strand comprising the unique overhang sequence (e.g., the bottom strand of splints shown in **Fig 26**) can be provided in each well of one or more plates. In **Fig 27**, two 96-well plates of splint strands comprising a unique overhang sequence are provided for each of the four splint types, for a total of eight plates. Of the eight plates, two plates (**2601a**, **2601b**) correspond to the bottom strand of splint **2601** comprising a unique overhang sequence ("AC") in **Fig 26,** two plates (**2603a, 2603b**) correspond to the bottom strand of splint **2603** in **Fig 26** comprising a unique overhang sequence ("CT"), two plates (**2605a**, **2605b**) correspond to the bottom strand of splint **2605** in **Fig 26** comprising a unique overhang sequence ("GA"), and two plates (**2607a, 2607b**) correspond to the bottom strand of splint **2607** in **Fig 26** comprising a unique overhang sequence ("TG"). The oligonucleotides in each 96-well plate (**2601a**, **2601b**, **2603a**, **2603b**, **2605a**, **2605b**, **2607a**, and **2607b**) can be transferred to another set of 96-well plates **2702**, with each plate transferred to its own separate plate (again, for a total of eight plates), and each well of each plate transferred to its corresponding well in the next plate.

[00222] The splint strand comprising a unique first partial barcode sequence (e.g., the upper strand of splints shown in **Fig 26**) and a first partial P5 sequence can be provided in one or more plates. In **Fig 27**, such splint strands are provided in two 96-well plates **2708a** and **2708b**, with each well of the two plates comprising an oligonucleotide with a unique first partial barcode sequence, for a total of 192 unique first partial barcode sequences across the two plates. Each well of plate **2708a** can be added to its corresponding well in four of the plates **2702** and each well of plate **2708b** can be added to its corresponding well in the other four of the plates **2702**. Thus, the two splint strands in each well can hybridize to generate a complete splint. After splint generation, each well of two of the 96-well plates **2702** in **Fig 27** comprises a splint configured as splint **2601,** splint **2603**, splint **2605**, or splint **2607** in **Fig 26** and a unique first partial barcode sequence, for a total of 192 unique first partial barcode sequences.

[00223] To each of the wells of the plates **2702**, beads **2709** comprising a second partial P5 sequence (e.g., similar or equivalent to sequence **2502** in **Fig 25**) can then be added. The splints in each well can hybridize with the second partial P5 sequence via the common overhang sequence **2609** of each splint. A ligation enzyme (e.g., a ligase) can then ligate the second partial P5 sequence to the splint strand comprising the remaining first partial P5 sequence and the first partial barcode sequence. First products are, thus, generated comprising beads linked to a sequence comprising a P5 sequence and a first partial barcode sequence, still hybridized with the splint strand comprising the overhang sequences. Following ligation, first products from the wells of each plate can be separately pooled to generate plate pools **2703**. The plate pools **2703** corresponding to each two-plate set (e.g., each set corresponding to a particular splint configuration) can also be separately pooled to generate first product pools **2704**, such that each first product pool **2704** comprises products generated from splints comprising only one unique overhang sequence. In **Fig 27**, four first product pools **2704** are generated, each corresponding to one of the four splint types used in the example. The products in each plate pool **2703** may be washed to remove unbound oligonucleotides, the products in each first product pool **2704** may be washed to remove unbound oligonucleotides, or washing may occur at both pooling steps. In some cases, plate pooling **2703** may be bypassed with the contents of each two-plate set entered directly into a first product pool **2704**.

[00224] Next, each first product pool **2704** can be aliquoted into each well of two 96-well plates **2705**, as depicted in **Fig 27**, for a total of eight plates (e.g., two plates per product pool **2704**). Separately, oligonucleotides that comprise a unique second partial barcode sequence, a terminal sequence complementary to one of the four unique overhang sequences, and any other sequence to be added (e.g., additional sequencing primer sites, random N-mers, etc.) can be provided in 96-well plates **2706**. Such oligonucleotides may, for example, comprise a sequence similar to sequence **2504** in **Fig 25**, except that the sequence comprises bases complementary to a unique overhang sequence at position **2505** instead of random bases. For example, for splint **2601** shown in **Fig 26**, the bases in position **2505** would be "TG", complementary to the unique overhang **2602** ("AC") of splint **2601**. Of the plates **2706**, sets of two plates can each comprise oligonucleotides comprising sequences complementary to one of the four unique overhang sequences, for a total of eight plates and four plate sets as shown in **Fig 27**. Plates **2706** can be

configured such that each well comprises a unique second partial barcode sequence, for a total of 768 unique second partial barcode sequences across the eight plates.

[00225] Each plate of plates **2706** can be paired with a corresponding plate of plates **2705**, based on the appropriate unique overhang sequence of first products entered into the plate of plates **2705**, as shown in **Fig 27**. Oligonucleotides in each well of the plate from plates **2706** can be added to its corresponding well in its corresponding plate from plates **2705**, such that each well comprises an aliquot of first products from the appropriate first product pool **2704** and oligonucleotides comprising a unique second barcode sequence and any other sequence (e.g., random N-mers) from plates **2706**. In each well of the plates **2705**, the unique overhang sequence of each first product can hybridize with an oligonucleotide comprising the second partial barcode sequence, via the oligonucleotide's bases complementary to the unique overhang sequence. A ligation enzyme (e.g., a ligase) can then ligate the oligonucleotides to the first products. Upon ligation, second products comprising complete barcode sequences are generated via joining of the first partial barcode sequence of the first products with the second partial barcode sequence of the second products. The second products obtained from plates **2705** can be removed and deposited into a common second product pool **2707**. The splint strands comprising the overhangs (as shown in **Fig 26**) can then be denatured in product pool **2707**, and the products washed to obtain final products. A total of 147,456 unique barcode sequences can be obtained (e.g., 192 first partial barcode sequences x 768 second partial barcode sequences) with base diversity in base positions complementary to unique overhang sequences used during ligations.

[00226] The above example with respect to splint sets is not meant to be limiting, nor is the number and type

(s) of plates used for combinatorial synthesis. A set of splints can comprise any suitable number of splints. Moreover, each set of splints may be designed with the appropriate first partial barcode sequence diversity depending upon, for example, the number of unique barcode sequences desired, the number of bases used to generate a barcode sequence, etc.

[00227] Using a combinatorial plate method, libraries of barcoded beads with high-diversity can be generated. For example, if two 384-well plates are used, each with oligonucleotides comprising partial barcode sequences pre-deposited in each well, it is possible that 384 x 384 or 147,456 unique barcode sequences can be generated. The combinatorial examples shown herein are not meant to be limiting as any suitable combination of plates may be used. For example, while in some cases, the barcode sequence segments added in each combinatorial step may be selected from the same sets of barcode sequence segments. However, in many cases, the barcode sequence segments added in each combinatorial step may be selected from partially or completely different sets of oligonucleotide sequences. For example, in some cases, a first oligonucleotide segment may include a barcode sequence from a first set of barcode sequences, e.g., 4-mer sequences, while the second oligonucleotide sequence may include barcode sequences from a partially or completely different set of barcode sequence segments, e.g., 4-mer sequences, 6-mer sequences, 8-mer sequences, etc., or even sequences of mixed lengths, e.g., where the second oligonucleotide segment is selected form a set of oligonucleotides having barcode sequences

having varied lengths and sequences, to generate multiparameter variability in the generated barcodes, e.g., sequence and length.,

[00228] With reference to the example above, for example, the number and type of plates (and barcodes) used for each step in a combinatorial method does not have to be the same. For example, a 384 well plate may be used for a first step and a 96 well plate may be used for a second step for a total of 36,864 unique barcode sequences generated. Furthermore, the number of bases of a full barcode sequence added in each combinatorial step does not need to be the same. For example, in a first combinatorial step, 4 bases of a 12 base barcode sequence may be added, with the remaining 8 bases added in a second combinatorial step. Moreover, the number of combinatorial steps used to generate a full barcode sequence may also vary. In some cases, about 2, 3, 4, 5, 6, 7, 8, 9, or 10 combinatorial steps are used.

[00229] The primer extension reactions and ligation reactions can be conducted with standard techniques and reagents in the multiwell plates. For example, the polymer, poly-ethylene glycol (PEG), may be present during the single-stranded ligation reaction at a concentration of about 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, or 75%. In some cases, the PEG may be present during the ligation reaction at a concentration of more than about 6%, 10%, 18%, 20%, 30%, 36%, 40%, 50% or more. In some cases, the PEG may be present during the ligation reaction in the second plate at a concentration of less than about 6%, 10%, 18%, 20%, 30%, 36%, 40%, or 50%.

[00230] The methods provided herein may reduce nucleotide bias in ligation reactions. Better results may occur when the first extension in the first well plate may be run to completion. For the single-strand ligation step in the second well plate, no competition may be present when only one type of oligonucleotide sequence is used. The partitioning in wells method for attaching content to beads may avoid misformed adaptors with 8N ends, particularly when the first extension in the first well plate is run to completion.

[00231] Potential modifications to the partitioning in wells process may include replacing the single-strand ligation step with PCR by providing the second oligonucleotide sequence with degenerate bases, modifying the first oligonucleotide sequence to be longer than the second oligonucleotide sequence, and/or adding a random N-mer sequence in a separate bulk reaction after the single-strand ligation step, as this may save synthesis costs and may reduce N-mer sequence bias.

[00232] In some cases, the following sequence of processes may be used to attach a barcode sequence to a bead. The barcode sequence may be mixed with suitable PCR reagents and a plurality of beads in aqueous fluid. The aqueous fluid may be emulsified within an immiscible fluid, such as an oil, to form an emulsion. The emulsion may generate individual fluidic droplets containing the barcode sequence, the bead, and PCR reagents. Individual fluidic droplets may be exposed to thermocycling conditions, in which the multiple rounds of temperature cycling permits priming and extension of barcode sequences. The emulsion containing the fluidic droplets may be broken by continuous phase exchange, described elsewhere in this disclosure. Resulting barcoded beads suspended in aqueous solution may be sorted by magnetic separation or other sorting methods to obtain a collection of purified barcoded beads in aqueous fluid.

[00233] In some cases, the following sequence of processes may be used to attach an N-mer sequence to a bead. The N-mer sequence may be mixed with suitable PCR reagents and a plurality of pooled barcoded beads in aqueous fluid. The aqueous fluid may be heated to permit hybridization and extension of the N-mer sequence. Additional heating may permit removal of the complement strand.

[00234] The PCR reagents may include any suitable PCR reagents. In some cases, dUTPs may be substituted for dTTPs during the primer extension or other amplification reactions, such that oligonucleotide products comprise uracil containing nucleotides rather than thymine containing nucleotides. This uracil-containing section of the universal sequence may later be used together with a polymerase that will not accept or process uracil-containing templates to mitigate undesired amplification products.

[00235] Amplification reagents may include a universal primer, universal primer binding site, sequencing primer, sequencing primer binding site, universal read primer, universal read binding site, or other primers compatible with a sequencing device, *e.g.*, an Illumina sequencer, Ion Torrent sequencer, etc. The amplification reagents may include P5, non cleavable 5'acrydite-P5, a cleavable 5' acrydite-SS-P5, R1c, Biotin R1c, sequencing primer, read primer, P5_Universal, P5_U, 52-BioR1-rc, a random N-mer sequence, a universal read primer, etc. In some cases, a primer may contain a modified nucleotide, a locked nucleic acid (LNA), an LNA nucleotide, a uracil containing nucleotide, a nucleotide containing a non-native base, a blocker oligonucleotide, a blocked 3' end, 3'ddCTP. **Fig 19** provides additional examples.

[00236] As described herein, in some cases oligonucleotides comprising barcodes are partitioned such that each bead is partitioned with, on average, less than one unique oligonucleotide sequence, less than two unique oligonucleotide sequences, less than three unique oligonucleotide sequences, less than four unique oligonucleotide sequences, less than five unique oligonucleotide sequences, or less than ten unique oligonucleotide sequences. Therefore, in some cases, a fraction of the beads does not contain an oligonucleotide template and therefore cannot contain an amplified oligonucleotide. Thus, it may be desirable to separate beads comprising oligonucleotides from beads not comprising oligonucleotides. In some cases, this may be done using a capture moiety.

[00237] In some embodiments, a capture moiety may be used with isolation methods such as magnetic separation to separate beads containing barcodes from beads, which may not contain barcodes. As such, in some cases, the amplification reagents may include capture moieties attached to a primer or probe. Capture moieties may allow for sorting of labeled beads from non-labeled beads to confirm attachment of primers and downstream amplification products to a bead. Exemplary capture moieties include biotin, streptavidin, glutathione-S-transferase (GST), cMyc, HA, etc. The capture moieties may be, or include, a fluorescent label or magnetic label. The capture moiety may comprise multiple molecules of a capture moiety, *e.g.*, multiple molecules of biotin, streptavidin, etc. In some cases, an amplification reaction may make use of capture primers attached to a capture moiety (as described elsewhere herein), such that the primer hybridizes with amplification products and the capture moiety is integrated into additional amplified oligonucleotides during additional cycles of the amplification reaction. In other cases, a probe

comprising a capture moiety may be hybridized to amplified oligonucleotides following the completion of an amplification reaction such that the capture moiety is associated with the amplified oligonucleotides.

[00238] A capture moiety may be a member of binding pair, such that the capture moiety can be bound with its binding pair during separation. For example, beads may be generated that comprise oligonucleotides that comprise a capture moiety that is a member of a binding pair (e.g., biotin). The beads may be mixed with capture beads that comprise the other member of the binding pair (e.g., streptavidin), such that the two binding pair members bind in the resulting mixture. The bead-capture bead complexes may then be separated from other components of the mixture using any suitable means, including, for example centrifugation and magnetic separation (e.g., including cases where the capture bead is a magnetic bead).

[00239] In many cases as described, individual beads will generally have oligonucleotides attached thereto, that have a common overall barcode sequence segment. As described herein, where a bead includes oligonucleotides having a common barcode sequence, it is generally meant that of the oligonucleotides coupled to a given bead, a significant percentage, e.g., greater than 70%, greater than 80%, greater than 90%, greater than 95% or even greater than 99% of the oligonucleotides of or greater than a given length, e.g., including the full expected length or lengths of final oligonucleotides and excluding unreacted anchor sequences or partial barcode sequences, include the same or identical barcode sequence segments. This barcode sequence segment or domain (again, which may be comprised of two or more sequence segments separated by one or more bases) may be included among other common or variable sequences or domains within a single bead. Also as described, the overall population of beads will include beads having large numbers of different barcode sequence segments. In many cases, however, a number of separate beads within a given bead population may include the same barcode sequence segment. In particular, a barcode sequence library having 1000, 10,000, 1,000,000, 10,000,000 or more different sequences, may be represented in bead populations of greater than 100,000, 1,000,000, 10,000,000, 100,000,000, 1 billion, 10 billion, 100 billion or more discrete beads, such that the same barcode sequence is represented multiple times within a given bead population or subpopulation. For example, the same barcode sequence may be present on two or more beads within a given analysis, 10 or more beads, 100 or more beads, etc..

[00240] A capture device, such as a magnetic bead, with a corresponding linkage, such as streptavidin, may be added to bind the capture moiety, for example, biotin. The attached magnetic bead may then enable isolation of the barcoded beads by, for example, magnetic sorting. Magnetic beads may also be coated with other linking entities besides streptavidin, including nickel-IMAC to enable the separation of His-tagged fusion proteins, coated with titanium dioxide to enable the separation of phosphorylated peptides, or coated with amine-reactive NHS-ester groups to immobilize protein or other ligands for separation.

[00241] In some embodiments, the capture moiety may be attached to a primer, to an internal sequence, to a specific sequence within the amplified product, to a barcode sequence, to a universal sequence, or to a complementary sequence. Capture moieties may be attached by PCR amplification or ligation. Capture

moieties may include a universal tag such as biotin attached to a specific target such as a primer before added to the bead population. In other cases, capture moieties may include a specific tag that recognizes a specific sequence or protein or antibody that may be added to the bead population independently. In some embodiments, the capture moieties may be pre-linked to a sorting bead, such as a magnetic bead. In some cases, the capture moiety may be a fluorescent label, which may enable sorting via fluorescence-activated cell sorting (FACS).

[00242] In some cases, a nucleic acid label (*e.g.*, fluorescent label) may be used to identify fluidic droplets, emulsions, or beads that contain oligonucleotides. Sorting (e.g., via flow cytometry) of the labeled droplets or beads may then be performed in order to isolate beads attached to amplified oligonucleotides. Exemplary stains include intercalating dyes, minor-groove binders, major groove binders, external binders, and bis-intercalators. Specific examples of such dyes include SYBR green, SYBR blue, DAPI, propidium iodide, SYBR gold, ethidium bromide, propidium iodide, imidazoles (*e.g.*, Hoechst 33258, Hoechst 33342, Hoechst 34580, and DAPI), 7-AAD, SYTOX Blue, SYTOX Green, SYTOX Orange, POPO-1, POPO-3, YOYO-1, YOYO-3, TOTO-1, TOTO-3, JOJO-1, LOLO-1, BOBO-1, BOBO-3, PO-PRO-1, PO-PRO-3, BO-PRO-1, BO-PRO-3, TO-PRO-1, TO-PRO-3, TO-PRO-5, JO-PRO-1, LO-PRO-1, YO-PRO-1, YO-PRO-3, PicoGreen, OliGreen, RiboGreen, EvaGreen, SYBR Green, SYBR Green II, SYBR DX, SYTO-40, -41, -42, -43, -44, -45 (blue), SYTO-13, -16, -24, -21, -23, -12, -11, -20, -22, -15, -14, -25 (green), SYTO-81, -80, -82, -83, -84, -85 (orange), SYTO-64, -17, -59, -61, -62, -60, and -63 (red).

Multi-Functional Beads

[00243] Beads may be linked to a variety of species (including non-nucleic acid species) such that they are multi-functional. For example, a bead may be linked to multiple types of oligonucleotides comprising a barcode sequence and an N-mer (e.g., a random N-mer or a targeted N-mer as described below). Each type of oligonucleotide may differ in its barcode sequence, its N-mer, or any other sequence of the oligonucleotide. Moreover, each bead may be linked to oligonucleotides comprising a barcode sequence and an N-mer and may also be linked to a blocker oligonucleotide capable of blocking the oligonucleotides comprising a barcode sequence and an N-mer. Loading of the oligonucleotide blocker and oligonucleotide comprising a barcode sequence and an N-mer may be completed at distinct ratios in order to obtain desired stoichiometries of oligonucleotide blocker to oligonucleotide comprising a barcode sequence and an N-mer. In general, a plurality of species may be loaded to beads at distinct ratios in order to obtain desired stoichiometries of the species on the beads.

[00244] Moreover, a bead may also be linked to one or more different types of multi-functional oligonucleotides. For example, a multi-functional oligonucleotide may be capable of functioning as two or more of the following: a primer, a tool for ligation, an oligonucleotide blocker, an oligonucleotide capable of hybridization detection, a reporter oligonucleotide, an oligonucleotide probe, a functional oligonucleotide, an enrichment primer, a targeted primer, a non-specific primer, and a fluorescent probe. Oligonucleotides that function as fluorescent probes may be used, for example, for bead detection or

characterization (e.g., quantification of number of beads, quantification of species (e.g., primers, linkers, etc.) attached to beads, determination of bead size/topology, determination of bead porosity, etc.).

[00245] Other non-limiting examples of species that may also be attached or coupled to beads include whole cells, chromosomes, polynucleotides, organic molecules, proteins, polypeptides, carbohydrates, saccharides, sugars, lipids, enzymes, restriction enzymes, ligases, polymerases, barcodes, adapters, small molecules, antibodies, antibody fragments, fluorophores, deoxynucleotide triphosphates (dNTPs), dideoxynucleotide triphosphates (ddNTPs), buffers, acidic solutions, basic solutions, temperature-sensitive enzymes, pH-sensitive enzymes, light-sensitive enzymes, metals, metal ions, magnesium chloride, sodium chloride, manganese, aqueous buffer, mild buffer, ionic buffer, inhibitors, saccharides, oils, salts, ions, detergents, ionic detergents, non-ionic detergents, oligonucleotides, nucleotides, DNA, RNA, peptide polynucleotides, complementary DNA (cDNA), double stranded DNA (dsDNA), single stranded DNA (ssDNA), plasmid DNA, cosmid DNA, chromosomal DNA, genomic DNA, viral DNA, bacterial DNA, mtDNA (mitochondrial DNA), mRNA, rRNA, tRNA, nRNA, siRNA, snRNA, snoRNA, scaRNA, microRNA, dsRNA, ribozyme, riboswitch and viral RNA, a locked nucleic acid (LNA) in whole or part, locked nucleic acid nucleotides, any other type of nucleic acid analogue, proteases, nucleases, protease inhibitors, nuclease inhibitors, chelating agents, reducing agents, oxidizing agents, probes, chromophores, dyes, organics, emulsifiers, surfactants, stabilizers, polymers, water, small molecules, pharmaceuticals, radioactive molecules, preservatives, antibiotics, aptamers, and combinations thereof. Both additional oligonucleotide species and other types of species may be coupled to beads by any suitable method including covalent and non-covalent means (e.g., ionic bonds, van der Waals interactions, hydrophobic interactions, encapsulation, diffusion of the species into the bead, etc.). In some cases, an additional species may be a reactant used for a reaction comprising another type of species on the bead. For example, an additional species coupled to a bead may be a reactant suitable for use in an amplification reaction comprising an oligonucleotide species also attached to the bead.

[00246] In some cases, a bead may comprise one or more capture ligands each capable of capturing a particular type of sample component, including components that may comprise nucleic acid. For example, a bead may comprise a capture ligand capable of capturing a cell from a sample. The capture ligand may be, for example, an antibody, antibody fragment, receptor, protein, peptide, small molecule or any other species targeted toward a species unique to and/or over-expressed on the surface of a particular cell. Via interactions with the cell target, the particular cell type can be captured from a sample such that it remains bound to the bead. A bead bound to a cell can be entered into a partition as described elsewhere herein to barcode nucleic acids obtained from the cell. In some cases, capture of a cell from a sample may occur in a partition. Lysis agents, for example, can be included in the partition such in order to release the nucleic acid from the cell. The released nucleic acid can be barcoded and processed using any of the methods described herein.

III. Barcode Libraries

[00247] Beads may contain one or more attached barcode sequences. The barcode sequences attached to a single bead may be identical or different. In some cases, each bead may be attached to about 1, 5, 10,

50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 identical barcode sequences. In some cases, each bead may be to about 1, 5, 10, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 different barcode sequences. In some cases, each bead may be attached to at least about 1, 5, 10, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 30000000, 40000000, 50000000, 60000000, 70000000, 80000000, 90000000, 100000000, 200000000, 300000000, 400000000, 500000000, 600000000, 700000000, 800000000, 900000000, 1000000000, 2000000000, 3000000000, 4000000000, 5000000000, 6000000000, 7000000000, 8000000000, 9000000000, 10000000000, 20000000000, 30000000000, 40000000000, 50000000000, 60000000000, 70000000000, 80000000000, 90000000000, 100000000000 or more identical barcode sequences. In some cases, each bead may be attached to at least about 1, 5, 10, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 30000000, 40000000, 50000000, 60000000, 70000000, 80000000, 90000000, 100000000, 200000000, 300000000, 400000000, 500000000, 600000000, 700000000, 800000000, 900000000, 1000000000, 2000000000, 3000000000, 4000000000, 5000000000, 6000000000, 7000000000, 8000000000, 9000000000, 10000000000, 20000000000, 30000000000, 40000000000, 50000000000, 60000000000, 70000000000, 80000000000, 90000000000, 100000000000 or more different barcode sequences. In some cases, each bead may be attached to less than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 identical barcode sequences. In some cases, each bead may be attached to less than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 different barcode sequences.

[00248] An individual barcode library may comprise one or more barcoded beads. In some cases, an individual barcode library may comprise about 1, 5, 10, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 100000000, 500000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 individual barcoded beads. In some cases, each library may comprise at least about 1, 5, 10, 50, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 30000000, 40000000, 50000000, 60000000, 70000000, 80000000, 90000000, 100000000, 200000000, 300000000, 400000000,

500000000, 600000000, 700000000, 800000000, 900000000, 1000000000, 2000000000, 3000000000, 4000000000, 5000000000, 6000000000, 7000000000, 8000000000, 9000000000, 10000000000, 20000000000, 30000000000, 40000000000, 50000000000, 60000000000, 70000000000, 80000000000, 90000000000, 100000000000 or more individual barcoded beads. In some cases, each library may comprise less than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, 1000000000, 5000000000, 10000000000, 50000000000, or 100000000000 individual barcoded beads. The barcoded beads within the library may have the same sequences or different sequences.

[00249] In some embodiments, each bead may have a unique barcode sequence. However, the number of beads with unique barcode sequences within a barcode library may be limited by combinatorial limits. For example, using four different nucleotides, if a barcode is 12 nucleotides in length, than the number of unique constructs may be limited to $4^{12} = 16777216$ unique constructs. Since barcode libraries may comprise many more beads than 1677216, there may be some libraries with multiple copies of the same barcode. In some embodiments, the percentage of multiple copies of the same barcode within a given library may be 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% ,10%, 15%, 20%, 25%, 30%, 40%, or 50%. In some cases, the percentage of multiple copies of the same barcode within a given library may be more than 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% ,10%, 15%, 20%, 25%, 30%, 40%, 50% or more. In some cases, the percentage of multiple copies of the same barcode within a given library may be less than 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% ,10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 25%, 30%, 40%, or 50%.

[00250] In some embodiments, each bead may comprise one unique barcode sequence but multiple different random N-mers. In some cases, each bead may have one or more different random N-mers. Again, the number of beads with different random N-mers within a barcode library may be limited by combinatorial limits. For example, using four different nucleotides, if an N-mer sequence is 12 nucleotides in length, than the number of different constructs may be limited to $4^{12} = 16777216$ different constructs. Since barcode libraries may comprise many more beads than 16777216, there may be some libraries with multiple copies of the same N-mer sequence. In some embodiments, the percentage of multiple copies of the same N-mer sequence within a given library may be 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% ,10%, 15%, 20%, 25%, 30%, 40%, or 50%. In some cases, the percentage of multiple copies of the same N-mer sequence within a given library may be more than 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% ,10%, 15%, 20%, 25%, 30%, 40%, 50% or more. In some cases, the percentage of multiple copies of the same N-mer sequence within a given library may be less than 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% ,10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 25%, 30%, 40%, or 50%.

[00251] In some embodiments, the unique identifier sequence within the barcode may be different for each primer within each bead. In some cases, the unique identifier sequence within the barcode sequence may be the same for each primer within each bead.

IV. Combining Barcoded Beads with Sample

Types of Samples

[00252] The methods, compositions, devices, and kits of this disclosure may be used with any suitable sample or species. A sample (e.g., sample material, component of a sample material, fragment of a sample material, etc.) or species can be, for example, any substance used in sample processing, such as a reagent or an analyte. Exemplary samples can include one or more of whole cells, chromosomes, polynucleotides, organic molecules, proteins, nucleic acids, polypeptides, carbohydrates, saccharides, sugars, lipids, enzymes, restriction enzymes, ligases, polymerases, barcodes (e.g., including barcode sequences, nucleic acid barcode sequences, barcode molecules), adaptors, small molecules, antibodies, fluorophores, deoxynucleotide triphosphate (dNTPs), dideoxynucleotide triphosphates (ddNTPs), buffers, acidic solutions, basic solutions, temperature-sensitive enzymes, pH-sensitive enzymes, light-sensitive enzymes, metals, metal ions, magnesium chloride, sodium chloride, manganese, aqueous buffer, mild buffer, ionic buffer, inhibitors, oils, salts, ions, detergents, ionic detergents, non-ionic detergents, oligonucleotides, template nucleic acid molecules (e.g., template oligonucleotides, template nucleic acid sequences), nucleic acid fragments, template nucleic acid fragments (e.g., fragments of a template nucleic acid generated from fragmenting a template nucleic acid during fragmentation, fragments of a template nucleic acid generated from a nucleic acid amplification reaction), nucleotides, DNA, RNA, peptide polynucleotides, complementary DNA (cDNA), double stranded DNA (dsDNA), single stranded DNA (ssDNA), plasmid DNA, cosmid DNA, chromosomal DNA, genomic DNA (gDNA), viral DNA, bacterial DNA, mtDNA (mitochondrial DNA), mRNA, rRNA, tRNA, nRNA, siRNA, snRNA, snoRNA, scaRNA, microRNA, dsRNA, ribozyme, riboswitch and viral RNA, proteases, locked nucleic acids in whole or part, locked nucleic acid nucleotides, nucleases, protease inhibitors, nuclease inhibitors, chelating agents, reducing agents, oxidizing agents, probes, chromophores, dyes, organics, emulsifiers, surfactants, stabilizers, polymers, water, pharmaceuticals, radioactive molecules, preservatives, antibiotics, aptamers, and the like. In summary, the samples that are used will vary depending on the particular processing needs.

[00253] Samples may be derived from human and non-human sources. In some cases, samples are derived from mammals, non-human mammals, rodents, amphibians, reptiles, dogs, cats, cows, horses, goats, sheep, hens, birds, mice, rabbits, insects, slugs, microbes, bacteria, parasites, or fish. Samples may be derived from a variety of cells, including but not limited to: eukaryotic cells, prokaryotic cells, fungi cells, heart cells, lung cells, kidney cells, liver cells, pancreas cells, reproductive cells, stem cells, induced pluripotent stem cells, gastrointestinal cells, blood cells, cancer cells, bacterial cells, bacterial cells isolated from a human microbiome sample, etc. In some cases, a sample may comprise the contents of a cell, such as, for example, the contents of a single cell or the contents of multiple cells. Examples of single cell applications of the methods and systems described herein are set forth in U.S. Provisional Patent Application No. _____ (Attorney Docket No. 43487-728.101), filed of even date herewith. Samples may also be cell-free, such as circulating nucleic acids (*e.g.*, DNA, RNA).

[00254] A sample may be naturally-occurring or synthetic. A sample may be obtained from any suitable location, including from organisms, whole cells, cell preparations and cell-free compositions from any organism, tissue, cell, or environment. A sample may be obtained from environmental biopsies, aspirates, formalin fixed embedded tissues, air, agricultural samples, soil samples, petroleum samples, water samples, or dust samples. In some instances, a sample may be obtained from bodily fluids, which may include blood, urine, feces, serum, lymph, saliva, mucosal secretions, perspiration, central nervous system fluid, vaginal fluid, or semen. Samples may also be obtained from manufactured products, such as cosmetics, foods, personal care products, and the like. Samples may be the products of experimental manipulation including recombinant cloning, polynucleotide amplification, polymerase chain reaction (PCR) amplification, purification methods (such as purification of genomic DNA or RNA), and synthesis reactions.

Methods of Attaching Barcodes to Samples

[00255] Barcodes (or other oligonucleotides, *e.g.* random N-mers) may be attached to a sample by joining the two nucleic acid segments together through the action of an enzyme. This may be accomplished by primer extension, polymerase chain reaction (PCR), another type of reaction using a polymerase, or by ligation using a ligase. When the ligation method is used to attach a sample to a barcode, the samples may or may not be fragmented prior to the ligation step. In some cases, the oligonucleotides (*e.g.*, barcodes, random N-mers) are attached to a sample while the oligonucleotides are still attached to the beads. In some cases, the oligonucleotides (*e.g.*, barcodes, random N-mers) are attached to a sample after the oligonucleotides are released from the beads, *e.g.*, by cleavage of the oligonucleotides comprising the barcodes from the beads and/or through degradation of the beads.

[00256] The oligonucleotides may include one or more random N-mer sequences. A collection of unique random N-mer sequences may prime random portions of a DNA segment, thereby amplifying a sample (*e.g.*, a whole genome). The resulting product may be a collection of barcoded fragments representative of the entire sample (*e.g.*, genome).

[00257] The samples may or may not be fragmented before ligation to barcoded beads. DNA fragmentation may involve separating or disrupting DNA strands into small pieces or segments. A variety of methods may be employed to fragment DNA including restriction digest or various methods of generating shear forces. Restriction digest may utilize restriction enzymes to make intentional cuts in a DNA sequence by blunt cleavage to both strands or by uneven cleavage to generate sticky ends. Examples of shear-force mediated DNA strand disruption may include sonication, acoustic shearing, needle shearing, pipetting, or nebulization. Sonication, is a type of hydrodynamic shearing, exposing DNA sequences to short periods of shear forces, which may result in about 700 bp fragment sizes. Acoustic shearing applies high-frequency acoustic energy to the DNA sample within a bowl-shaped transducer. Needle shearing generates shear forces by passing DNA through a small diameter needle to physically tear DNA into smaller segments. Nebulization forces may be generated by sending DNA

through a small hole of an aerosol unit in which resulting DNA fragments are collected from the fine mist exiting the unit.

[00258] In some cases, a ligation reaction is used to ligate oligonucleotides to sample. The ligation may involve joining together two nucleic acid segments, such as a barcode sequence and a sample, by catalyzing the formation of a phosphodiester bond. The ligation reaction may include a DNA ligase, such as an *E.coli* DNA ligase, a T4 DNA ligase, a mammalian ligase such as DNA ligase I, DNA ligase III, DNA ligase IV, thermostable ligases, or the like. The T4 DNA ligase may ligate segments containing DNA, oligonucleotides, RNA, and RNA-DNA hybrids. The ligation reaction may not include a DNA ligase, utilizing an alternative such as a topoisomerase. To ligate a sample to a barcode sequence, utilizing a high DNA ligase concentration and including PEG may achieve rapid ligation. The optimal temperature for DNA ligase, which may be 37 °C, and the melting temperature of the DNA to be ligated, which may vary, may be considered to select for a favorable temperature for the ligation reaction. The sample and barcoded beads may be suspended in a buffer to minimize ionic effects that may affect ligation.

[00259] Although described in terms of ligation or direct attachment of a barcode sequence to a sample nucleic acid component, above, the attachment of a barcode to a sample nucleic acid, as used herein, also encompasses the attachment of a barcode sequence to a complement of a sample, or a copy or complement of that complement, e.g., when the barcode is associated with a primer sequence that is used to replicate the sample nucleic acid, as is described in greater detail elsewhere herein. In particular, where a barcode containing primer sequence is used in a primer extension reaction using the sample nucleic acid (or a replicate of the sample nucleic acid) as a template, the resulting extension product, whether a complement of the sample nucleic acid or a duplicate of the sample nucleic acid, will be referred to as having the barcode sequence attached to it.

[00260] In some cases, sample is combined with the barcoded beads (either manually or with the aid of a microfluidic device) and the combined sample and beads are partitioned, such as in a microfluidic device. The partitions may be aqueous droplets within a water-in-oil emulsion. When samples are combined with barcoded beads, on average less than two target analytes may be present in each fluidic droplet. In some embodiments, on average, less than three target analytes may appear per fluidic droplet. In some cases, on average, more than two target analytes may appear per fluidic droplet. In other cases, on average, more than three target analytes may appear per fluidic droplet. In some cases, one or more strands of the same target analyte may appear in the same fluidic droplet. In some cases, less than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 50, 100, 1000, 5000, 10000, or 100000 target analytes are present within a fluidic droplet. In some cases, greater than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 50, 100, 1000, 5000, 10000, or 100000 target analytes are present within a fluidic droplet. The partitions described herein are often characterized by having extremely small volumes. For example, in the case of droplet based partitions, the droplets may have overall volumes that are less than 1000 pL, less than 900 pL, less than 800 pL, less than 700 pL, less than 600 pL, less than 500 pL, less than 400pL, less than 300 pL, less than 200 pL, less than 100pL, less than 50 pL, less than 20 pL, less than 10 pL, or even less than 1 pL. Where co-partitioned with beads, it will be

appreciated that the sample fluid volume within the partitions may be less than 90% of the above described volumes, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, or even less than 10% the above described volumes.

[00261] When samples are combined with barcoded beads, on average less than one bead may be present in each fluidic droplet. In some embodiments, on average, less than two beads may be present in each fluidic droplet. In some embodiments, on average, less than three beads may be present per fluidic droplet. In some cases, on average, more than one bead may be present in each fluidic droplet. In other cases, on average, more than two beads may appear be present in each fluidic droplet. In other cases, on average, more than three beads may be present per fluidic droplet. In some embodiments, a ratio of on average less than one barcoded bead per fluidic droplet may be achieved using limiting dilution technique. Here, barcoded beads may be diluted prior to mixing with the sample, diluted during mixing with the sample, or diluted after mixing with the sample.

[00262] The number of different barcodes or different sets of barcodes (e.g., different sets of barcodes, each different set coupled to a different bead) that are partitioned may vary depending upon, for example, the particular barcodes to be partitioned and/or the application. Different sets of barcodes may be, for example, sets of identical barcodes where the identical barcodes differ between each set. Or different sets of barcodes may be, for example, sets of different barcodes, where each set differs in its included barcodes. In some cases, different barcodes are partitioned by attaching different barcodes to different beads (e.g., gel beads). In some cases, different sets of barcodes are partitioned by disposing each different set in a different partition. In some cases, though a partition may comprise one or more different barcode sets. For example, each different set of barcodes may be coupled to a different bead (e.g., a gel bead). Each different bead may be partitioned into a fluidic droplet, such that each different set of barcodes is partitioned into a different fluidic droplet. For example, about 1, 5, 10, 50, 100, 1000, 10000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, 100000000, or more different barcodes or different sets of barcodes may be partitioned. In some examples, at least about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, 100000000, or more different barcodes or different sets of barcodes may be partitioned. In some examples, less than about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, or 100000000 different barcodes or different sets of barcodes may be partitioned. In some examples, about 1-5, 5-10, 10-50, 50-100, 100-1000, 1000-10000, 10000-100000, 100000-1000000, 10000-1000000, 10000-10000000, or 10000-100000000 different barcodes or different sets of barcodes may be partitioned.

[00263] Barcodes may be partitioned at a particular density. For example, barcodes may be partitioned so

that each partition contains about 1, 5, 10, 50, 100, 1000, 10000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, or 100000000 barcodes per partition. Barcodes may be partitioned so that each partition contains at least about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, 100000000, or more barcodes per partition. Barcodes may be partitioned so that each partition contains less than about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, or 100000000 barcodes per partition. Barcodes may be partitioned such that each partition contains about 1-5, 5-10, 10-50, 50-100, 100-1000, 1000-10000, 10000-100000, 100000-1000000, 10000-1000000, 10000-10000000, or 10000-100000000 barcodes per partition. In some cases, partitioned barcodes may be coupled to one or more beads, such as, for example, a gel bead. In some cases, the partitions are fluidic droplets.

[00264] Barcodes may be partitioned such that identical barcodes are partitioned at a particular density. For example, identical barcodes may be partitioned so that each partition contains about 1, 5, 10, 50, 100, 1000, 10000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, or 100000000 identical barcodes per partition. Barcodes may be partitioned so that each partition contains at least about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, 100000000, or more identical barcodes per partition. Barcodes may be partitioned so that each partition contains less than about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, or 100000000 identical barcodes per partition. Barcodes may be partitioned such that each partition contains about 1-5, 5-10, 10-50, 50-100, 100-1000, 1000-10000, 10000-100000, 100000-1000000, 10000-1000000, 10000-10000000, or 10000-100000000 identical barcodes per partition. In some cases, partitioned identical barcodes may be coupled to a bead, such as, for example, a gel bead. In some cases, the partitions are fluidic droplets.

[00265] Barcodes may be partitioned such that different barcodes are partitioned at a particular density. For example, different barcodes may be partitioned so that each partition contains about 1, 5, 10, 50, 100, 1000, 10000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000,

6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, or 100000000 different barcodes per partition. Barcodes may be partitioned so that each partition contains at least about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, 100000000, or more different barcodes per partition. Barcodes may be partitioned so that each partition contains less than about 1, 5, 10, 50, 100, 1000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200,000, 300,000, 400,000, 500,000, 600,000, 700,000, 800,000, 900,000, 1000000, 2000000, 3000000, 4000000, 5000000, 6000000, 7000000, 8000000, 9000000, 10000000, 20000000, 50000000, or 100000000 different barcodes per partition. Barcodes may be partitioned such that each partition contains about 1-5, 5-10, 10-50, 50-100, 100-1000, 1000-10000, 10000-100000, 100000-1000000, 10000-1000000, 10000-10000000, or 10000-100000000 different barcodes per partition. In some cases, partitioned different barcodes may be coupled to a bead, such as, for example, a gel bead. In some cases, the partitions are fluidic droplets.

[00266] The number of partitions employed to partition barcodes or different sets of barcodes may vary, for example, depending on the application and/or the number of different barcodes or different sets of barcodes to be partitioned. For example, the number of partitions employed to partition barcodes or different sets of barcodes may be about 5, 10, 50, 100, 250, 500, 750, 1000, 1500, 2000, 2500, 5000, 7500, or 10,000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100,000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1,000,000, 2000000, 3000000, 4000000, 5000000, 10000000, 20000000 or more. The number of partitions employed to partition barcodes or different sets of barcodes may be at least about 5, 10, 50, 100, 250, 500, 750, 1000, 1500, 2000, 2500, 5000, 7500, 10,000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 2000000, 3000000, 4000000, 5000000, 10000000, 20000000 or more. The number of partitions employed to partition barcodes or different sets of barcodes may be less than about 5, 10, 50, 100, 250, 500, 750, 1000, 1500, 2000, 2500, 5000, 7500, 10,000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 2000000, 3000000, 4000000, 5000000, 10000000, or 20000000. The number of partitions employed to partition barcodes may be about 5-10000000, 5-5000000, 5-1,000,000, 10-10,000, 10-5,000, 10-1,000, 1,000-6,000, 1,000-5,000, 1,000-4,000, 1,000-3,000, or 1,000-2,000. In some cases, the partitions may be fluidic droplets.

[00267] As described above, different barcodes or different sets of barcodes (e.g., each set comprising a plurality of identical barcodes or different barcodes) may be partitioned such that each partition generally comprises a different barcode or different barcode set. In some cases, each partition may comprise a different set of identical barcodes, such as an identical set of barcodes coupled to a bead (e.g., a gel bead). Where different sets of identical barcodes are partitioned, the number of identical barcodes per partition may vary. For example, about 100,000 or more different sets of identical barcodes (e.g., a set of identical barcodes attached to a bead) may be partitioned across about 100,000 or more different partitions, such

that each partition comprises a different set of identical barcodes (e.g., each partition comprises a bead coupled to a different set of identical barcodes). In each partition, the number of identical barcodes per set of barcodes may be about 1,000,000 or more identical barcodes (e.g., each partition comprises 1,000,000 or more identical barcodes coupled to one or more beads). In some cases, the number of different sets of barcodes may be equal to or substantially equal to the number of partitions or may be less than the number of partitions. Any suitable number of different barcodes or different barcode sets, number of barcodes per partition, and number of partitions may be combined. Thus, as will be appreciated, any of the above-described different numbers of barcodes may be provided with any of the above-described barcode densities per partition, and in any of the above-described numbers of partitions.

Microfluidic Devices and Droplets

[00268] In some cases, this disclosure provides devices for making beads and for combining beads (or other types of partitions) with samples, e.g., for co-partitioning sample components and beads. Such a device may be a microfluidic device (e.g., a droplet generator). The device may be formed from any suitable material. In some examples, a device may be formed from a material selected from the group consisting of fused silica, soda lime glass, borosilicate glass, poly(methyl methacrylate) PMMA, PDMS, sapphire, silicon, germanium, cyclic olefin copolymer, polyethylene, polypropylene, polyacrylate, polycarbonate, plastic, thermosets, hydrogels, thermoplastics, paper, elastomers, and combinations thereof.

[00269] A device may be formed in a manner that it comprises channels for the flow of fluids. Any suitable channels may be used. In some cases, a device comprises one or more fluidic input channels (e.g., inlet channels) and one or more fluidic outlet channels. In some embodiments, the inner diameter of a fluidic channel may be about 10μm, 20μm, 30μm, 40μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 85μm, 90μm, 100μm, 125μm, or 150μm. In some cases, the inner diameter of a fluidic channel may be more than 10μm, 20μm, 30μm, 40μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 85μm, 90μm, 100μm, 125μm, 150μm or more. In some embodiments, the inner diameter of a fluidic channel may be less than about 10μm, 20μm, 30μm, 40μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 85μm, 90μm, 100μm, 125μm, or 150μm. Volumetric flow rates within a fluidic channel may be any flow rate known in the art.

[00270] As described elsewhere herein, the microfluidic device may be utilized to form beads by forming a fluidic droplet comprising one or more gel precursors, one or more crosslinkers, optionally an initiator, and optionally an aqueous surfactant. The fluidic droplet may be surrounded by an immiscible continuous fluid, such as an oil, which may further comprise a surfactant and/or an accelerator.

[00271] In some embodiments, the microfluidic device may be used to combine beads (e.g., barcoded beads or other type of first partition, including any suitable type of partition described herein) with sample (e.g., a sample of nucleic acids) by forming a fluidic droplet (or other type of second partition, including any suitable type of partition described herein) comprising both the beads and the sample. The fluidic droplet may have an aqueous core surrounded by an oil phase, such as, for example, aqueous droplets within a water-in-oil emulsion. The fluidic droplet may contain one or more barcoded beads, a sample,

amplification reagents, and a reducing agent. In some cases, the fluidic droplet may include one or more of water, nuclease-free water, acetonitrile, beads, gel beads, polymer precursors, polymer monomers, polyacrylamide monomers, acrylamide monomers, degradable crosslinkers, non-degradable crosslinkers, disulfide linkages, acrydite moieties, PCR reagents, primers, polymerases, barcodes, polynucleotides, oligonucleotides, nucleotides, DNA, RNA, peptide polynucleotides, complementary DNA (cDNA), double stranded DNA (dsDNA), single stranded DNA (ssDNA), plasmid DNA, cosmid DNA, chromosomal DNA, genomic DNA, viral DNA, bacterial DNA, mtDNA (mitochondrial DNA), mRNA, rRNA, tRNA, nRNA, siRNA, snRNA, snoRNA, scaRNA, microRNA, dsRNA, probes, dyes, organics, emulsifiers, surfactants, stabilizers, polymers, aptamers, reducing agents, initiators, biotin labels, fluorophores, buffers, acidic solutions, basic solutions, light-sensitive enzymes, pH-sensitive enzymes, aqueous buffer, oils, salts, detergents, ionic detergents, non-ionic detergents, and the like. In summary, the composition of the fluidic droplet will vary depending on the particular processing needs.

[00272] The fluidic droplets may be of uniform size or heterogeneous size. In some cases, the diameter of a fluidic droplet may be about 1μm, 5μm, 10μm, 20μm, 30μm, 40μm, 45μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 90μm, 100μm, 250μm, 500μm, or 1mm. In some cases, a fluidic droplet may have a diameter of at least about 1μm, 5μm, 10μm, 20μm, 30μm, 40μm, 45μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 90μm, 100μm, 250μm, 500μm, 1mm or more. In some cases, a fluidic droplet may have a diameter of less than about 1μm, 5μm, 10μm, 20μm, 30μm, 40μm, 45μm, 50μm, 60μm, 65μm, 70μm, 75μm, 80μm, 90μm, 100μm, 250μm, 500μm, or 1mm. In some cases, fluidic droplet may have a diameter in the range of about 40-75μm, 30-75μm, 20-75μm, 40-85μm, 40-95μm, 20-100μm, 10-100μm, 1-100μm, 20-250μm, or 20-500μm.

[00273] In some embodiments, the device may comprise one or more intersections of two or more fluid input channels. For example, the intersection may be a fluidic cross. The fluidic cross may comprise two or more fluidic input channels and one or more fluidic outlet channels. In some cases, the fluidic cross may comprise two fluidic input channels and two fluidic outlet channels. In other cases, the fluidic cross may comprise three fluidic input channels and one fluidic outlet channel. In some cases, the fluidic cross may form a substantially perpendicular angle between two or more of the fluidic channels forming the cross.

[00274] In some cases, a microfluidic device may comprise a first and a second input channel that meet at a junction that is fluidly connected to an output channel. In some cases, the output channel may be, for example, fluidly connected to a third input channel at a junction. In some cases, a fourth input channel may be included and may intersect the third input channel and outlet channel at a junction. In some cases, a microfluidic device may comprise first, second, and third input channels, wherein the third input channel intersects the first input channel, the second input channel, or a junction of the first input channel and the second input channel.

[00275] As described elsewhere herein, the microfluidic device may be used to generate gel beads from a liquid. For example, in some embodiments, an aqueous fluid comprising one or more gel precursors, one or more crosslinkers and optionally an initiator, optionally an aqueous surfactant, and optionally an

alcohol within a fluidic input channel may enter a fluidic cross. Within a second fluidic input channel, an oil with optionally a surfactant and an accelerator may enter the same fluidic cross. Both aqueous and oil components may be mixed at the fluidic cross causing aqueous fluidic droplets to form within the continuous oil phase. Gel precursors within fluidic droplets exiting the fluidic cross may polymerize forming beads.

[00276] As described elsewhere herein, the microfluidic device (e.g., a droplet generator) may be used to combine sample with beads (e.g., a library of barcoded beads) as well as an agent capable of degrading the beads (*e.g.*, reducing agent if the beads are linked with disulfide bonds), if desired. In some embodiments, a sample (e.g., a sample of nucleic acids) may be provided to a first fluidic input channel that is fluidly connected to a first fluidic cross (e.g., a first fluidic junction). Pre-formed beads (e.g., barcoded beads, degradable barcoded beads) may be provided to a second fluidic input channel that is also fluidly connected to the first fluidic cross, where the first fluidic input channel and second fluidic input channel meet. The sample and beads may be mixed at the first fluidic cross to form a mixture (e.g., an aqueous mixture). In some cases, a reducing agent may be provided to a third fluidic input channel that is also fluidly connected to the first fluidic cross and meets the first and second fluidic input channel at the first fluidic cross. The reducing agent can then be mixed with the beads and sample in the first fluidic cross. In other cases, the reducing agent may be premixed with the sample and/or the beads before entering the microfluidic device such that it is provided to the microfluidic device through the first fluidic input channel with the sample and/or through the second fluidic input channel with the beads. In other cases, no reducing agent may be added.

[00277] In some embodiments, the sample and bead mixture may exit the first fluidic cross through a first outlet channel that is fluidly connected to the first fluidic cross (and, thus, any fluidic channels forming the first fluidic cross). The mixture may be provided to a second fluidic cross (e.g., a second fluidic junction) that is fluidly connected to the first outlet channel. In some cases, an oil (or other suitable immiscible) fluid may enter the second fluidic cross from one or more separate fluidic input channels that are fluidly connected to the second fluidic cross (and, thus, any fluidic channels forming the cross) and that meet the first outlet channel at the second fluidic cross. In some cases, the oil (or other suitable immiscible fluid) may be provided in one or two separate fluidic input channels fluidly connected to the second fluidic cross (and, thus, the first outlet channel) that meet the first outlet channel and each other at the second fluidic cross. Both components, the oil and the sample and bead mixture, may be mixed at the second fluidic cross. This mixing partitions the sample and bead mixture into a plurality of fluidic droplets (e.g., aqueous droplets within a water-in-oil emulsion), in which at least a subset of the droplets that form encapsulate a barcoded bead (e.g., a gel bead). The fluidic droplets that form may be carried within the oil through a second fluidic outlet channel exiting from the second fluidic cross. In some cases, fluidic droplets exiting the second outlet channel from the second fluidic cross may be partitioned into wells for further processing (e.g., thermocycling).

[00278] In many cases, it will be desirable to control the occupancy rate of resulting droplets (or second partitions) with respect to beads (or first partitions). Such control is described in, for example, U.S.

Provisional patent application No. 61/977,804, filed April 4, 2014, the full disclosure of which is incorporated herein by reference in its entirety for all purposes. In general, the droplets (or second partitions) will be formed such that at least 50%, 60%, 70%, 80%, 90% or more droplets (or second partitions) contain no more than one bead (or first partition). Additionally, or alternatively, the droplets (or second partitions) will be formed such that at least 50%, 60%, 70%, 80%, 90% or more droplets (or second partitions) include exactly one bead (or first partition). In some cases, the resulting droplets (or second partitions) may each comprise, on average, at most about one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, or twenty beads (or first partitions). In some cases, the resulting droplets (or second partitions) may each comprise, on average, at least about one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, or more beads (or first partitions).

[00279] In some embodiments, samples may be pre-mixed with beads (e.g., degradable beads) comprising barcodes and any other reagent (e.g., reagents necessary for sample amplification, a reducing agent, etc.) prior to entry of the mixture into a microfluidic device to generate an aqueous reaction mixture. Upon entry of the aqueous mixture to a fluidic device, the mixture may flow from a first fluidic input channel and into a fluidic cross. In some cases, an oil phase may enter the fluidic cross from a second fluidic input channel (e.g., a fluidic channel perpendicular to or substantially perpendicular to the first fluidic input channel) also fluidly connected to the fluidic cross. The aqueous mixture and oil may be mixed at the fluidic cross, such that an emulsion (*e.g.* a gel-water-oil emulsion) forms. The emulsion can comprise a plurality of fluidic droplets (e.g., droplets comprising the aqueous reaction mixture) in the continuous oil phase. In some cases, each fluidic droplet may comprise a single bead (e.g., a gel bead attached to a set of identical barcodes), an aliquot of sample, and an aliquot of any other reagents (e.g., reducing agents, reagents necessary for amplification of the sample, etc.). In some cases, though, a fluidic droplet may comprise a plurality of beads. Upon droplet formation, the droplet may be carried via the oil continuous phase through a fluidic outlet channel exiting from the fluidic cross. Fluidic droplets exiting the outlet channel may be partitioned into wells for further processing (e.g., thermocycling).

[00280] In cases where a reducing agent may be added to the sample prior to entering the microfluidic device or may be added at the first fluidic cross, the fluidic droplets formed at the second fluidic cross may contain the reducing agent. In this case, the reducing agent may degrade or dissolve the beads contained within the fluidic droplet as the droplet travels through the outlet channel leaving the second fluidic cross.

[00281] In some embodiments, a microfluidic device may contain three discrete fluidic crosses in parallel. Fluidic droplets may be formed at any one of the three fluidic crosses. Sample and beads may be combined within any one of the three fluidic crosses. A reducing agent may be added at any one of the three fluidic crosses. An oil may be added at any one of the three fluidic crosses.

[00282] The methods, compositions, devices, and kits of this disclosure may be used with any suitable oil. In some embodiments, an oil may be used to generate an emulsion. The oil may comprise fluorinated oil, silicon oil, mineral oil, vegetable oil, and combinations thereof.

[00283] In some embodiments, the aqueous fluid within the microfluidic device may also contain an alcohol. For example, an alcohol may be glycerol, ethanol, methanol, isopropyl alcohol, pentanol, ethane, propane, butane, pentane, hexane, and combinations thereof. The alcohol may be present within the aqueous fluid at about 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, or 20% (v/v). In some cases, the alcohol may be present within the aqueous fluid at least about 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20% or more (v/v). In some cases, the alcohol may be present within the aqueous fluid for less than about 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, or 20% (v/v).

[00284] In some embodiments, the oil may also contain a surfactant to stabilize the emulsion. For example, a surfactant may be a fluorosurfactant, Krytox lubricant, Krytox FSH, an engineered fluid, HFE-7500, a silicone compound, a silicon compound containing PEG, such as bis krytox peg (BKP). The surfactant may be present at about 0.1%, 0.5%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, 2%, 5%, or 10% (w/w). In some cases, the surfactant may be present at least about 0.1%, 0.5%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, 2%, 5%, 10% (w/w) or more. In some cases, the surfactant may be present for less than about 0.1%, 0.5%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, 2%, 5%, or 10% (w/w).

[00285] In some embodiments, an accelerator and/or initiator may be added to the oil. For example, an accelerator may be Tetramethylethylenediamine (TMEDA or TEMED). In some cases, an initiator may be ammonium persulfate or calcium ions. The accelerator may be present at about 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, or 2% (v/v). In some cases, the accelerator may be present at least about 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, or 2% (v/v) or more. In some cases, the accelerator may be present for less than about 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, or 2% (v/v).

V. Amplification

[00286] DNA amplification is a method for creating multiple copies of small or long segments of DNA. The methods, compositions, devices, and kits of this disclosure may use DNA amplification to attach one or more desired oligonucleotide sequences to individual beads, such as a barcode sequence or random N-mer sequence. DNA amplification may also be used to prime and extend along a sample of interest, such as genomic DNA, utilizing a random N-mer sequence, in order to produce a fragment of the sample sequence and couple the barcode associated with the primer to that fragment.

[00287] For example, a nucleic acid sequence may be amplified by co-partitioning a template nucleic acid sequence and a bead comprising a plurality of attached oligonucleotides (e.g., releasably attached oligonucleotides) into a partition (e.g., a droplet of an emulsion, a microcapsule, or any other suitable type of partition, including a suitable type of partition described elsewhere herein). The attached

oligonucleotides can comprise a primer sequence (e.g., a variable primer sequence such as, for example, a random N-mer, or a targeted primer sequence such as, for example, a targeted N-mer) that is complementary to one or more regions of the template nucleic acid sequence and, in addition, may also comprise a common sequence (e.g., such as a barcode sequence). The primer sequence can be annealed to the template nucleic acid sequence and extended (e.g., in a primer extension reaction or any other suitable nucleic acid amplification reaction) to produce one or more first copies of at least a portion of the template nucleic acid, such that the one or more first copies comprises the primer sequence and the common sequence. In cases where the oligonucleotides comprising the primer sequence are releasably attached to the bead, the oligonucleotides may be released from the bead prior to annealing the primer sequence to the template nucleic acid sequence. Moreover, in general, the primer sequence may be extended via a polymerase enzyme (e.g., a strand displacing polymerase enzyme as described elsewhere herein, an exonuclease deficient polymerase enzyme as described elsewhere herein, or any other type of suitable polymerase, including a type of polymerase described elsewhere herein) that is also provided in the partition. Furthermore, the oligonucleotides releasably attached to the bead may be exonuclease resistant and, thus, may comprise one or more phosphorothioate linkages as described elsewhere herein. In some cases, the one or more phosphorothioate linkages may comprise a phosphorothioate linkage at a terminal internucleotide linkage in the oligonucleotides.

[00288] In some cases, after the generation of the one or more first copies, the primer sequence can be annealed to one or more of the first copies and the primer sequence again extended to produce one or more second copies. The one or more second copies can comprise the primer sequence, the common sequence, and may also comprise a sequence complementary to at least a portion of an individual copy of the one or more first copies, and/or a sequence complementary to the variable primer sequence. The aforementioned steps may be repeated for a desired number of cycles to produce amplified nucleic acids.

[00289] The oligonucleotides described above may comprise a sequence segment that is not copied during an extension reaction (such as an extension reaction that produces the one or more first or second copies described above). As described elsewhere herein, such a sequence segment may comprise one or more uracil containing nucleotides and may also result in the generation of amplicons that form a hairpin (or partial hairpin) molecule under annealing conditions.

[00290] In another example, a plurality of different nucleic acids can be amplified by partitioning the different nucleic acids into separate first partitions (e.g., droplets in an emulsion) that each comprise a second partition (e.g., beads, including a type of bead described elsewhere herein). The second partition may be releasably associated with a plurality of oligonucleotides. The second partition may comprise any suitable number of oligonucleotides (e.g., more than 1,000 oligonucleotides, more than 10,000 oligonucleotides, more than 100,000 oligonucleotides, more than 1,000,000 oligonucleotides, more than 10,000,000 oligonucleotides, or any other number of oligonucleotides per partition described herein). Moreover, the second partitions may comprise any suitable number of different barcode sequences (e.g., at least 1,000 different barcode sequences, at least 10,000 different barcode sequences, at least 100,000

different barcode sequences, at least 1,000,000 different barcode sequences, at least 10,000,000 different barcode sequence, or any other number of different barcode sequences described elsewhere herein).

[00291] Furthermore, the plurality of oligonucleotides associated with a given second partition may comprise a primer sequence (e.g., a variable primer sequence, a targeted primer sequence) and a common sequence (e.g., a barcode sequence). Moreover, the plurality of oligonucleotides associated with different second partitions may comprise different barcode sequences. Oligonucleotides associated with the plurality of second partitions may be released into the first partitions. Following release, the primer sequences within the first partitions can be annealed to the nucleic acids within the first partitions and the primer sequences can then be extended to produce one or more copies of at least a portion of the nucleic acids with the first partitions. In general, the one or more copies may comprise the barcode sequences released into the first partitions.

Amplification within Droplets and Sample Indexing

[00292] Nucleic acid (e.g., DNA) amplification may be performed on contents within fluidic droplets. As described herein, fluidic droplets may contain oligonucleotides attached to beads. Fluidic droplets may further comprise a sample. Fluidic droplets may also comprise reagents suitable for amplification reactions which may include Kapa HiFi Uracil Plus, modified nucleotides, native nucleotides, uracil containing nucleotides, dTTPs, dUTPs, dCTPs, dGTPs, dATPs, DNA polymerase, Taq polymerase, mutant proof reading polymerase, 9 degrees North, modified (NEB), exo (-), exo (-) Pfu, Deep Vent exo (-), Vent exo (-), and acyclonucleotides (acyNTPS).

[00293] Oligonucleotides attached to beads within a fluidic droplet may be used to amplify a sample nucleic acid such that the oligonucleotides become attached to the sample nucleic acid. The sample nucleic acids may comprise virtually any nucleic acid sought to be analyzed, including, for example, whole genomes, exomes, amplicons, targeted genome segments e.g., genes or gene families, cellular nucleic acids, circulating nucleic acids, and the like, and, as noted above, may include DNA (including gDNA, cDNA, mtDNA, etc.) RNA (e.g., mRNA, rRNA, total RNA, etc.). Preparation of such nucleic acids for barcoding may generally be accomplished by methods that are readily available, e.g., enrichment or pull-down methods, isolation methods, amplification methods etc. In order to amplify a desired sample, such as gDNA, the random N-mer sequence of an oligonucleotide within the fluidic droplet may be used to prime the desired target sequence and be extended as a complement of the target sequence. In some cases, the oligonucleotide may be released from the bead in the droplet, as described elsewhere herein, prior to priming. For these priming and extension processes, any suitable method of DNA amplification may be utilized, including polymerase chain reaction (PCR), digital PCR, reverse-transcription PCR, multiplex PCR, nested PCR, overlap-extension PCR, quantitative PCR, multiple displacement amplification (MDA), or ligase chain reaction (LCR). In some cases, amplification within fluidic droplets may be performed until a certain amount of sample nucleic acid comprising barcode may be produced. In some cases, amplification may be performed for about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 cycles. In some cases, amplification may be performed for more than

about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 cycles, or more. In some cases, amplification may be performed for less than about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 cycles.

[00294] An exemplary amplification and barcoding process as described herein, is schematically illustrated in **Fig 38**. As shown, oligonucleotides that include a barcode sequence are co-partitioned in, e.g., a droplet **3802** in an emulsion, along with a sample nucleic acid **3804**. As noted elsewhere herein, the oligonucleotides **3808** may be provided on a bead **3806** that is co-partitioned with the sample nucleic acid **3804**, which oligonucleotides are preferably releasable from the bead **3806**, as shown in panel A. The oligonucleotides **3808** include a barcode sequence **3812**, in addition to one or more functional sequences, e.g., sequences **3810**, **3814** and **3816**. For example, oligonucleotide **3808** is shown as comprising barcode sequence **3812**, as well as sequence **3810** that may function as an attachment or immobilization sequence for a given sequencing system, e.g., a P5 sequence used for attachment in flow cells of an Illumina Hiseq or Miseq system. As shown, the oligonucleotides also include a primer sequence **3816**, which may include a random or targeted N-mer for priming replication of portions of the sample nucleic acid **3804**. Also included within oligonucleotide **3808** is a sequence **3814** which may provide a sequencing priming region, such as a "read1" or R1 priming region, that is used to prime polymerase mediated, template directed sequencing by synthesis reactions in sequencing systems. In many cases, the barcode sequence **3812**, immobilization sequence **3810** and R1 sequence **3814** will be common to all of the oligonucleotides attached to a given bead. The primer sequence **3816** may vary for random N-mer primers, or may be common to the oligonucleotides on a given bead for certain targeted applications.

[00295] Based upon the presence of primer sequence **3816**, the oligonucleotides are able to prime the sample nucleic acid as shown in panel B, which allows for extension of the oligonucleotides **3808** and **3808a** using polymerase enzymes and other extension reagents also co-partitioned with the bead **3806** and sample nucleic acid **3804**. As described elsewhere herein, these polymerase enzymes may include thermostable polymerases, e.g., where initial denaturation of double stranded sample nucleic acids within the partitions is desired. Alternatively, denaturation of sample nucleic acids may precede partitioning, such that single stranded target nucleic acids are deposited into the partitions, allowing the use of non-thermostable polymerase enzymes, e.g., Klenow, phi29, Pol 1, and the like, where desirable. As shown in panel C, following extension of the oligonucleotides that, for random N-mer primers, would anneal to multiple different regions of the sample nucleic acid **3804**; multiple overlapping complements or fragments of the nucleic acid are created, e.g., fragments **3818** and **3820**. Although including sequence portions that are complementary to portions of sample nucleic acid, e.g., sequences **3822** and **3824**, these constructs are generally referred to herein as comprising fragments of the sample nucleic acid **3804**, having the attached barcode sequences. In some cases, it may be desirable to artificially limit the size of the replicate fragments that are produced in order to maintain manageable fragment sizes from the first amplification steps. In some cases, this may be accomplished by mechanical means, as described above, e.g., using fragmentation systems like a Covaris system, or it may be accomplished by incorporating

random extension terminators, e.g., at low concentrations, to prevent the formation of excessively long fragments.

[00296] These fragments may then be subjected to sequence analysis, or they may be further amplified in the process, as shown in panel D. For example, additional oligonucleotides, e.g., oligonucleotide **3808b**, also released from bead **3806**, may prime the fragments **3818** and **3820**. This shown in for fragment **3818**. In particular, again, based upon the presence of the random N-mer primer **3816b** in oligonucleotide **3808b** (which in many cases will be different from other random N-mers in a given partition, e.g., primer sequence **3816**), the oligonucleotide anneals with the fragment **3818**, and is extended to create a complement **3826** to at least a portion of fragment **3818** which includes sequence **3828**, that comprises a duplicate of a portion of the sample nucleic acid sequence. Extension of the oligonucleotide **3808b** continues until it has replicated through the oligonucleotide portion **3808** of fragment **3818**. As noted elsewhere herein, and as illustrated in panel D, the oligonucleotides may be configured to prompt a stop in the replication by the polymerase at a desired point, e.g., after replicating through sequences **3816** and **3814** of oligonucleotide **3808** that is included within fragment **3818**. As described herein, this may be accomplished by different methods, including, for example, the incorporation of different nucleotides and/or nucleotide analogues that are not capable of being processed by the polymerase enzyme used. For example, this may include the inclusion of uracil containing nucleotides within the sequence region **3812** to cause a non-uracil tolerant polymerase to cease replication of that region. As a result, a fragment **3826** is created that includes the full-length oligonucleotide **3808b** at one end, including the barcode sequence **3812**, the attachment sequence **3810**, the R1 primer region **3814**, and the random n-mer sequence **3816b**. At the other end of the sequence will be included the complement **3816'** to the random n-mer of the first oligonucleotide **3808**, as well as a complement to all or a portion of the R1 sequence, shown as sequence **3814'**. The R1 sequence **3814** and its complement **3814'** are then able to hybridize together to form a partial hairpin structure **3828**. As will be appreciated because the random-n-mers differ among different oligonucleotides, these sequences and their complements would not be expected to participate in hairpin formation, e.g., sequence **3816'**, which is the complement to random N-mer **3816**, would not be expected to be complementary to random n-mer sequence **3816b**. This would not be the case for other applications, e.g., targeted primers, where the N-mers may be common among oligonucleotides within a given partition.

[00297] By forming these partial hairpin structures, it allows for the removal of first level duplicates of the sample sequence from further replication, e.g., preventing iterative copying of copies. The partial hairpin structure also provides a useful structure for subsequent processing of the created fragments, e.g. ,fragment **3826**.

[00298] Following attachment of the barcode to the sample, additional amplification steps (*e.g.* PCR) may be performed to amplify the barcoded fragments prior to sequencing, as well as to optionally add additional functional sequences to those barcoded fragments, e.g., additional primer binding sites (*e.g.* Read2 sequence primer, Index primer) that is compatible with a sequencing device (*e.g.* Illumina MiSeq) and optionally, one or more additional barcode sequences (e.g., see **Fig 14C**), as well as other functional sequences, e.g., additional immobilization sequences or their complements, e.g., P7 sequences. In some

cases, an additional barcode sequence may serve as a sample index, with the original barcode and sample index permitting multiplexed sequencing (e.g., simultaneous molecular tagging and sample identification). The original barcode can be used during sequencing to align a sequence read corresponding to the nucleic acid molecule associated with the barcode (e.g., identified via the barcode). A different sample index can be included in sequencer-ready products generated from each different sample. Thus, the sample index can be used during sequencing for identifying the sample to which a particular sequence read belongs and multiplexing can be achieved.

[00299] In some cases, a sample index can be added to a sample nucleic acid after the addition of the original barcode to the sample nucleic acid, with or without the use of partitions or the generation of additional partitions. In some cases, the sample index is added in bulk. In some cases, the addition of a sample index to a sample nucleic acid may occur prior to the addition of a barcode to the sample nucleic acid. In some cases, the addition of a sample index to a sample nucleic acid may occur simultaneous to or in parallel to the addition of a sample index to the sample nucleic acid.

[00300] In some cases, a sample index may be added to a sample nucleic acid after addition of a barcode sequence to the sample nucleic acid. For example, as described elsewhere herein, amplification methods may be used to attach a barcode sequence and other sequences (e.g., P5, R1, etc.) to a sample nucleic acid. In some cases, a random amplification scheme, such as Partial Hairpin Amplification for Sequencing (PHASE – as described elsewhere herein), for example, may aid in attaching a barcode sequence and other sequences to a sample nucleic acid. In one example, a plurality of primers, each comprising a different random N-mer, a sequencer attachment or immobilization site (e.g., P5), a barcode sequence (e.g., an identical barcode sequence), and a sequencing primer binding site (e.g., R1) are used to randomly prime and amplify a sample nucleic acid. Any of the sequencer primer binding site, the barcode sequence, and/or sequencing primer binding site may comprise uracil containing nucleotides. The primer may also include an oligonucleotide blocker hybridized to the primer at one or more sequences of the primer to ensure that priming of the sample nucleic acid occurs only via the random N-mer. A schematic representation of an example primer is as follows (oligonucleotide blocker not shown):

P5-Barcode-R1-RandomNMer

[00301] Random priming of the sample nucleic acid and multiple rounds of amplification can generate amplicons comprising a portion of the sample nucleic acid linked at one end to the sequencer attachment or immobilization site (e.g., P5), the barcode, the sequencing primer binding site (e.g., R1), and the random N-mer. At its other end, the portion of the sample nucleic acid can be linked to a region (e.g., R1c, or R1c partial) that is complementary or partially complementary to the sequencing primer binding site. A schematic representation of an example sequence (in a linear configuration) is as follows:

P5-Barcode-R1-RandomNmer-Insert-R1c,partial

where "Insert" corresponds to the portion of the sample nucleic acid copied during amplification. The

sequencing primer binding site (e.g., R1) and its partial complement (e.g., R1c, partial) at the opposite end of the portion of the copied sample nucleic acid (Insert) can intramolecularly hybridize to form a partial hairpin structure as described elsewhere herein.

[00302] Following creation of the barcoded fragments of the sample nucleic acid, and as noted above, it may be desirable to further amplify those fragments, as well as attach additional functional sequences to the amplified, barcoded fragments. This amplification may be carried out using any suitable amplification process, including, e.g., PCR, LCR, linear amplification, or the like. Typically, this amplification may be initiated using targeted primers that prime against the known terminal sequences in the created fragments, e.g., priming against one or both of the attachment sequence **3810**, in **Fig 38**, and sequence **3814'**. Further by incorporating additional functional sequences within these primers, e.g., additional attachment sequences such as P7, additional sequencing primers, e.g., a read 2 or R2 priming sequence, as well as optional sample indexing sequences, one can further configure the amplified barcoded fragments.

[00303] By way of example, following generation of partial hairpin amplicons, intramolecular hybridization of the partial hairpin amplicons can be disrupted by contacting the partial hairpin amplicons with a primer that is complementary to the duplex portion of the hairpin, e.g., sequence **3814'**, in order to disrupt the hairpin and prime extension along the hairpin structure. In many cases, it will be desirable to provide these primers with a stronger hybridization affinity than the hairpin structure in order to preferentially disrupt that hairpin. As such, in at least one example, the primer comprises a locked nucleic acid (LNAs) or locked nucleic acid nucleotides. LNAs include nucleotides where the ribonucleic acid base comprises a molecular bridge connecting the 2'-oxygen and 4'-carbon of the nucleotide's ribose moiety. LNAs generally have higher melting temperatures and lower hybridization energies. Accordingly, LNAs can favorably compete with intramolecular hybridization of the partial hairpin amplicons by binding to any of the hybridized sequences of a partial hairpin amplicon. Subsequent amplification of the disrupted amplicons via primers comprising LNAs and other primers can generate linear products comprising any additional sequences (including a sample index) to be added to the sequence.

[00304] For the example partial hairpin P5-Barcode-R1-RandomNmer-Insert-R1c,partial configuration described above, the partial hairpin can be contacted with a primer comprising LNAs and a sequence complementary to R1c,partial (e.g., see **Fig 14C**). The primer may also comprise the complement of any additional sequence to be added to the construct. For example, the additional sequence (e.g., R2partial) may be a sequence that, when coupled to R1c,partial, generates an additional sequencing primer binding site (e.g., R2). Hybridization of the primer with the partial hairpin can disrupt the partial hairpin's intramolecular hybridization and linearize the construct. Hybridization may occur, for example, such that the primer hybridizes with R1c,partial via its complementary sequence (e.g., see **Fig 14C**). Extension of the primer can generate a construct comprising the primer linked to a sequence complementary to the linearized partial hairpin amplicon. A schematic of an example construct is as follows:

P5c-Barcode,c-R1c-RandomNmer,c-Insert,c-R1,partial-R2partial,c

where P5c corresponds to the complement of P5, Barcode,c corresponds to the complement of the barcode, RandomNmer,c corresponds to the complement of the random N-mer, Insert,c corresponds to the complement of the portion of the Insert, and R1,partial-R2partial,c corresponds to the complement of R2.

[00305] Upon a further round of amplification with a second primer (e.g., P5, hybridizing at P5c), a linear construct comprising the partial hairpin amplicon sequence and a sequence complementary to the primer can be generated. A schematic representation of an example configuration is as follows:

P5-Barcode-R1-RandomNmer-Insert-R1c,partial-R2partial or

P5-Barcode-R1-RandomNmer-Insert-R2

where the combined sequence of R1c,partial and R2partial can correspond to an additional sequencing primer binding site (e.g., R2).

[00306] Additional sequences can be added to the construct using additional rounds of such amplification, for however many additional sequences/rounds of amplification are desired. For the example P5-Barcode-R1-RandomNmer-Insert-R2 construct described above, a primer comprising a sequence complementary to R2 (e.g., R2c), the complement of a sample index sequence (e.g., SIc, SampleBarcode), and the complement of an additional sequencer primer binding site sequence (e.g., P7c) can be hybridized to the construct at R2, via R2c of the primer (e.g., see **Fig 14C**). Extension of the primer can generate a construct comprising the primer linked to a sequence complementary to the construct. A schematic representation of an example configuration is as follows:

P5c-Barcode,c-R1c-RandomNmer,c-Insert,c-R2,c-SIc-P7c

[00307] Upon a further round of amplification with a second primer (e.g., P5, hybridizing at P5c), a sequencer-ready construct comprising the construct sequence and a sequence complementary to the primer can be generated. A schematic representation of an example configuration of such a sequencer-ready construct is as follows:

[00308] P5-Barcode-R1-RandomNmer-Insert-R2-SampleIndex-P7As an alternative, the starting primer may comprise a barcode sequence, P7, and R2 (instead of P5 and R1). A schematic representation of an example primer is as follows:

P7-Barcode-R2-RandomNmer

[00309] Using an analogous amplification scheme as described above (e.g., amplification with primers comprising LNAs, additional rounds of amplification, etc.), an insert comprising a portion of a sample

nucleic acid to be sequenced, P5, R1, and a sample index can be added to the primer to generate a sequencer-ready product. A schematic representation of an example product is as follows:

P7-Barcode-R2-RandomNmer-Insert-R1-SampleIndex-P5

[00310] In other cases, a sample index may be added to a sample nucleic acid concurrently with the addition of a barcode sequence to the sample nucleic acid. For example, a primer used to generate a barcoded sample nucleic acid may comprise both a barcode sequence and a sample index, such that when the barcode is coupled to the sample nucleic acid, the sample index is coupled simultaneously. The sample index may be positioned anywhere in the primer sequence. In some cases, the primer may be a primer capable of generating barcoded sample nucleic acids via random amplification, such as PHASE amplification. Schematic representations of examples of such primers include:

P5-Barcode-R1-SampleIndex-RandomNmer

P5-Barcode-SampleIndex-R1-RandomNmer

P5-SampleIndex-Barcode-R1-RandomNmer

[00311] Upon random priming of a sample nucleic acid with a respective primer and amplification of the sample nucleic acid in the partition, partial hairpin amplicons comprising a barcode sequence and a sample index sequence can be generated. Schematic representations (shown in linear form) of examples of such partial hairpin amplicons generated from the above primers include, respectively:

P5-Barcode-R1-SampleIndex-RandomNmer-Insert-R1c,partial

P5-Barcode-SampleIndex-R1-RandomNmer-Insert-R1c,partial

P5-SampleIndex-Barcode-R1-RandomNmer-Insert-R1c,partial

R1c, partial can intramolecularly hybridize with its complementary sequence in R1 to form a partial hairpin amplicon.

[00312] By way of example, in some cases, following the generation of partial hairpin amplicons, additional sequences (e.g., functional sequences like R2 and P7 sequences) can be added to the partial hairpin amplicons, such as, for example, in bulk. In analogous fashion to amplification methods described elsewhere herein, primers that include these additional functional sequences may be used to prime the replication of the partial hairpin molecule, e.g., by priming against the 5' end of the partial hairpin, e.g., the R1c sequence, described above. In many cases, it will be desirable to provide a higher affinity primer sequence, e.g., to outcompete rehybridization of the hairpin structure, in order to provide greater priming and replication. In such cases, tighter binding primer sequences, e.g., that include in their sequence one or more higher affinity nucleotide analogues, like LNAs or the like, may be used to disrupt

partial hairpin amplicons and add additional sequences to the amplicons. For example, with reference to the example described above, a primer may comprise LNAs, a sequence complementary to R1c,partial and a sequence comprising the complement to R2partial, such that when the primer is extended and the resulting product further amplified via a P5 primer, R1c,partial and R2partial are joined to generate R2. Schematic representations of examples of such constructs generated from the above primers include, respectively:

P5-Barcode-R1-SampleIndex-RandomNmer-Insert-R2

P5-Barcode-SampleIndex-R1-RandomNmer-Insert-R2

P5-SampleIndex-Barcode-R1-RandomNmer-Insert-R2

[00313] As noted above, additional rounds of amplification cycles may be used to add additional sequences to the constructs. For example, a primer may comprise a sequence complementary to R2 and a sequence comprising the complement to P7, such that when the primer is extended and the resulting product further amplified via a P5 primer, P7 is linked to R2 and a sequencer-ready construct is generated. Schematic representations of examples of such sequencer-ready constructs generated from the above primers include, respectively:

P5-Barcode-R1-SampleIndex-RandomNmer-Insert-R2-P7

P5-Barcode-SampleIndex-R1-RandomNmer-Insert-R2-P7

P5-SampleIndex-Barcode-R1-RandomNmer-Insert-R2-P7

[00314] Combining a barcode and a sample index into a primer capable of amplifying regions of a sample nucleic acid (e.g., via PHASE amplification) may enable parallelization of sample indexing. Sets of primers may be used to index nucleic acids from different samples. Each set of primers may be associated with nucleic acid molecules obtained from a particular sample and comprise primers comprising a diversity of barcode sequences and a common sample index sequence.

[00315] In some cases, it may be desirable to attach additional sequence segments to the 5' end of the partial hairpin molecules described herein, not only to provide additional functionality to the amplified fragment of the sample nucleic acid as described above, but also to ensure more efficient subsequent processing , e.g., amplification and/or sequencing, of those molecules. For example, where a partial hairpin molecule is subjected to extension reaction conditions, it may be susceptible to filling in of the partial hairpin structure, by priming its own 'filling in' reaction through extension at the 5' terminus. As a result, a complete hairpin structure may be created that is more difficult to amplify, by virtue of the greater stability of its duplex portion. In such cases, it may be desirable to preferentially attach additional sequence segment(s) that is not complementary to the opposing end sequence, in order to prevent the formation of a complete hairpin structure. In one exemplary process, the LNA primers described above for the amplification of the partial hairpin structures, may be provided with additional overhanging

sequence, including, e.g., the R2 complementary sequence described above, as well as potentially complementary sequences to other functional sequence components, e.g., attachment sequences like P7, sample index sequences, and the like. Subjecting the partial hairpin and primer to the extension reaction described above for amplification of that partial hairpin, will also result in extension of the partial hairpin along the overhanging sequence on the LNA primer. The extended sequence may comprise simply a non-complementary sequence, or it may comprise additional functional sequences, or their complements as noted above, such that the extension reaction results in attachment of those functional sequences to the 5' terminus of the partial hairpin structure.

[00316] In alternative aspects, additional sequence segments may be ligated to the 5' end of the partial hairpin structure where such sequence segments are not complementary to the non-overlapped portion of the hairpin structure. The foregoing are schematically illustrated in **Fig 40**. As shown in path A, a partial hairpin structure, when subjected to primer extension conditions, may act as its own primer and have its 5' sequence extended, as shown by the dashed arrow, until it forms a complete or nearly complete hairpin structure, e.g., with little or no overhang sequence. This full hairpin structure will possess far greater duplex stability, thereby potentially negatively impacting the ability to disrupt the hairpin structure to prime its replication, even when employing higher affinity primers, e.g., LNA containing primers/probes.

[00317] In order to minimize this possibility, as shown in both paths B and C, a separate sequence segment **4006** is added to the 5'end of the hairpin structure, to provide a partial hairpin with non-complementary tail sequences **4008**, in order to prevent the generation of the complete or nearly complete hairpin structure. As shown, this may be accomplished in a number of different ways. For example, in a first process shown in path B, an invading probe **4010** may be used to disrupt the partial hairpin structure and hybridize to sequence segment **4012**. Such invading probes may be provided with higher affinity binding than the inherent partial hairpin structure, e.g., through use of higher affinity nucleotide analogues such as LNAs or the like. In particular, that portion of the invader sequence **4010** that hybridizes to sequence segment **4012** may comprise LNAs within its sequence in the same fashion described herein for use with LNA primer sequences used in subsequent amplification.

[00318] Extension of the 5' portion of the partial hairpin (and sequence segment **4012**) as shown by the dashed arrow in path B, then appends the sequence **4006** to the 5' terminus of the partial hairpin structure to provide structure **4008**. Alternatively, sequence **4006** may be ligated to the 5' end of the partial hairpin structure **4002** (or sequence segment **4012**). As shown in path C, this achieved through the use of a splint sequence **4014** that is partially complementary to sequence **4006** and partially complementary to sequence **4012**, in order to hold sequence **4006** adjacent to sequence segment **4012** for ligation. As will be appreciated, the splint sequence **4014** may again utilize a higher affinity invading probe, like probe **4010**, to disrupt the hairpin structure and hybridize to sequence segment **4012**. In particular, again, that portion of splint sequence **4014** that is intended to hybridize to sequence segment **4012** may be provided with one or more LNA nucleotide analogues within its sequence, in order to preferentially disrupt the partial hairpin structure **4002**, and allow ligation of sequence **4006** to its 5' end.

[00319] In some cases, a microfluidic device (e.g., a microfluidic chip) may be useful in parallelizing sample indexing. Such a device may comprise parallel modules each capable of adding a barcode sequence and a sample index to nucleic acid molecules of a sample via primers comprising both the barcode sequence and the sample index. Each parallel module may comprise a primer set comprising a different sample index, such that the sample processed in each module is associated with a different sample index and set of barcodes. For example, a microfluidic device with 8 modules may be capable of sample indexing 8 different samples. Following barcoding and sample indexing via attachment of the sequences to a sample nucleic acid, bulk addition of additional sequences (e.g., R2, P7, other barcode sequences) via, for example, serial amplification can be used to generate sequencer-ready products as described elsewhere herein.

[00320] In some cases, sample indexing may be achieved during barcoding without the inclusion of a separate sample index sequence in a primer used to attach a barcode to a sample nucleic acid. In such cases, a barcode sequence, for example, may also serve as a sample index. An example configuration of a sequencer-ready construct with a sequence functioning as both a barcode sequence and a sample index is as follows:

P5-BSI-R1- RandomNmer-Insert-R2-P7

where "BSI" is the sequence functioning as both a barcode sequence and a sample index.

A sequencer-ready product may comprise a barcode sequence that can be used to align sequence reads and provide a sequence for a sample nucleic acid. The sequencer-ready product may be generated, for example, using PHASE amplification and subsequent bulk amplification as described elsewhere herein. Moreover, the barcode sequence may belong to a particular set of known barcode sequences. The set of barcode sequences may be associated with a particular sample, such that identification of the sample from which a particular sequencing read originates can be achieved via the read barcode sequence. Each sample can be associated with a set of known barcode sequences, with each barcode sequence set comprising barcode sequences that do not overlap with barcode sequence in other barcode sets associated with other samples. Thus, the uniqueness of a barcode sequence and its uniqueness amongst different sets of barcode sequences may be used for multiplexing.

[00321] For example, a sequencing read may comprise the barcode sequence "GAGCCG". Barcode sequence "GAGCCG" may be a barcode sequence in a set of known barcode sequences associated with Sample A. The sequence is not found in a set of known barcode sequences associated with another sample. Upon reading the sequence "GAGCCG", it can be determined that the sequence read is associated with Sample A because the sequence "GAGCCG" is unique to the set of barcode sequences associated with Sample A. Moreover, another sequencing read may comprise the barcode sequence "AGCAGA". Barcode sequence "AGCAGA" may be a barcode sequence in a set of known barcode sequences associated with Sample B. The sequence is not found in a set of known barcode sequences

associated with another sample. Upon reading the sequence "AGCAGA", it can be determined that the sequence read is associated with Sample B because "AGCAGA" is unique to the set of barcode sequences associated with Sample B.

[00322] In another example, a sample index sequence may be embedded in a random sequence of a primer used in one or more amplification reactions to attach a barcode to a sample nucleic acid. For example, a primer may comprise a barcode sequence and a random sequence that can be used to randomly prime a sample nucleic acid and attach the barcode sequence to the sample nucleic acid. In some cases, the random sequence may be a pseudo-random sequence such that particular bases of the random sequence are conserved between all primers. The pattern of the conserved bases may be used as a sample index, such that all sequencer-ready products obtained from a particular sample all comprise the conserved pattern of bases in the random sequence region. Each sample can be associated with a different pattern of conserved bases and, thus, multiplexing can be achieved. In some cases, the pattern is a contiguous sequence region of a pseudo-random sequence (e.g., "NNNATACNNN") or in other cases, the pattern is a non-contiguous sequence region of a pseudo-random sequence (e.g., "NCNGNNAANN"), where "N" corresponds to a random base. Moreover, any suitable number of bases may be conserved in a pseudo-random sequence in any pattern and the examples described herein are not meant to be limiting. An example configuration of a sequencer-ready construct with a sequence functioning as both a barcode sequence and a sample index is as follows:

P5-Barcode-R1- NQNQNNQQNN-Insert-R2-P7

where "Q" is a conserved base in the random region

[00323] For example, a sequencer-ready product may comprise a 10-mer pseudo-random sequence "NCNGNNAANN", where the second base ("C"), fourth base ("G"), seventh base ("A"), and eighth base ("A") of the pseudo-random sequence are conserved for all sequencer-ready products generated from Sample A. A sequencing read may comprise such a pattern of conserved bases in the random sequence region. Upon reading the conserved base pattern, it can be determined that the sequence read is associated with Sample A because the "NCNGNNAANN" conserved pattern of bases is associated with Sample A. Moreover, a sequencer-ready product may comprise a 10-mer pseudo-random sequence "NNGCNGNGNN", where the third base ("G"), fourth base ("C"), sixth base ("G"), and eighth base ("G") of the pseudo-random sequence are conserved for all sequencer-ready products generated from Sample B. A sequencing read may comprise such a pattern of conserved bases in the random sequence region. Upon reading the conserved base pattern, it can be determined that the sequence read is associated with Sample B because the "NNGCNGNGNN" conserved pattern of bases is associated with Sample B.

[00324] In other cases, a sample index may be added to a sample nucleic acid prior to the addition of a barcode sequence to the sample nucleic acid. For example, a sample nucleic acid may be pre-amplified in bulk such that resulting amplicons are attached to a sample index sequence prior to barcoding. For example, sample may be amplified with a primer comprising a sample index sequence such that the

sample index sequence can be attached to the sample nucleic acid. In some cases, the primer may be a random primer (e.g., comprising a random N-mer) and amplification may be random. Produced amplicons that comprise the sample index can then be barcoded using any suitable method, including barcoding methods described herein.

[00325] Sample nucleic acid molecules can be combined into partitions (e.g., droplets of an emulsion) with the primers described above. In some cases, each partition can comprise a plurality of sample nucleic acid molecules (e.g., smaller pieces of a larger nucleic acid). In some cases, no more than one copy of a unique sample nucleic acid molecule is present per partition. In some cases, each partition can generally comprise primers comprising an identical barcode sequence and a sample priming sequence (e.g., a variable random-Nmer, a targeted N-mer), with the barcode sequence generally differing between partitions. In such cases, each partition (and, thus, sample nucleic acid in the partition) can be associated with a unique barcode sequence and the unique barcode sequence can be used to determine a sequence for the barcoded sample nucleic acid generated in the partition.

[00326] In some cases, upon generation of barcoded sample nucleic acids, the barcoded sample nucleic acids can be released from their individual partitions, pooled, and subject to bulk amplification schemes to add additional sequences (e.g., additional sequencing primer binding sites, additional sequencer primer binding sites, additional barcode sequences, sample index sequences) common to all downstream sequencer-ready products. In cases where the partitions are droplets of an emulsion, the emulsion may be broken and the barcoded sample nucleic acids pooled. A sample index can be added in bulk to the released, barcoded sample nucleic acids, for example, using the serial amplification methods described herein. Where a sample index is added in bulk, each sequencer-ready product generated from the same sample will comprise the same sample index that can be used to identify the sample from which the read for the sequencer-ready product was generated. Where a sample index is added during barcoding, each primer used for barcoding may comprise an identical sample index sequence, such that each sequencer-ready product generated from the same sample will comprise the same sample index sequence.

[00327] Partitioning of sample nucleic acids to generate barcoded (or barcoded and sample indexed) sample nucleic acids and subsequent addition of additional sequences (e.g., including a sample index) to the barcoded sample nucleic acids can be repeated for each sample, using a different sample index for each sample. In some cases, a microfluidic droplet generator may be used to partition sample nucleic acids. In some cases, a microfluidic chip may comprise multiple droplet generators, such that a different sample can be processed at each droplet generator, permitting parallel sample indexing. Via each different sample index, multiplexing during sequencing can be achieved.

[00328] Upon the generation of sequencer-ready oligonucleotides, the sequencer-ready oligonucleotides can then be provided to a sequencing device for sequencing. Thus, for example, the entire sequence provided to the sequencing device may comprise one or more adaptors compatible with the sequencing device (*e.g.* P5, P7), one or more barcode sequences, one or more primer binding sites (*e.g.* Read1 (R1) sequence primer, Read2 (R2) sequencing primer, Index primer), an N-mer sequence, a universal sequence, the sequence of interest, and combinations thereof. The barcode sequence may be located at either end of

the sequence. In some cases, the barcode sequence may be located between P5 and Read1 sequence primer binding site. In other cases, the barcode sequence may be located between P7 and Read 2 sequence primer binding site. In some cases, a second barcode sequence may be located between P7 and Read 2 sequence primer binding site. The index sequence primer binding site may be utilized in the sequencing device to determine the barcode sequence.

[00329] The configuration of the various components (e.g., adaptors, barcode sequences, sample index sequences, sample sequence, primer binding sites, etc.) of a sequence to be provided to a sequencer device may vary depending on, for example the particular configuration desired and/or the order in which the various components of the sequence is added. Any suitable configuration for sequencing may be used and any sequences can be added to oligonucleotides in any suitable order. Additional sequences may be added to a sample nucleic acid prior to, during, and after barcoding of the sample nucleic acid. For example, a P5 sequence can be added to a sample nucleic acid during barcoding and P7 can be added in bulk amplification following barcoding of the sample nucleic acid. Alternatively, a P7 sequence can be added to a sample nucleic acid during barcoding and a P5 sequence can be added in bulk amplification following barcoding of the sample nucleic acid. Example configurations displayed as examples herein are not intended to be limiting. Moreover, the addition of sequence components to an oligonucleotide via amplification is also not meant to be limiting. Other methods, such as, for example, ligation may also be used. Furthermore, adaptors, barcode sequences, sample index sequences, primer binding sites, sequencer-ready products, etc. described herein are not meant to be limiting. Any type of oligonucleotide described herein, including sequencer-ready products, may be generated for any suitable type of sequencing platform (e.g., Illumina sequencing, Life Technologies Ion Torrent, Pacific Biosciences SMRT, Roche 454 sequencing, Life Technologies SOLiD sequencing, etc.) using methods described herein.

[00330] Sequencer-ready oligonucleotides can be generated with any adaptor sequence suitable for a particular sequencing platform using methods described herein. For example, sequencer-ready oligonucleotides comprising one or more barcode sequences and P1 and A adaptor sequences useful in Life Technologies Ion Torrent sequencing may be generated using methods described herein. In one example, beads (e.g., gel beads) comprising an acrydite moiety linked to a P1 sequence via a disulfide bond may be generated. A barcode construct may be generated that comprises a P1sequence, a barcode sequence, and a random N-mer sequence. The barcode construct may enter an amplification reaction (e.g., in a partition, such as a fluidic droplet) to barcode sample nucleic acid. Barcoded amplicons may then be subject to further amplification in bulk to add the A sequence and any other sequence desired, such as a sample index. Alternatively, P1 and A sequences can be interchanged such that A is added during sample barcoding and P1 is added in bulk. The complete sequence can then be entered into an Ion Torrent sequencer. Other adaptor sequences (e.g., P1 adaptor sequence for Life Technologies SOLiD sequencing, A and B adaptor sequences for Roche 454, etc.) for other sequencing platforms can be added in analogous fashion.

[00331] Although described herein as generating partial hairpin molecules, and in some cases, preventing formation of complete hairpins, in some cases, it may be desirable to provide complete hairpin fragments that include the barcode sequences described herein. In particular, such complete hairpin molecules may be further subjected to conventional sample preparation steps by treating the 3' and 5' end of the single hairpin molecule as one end of a double stranded duplex molecule in a conventional sequencing workflow. In particular, using conventional ligation steps, one could readily attach the appropriate adapter sequences to both the 3' and 5' end of the hairpin molecule in the same fashion as those are attached to the 3' and 5' termini of a duplex molecule. For example, in case of an Illumina based sequencing process, one could attach a standard Y adapter that includes the P5 and P7 adapters and R1 and R2 primer sequences, to one end of the hairpin as if it were one end of a duplex molecule, using standard Illumina protocols.

Methods for Reducing Undesired Amplification Products (Partial Hairpin Amplification for Sequencing (PHASE))

[00332] A random N-mer sequence may be used to randomly prime a sample, such as genomic DNA (gDNA). In some embodiments, the random N-mer may comprise a primer. In some cases, the random N-mer may prime a sample. In some cases, the random N-mer may prime genomic DNA. In some cases, the random N-mer may prime DNA fragments.

[00333] Additionally, a random N-mer sequence may also be attached to another oligonucleotide. This oligonucleotide may be a universal sequence and/or may contain one or more primer read sequences that may be compatible with a sequencing device (*e.g.* Read 1 primer site, Read 2 primer site, Index primer site), one or more barcode sequences, and one or more adaptor segments that may be compatible with a sequencing device (*e.g.* P5, P7). Alternatively, the oligonucleotide may comprise none of these and may include another sequence.

[00334] Via subsequent amplification methods, priming of a sample nucleic acid with a random N-mer may be used to attach an oligonucleotide sequence (e.g., an oligonucleotide sequence comprising a barcode sequence) linked to a random N-mer to the sample nucleic acid, including a sample nucleic acid to be sequenced. Utilizing random primers to prime a sample may introduce significant sequence read errors, due to, for example, the production of undesired amplification products.

[00335] To mitigate undesired amplification products, at least a subsection of an oligonucleotide sequence may be substituted with dUTPs or uracil containing nucleotides in place of dTTPs or thymine containing nucleotides, respectively. In some cases, substitution may be complete (e.g., all thymine containing nucleotides are substituted with uracil containing nucleotides), or may be partial such that a portion of an oligonucleotide's thymine containing nucleotides are substituted with uracil containing nucleotides. In some cases, thymine containing nucleotides in all but the last about 10 to about 20, last about 10 to 30, last about 10 to 40, or last about 5 to 40 nucleotides of an oligonucleotide sequence adjacent to a random N-mer sequence are substituted with dUTPs or uracil containing nucleotides. In addition, a polymerase that does not accept or process uracil-containing templates may be used for amplification of the sample nucleic acid. In this case, the non-uracil containing portion of about 10 to about 20 nucleotides may be

amplified and the remaining portion containing the dUTPs or uracil containing nucleotides may not be amplified. In some cases, the portion of an oligonucleotide sequence comprising dUTPs or uracil containing nucleotides may be adjacent to the N-mer sequence. In some cases, the portion of an oligonucleotide sequence comprising dUTPs or uracil containing nucleotides may be adjacent to the barcode sequence. Any portion of an oligonucleotide sequence, including an adaptor segment, barcode, or read primer sequence may comprise dUTPs or uracil containing nucleotides (e.g., substituted for thymine containing nucleotides), depending upon the configuration of the oligonucleotide sequence.

[00336] Moreover, the number and positioning of uracil containing nucleotide-for-thymine containing nucleotide substitutions in an oligonucleotide may be used, for example, to tune the size of partial hairpin products obtained with amplification methods described below and/or to tune the binding of the polymerase enzyme with a uracil containing primer sequence. Additionally, free uracil containing nucleotides, e.g., UTP or an analogue thereof, may also be provided within the reaction mixture, e.g., within the partition, at a desired concentration to mediate polymerase/uracil-primer binding kinetics. In some cases, smaller partial hairpin products may give rise to more accurate sequencing results. Accordingly, an oligonucleotide may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more uracil containing nucleotide-for-thymine containing nucleotide substitutions depending upon, for example, the desired length of partial hairpin products generated from the oligonucleotide.

[00337] Upon random priming of a sample nucleic acid with a random N-mer linked to an oligonucleotide sequence (e.g., an oligonucleotide sequence comprising uracil containing nucleotides described above) **Fig 15A**, a first round of amplification (e.g., using a polymerase that does not accept or process a uracil containing nucleotide as a template) may result in the attachment of the oligonucleotide sequence to a complement of the sample nucleic acid, **Fig 15B** and **Fig 15C**. Upon priming (via the random N-mer) and further amplification of the amplification product with another copy of the oligonucleotide sequence comprising the random N-mer (**Fig 15D**), an amplification product comprising the oligonucleotide sequence, a portion of the sample nucleic acid sequence, and a partial complementary oligonucleotide sequence (e.g., complementary to the portion of the oligonucleotide sequence not comprising uracil containing nucleotides) at an end of the amplification product opposite the oligonucleotide sequence, can be generated. The partial complementary oligonucleotide sequence and the oligonucleotide sequence can hybridize to form a partial hairpin that, in some cases, can no longer participate in nucleic acid amplification. A partial hairpin can be generated because a portion of the original oligonucleotide sequence comprising uracil containing nucleotides was not copied. Amplification can continue for a desired number of cycles (e.g., 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 cycles), up until all oligonucleotide sequences comprising random N-mers have been exhausted (**Fig 15E-G**).

[00338] In some embodiments, to ensure priming of sample nucleic acid (e.g., genomic DNA (gDNA)) with only a random N-mer and not portions of an attached oligonucleotide sequence, the oligonucleotide sequence may be blocked via hybridization of a blocker oligonucleotide (e.g., black dumbbell in **Fig 15**). A blocker oligonucleotide (also referred to as an oligonucleotide blocker elsewhere herein) may be

hybridized to any portion of an oligonucleotide sequence, including a barcode sequence, read primer site sequence, all or a portion of a uracil containing portion of the oligonucleotides, or all or any other portion of the oligonucleotides, or other sequence therein. A blocker oligonucleotide may be DNA or RNA. In some cases, a blocker oligonucleotide may comprise uracil containing nucleotide-for-thymine containing nucleotide substitutions. In some cases, all of the thymine containing nucleotides of a blocker oligonucleotide may be substituted with uracil containing nucleotides. In some cases, a portion of the thymine containing nucleotides of a blocker oligonucleotide may be substituted with uracil containing nucleotides. In some cases, a blocker oligonucleotide may comprise locked nucleic acid (LNA), an LNA nucleotide, bridged nucleic acid (BNA), and/or a BNA nucleotide. Moreover a blocker oligonucleotide may be of any suitable length necessary for blocker functionality. A blocker oligonucleotide may be of length suitable to block a portion of an oligonucleotide or may be of the same or of substantially the same length of an oligonucleotide it is designed to block. The blocker oligonucleotide may ensure that only random N-mers bind to the sample nucleic acid (*e.g.*, genomic DNA) and not other portions of the oligonucleotide sequence.

[00339] The stoichiometric ratio of a blocker oligonucleotide to oligonucleotide (e.g., blocker oligonucleotide:oligonucleotide) may vary. For example, the blocker oligonucleotide:oligonucleotide stoichiometric ratio may be about 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00, 1.05, 1.10, 1.15, 1.20, 1.25, 1.30, 1.35, 1.40, 1.45, 1.50, 1.55, 1.60, 1.65, 1.70, 1.75, 1.80, 1.85, 1.90, 1.95, 2.00, 2.10, 2.20, 2.30, 2.40, 2.50, 2.60, 2.70, 2.80, 2.90, 3.00, 3.50, 4.00, 4.50, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 10.0, 20, 30, 40, 50, 100 or more. In some cases, the blocker oligonucleotide:oligonucleotide stoichiometric ratio may be at least about 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00, 1.05, 1.10, 1.15, 1.20, 1.25, 1.30, 1.35, 1.40, 1.45, 1.50, 1.55, 1.60, 1.65, 1.70, 1.75, 1.80, 1.85, 1.90, 1.95, 2.00, 2.10, 2.20, 2.30, 2.40, 2.50, 2.60, 2.70, 2.80, 2.90, 3.00, 3.50, 4.00, 4.50, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 10.0, 20, 30, 40, 50, 100 or more. In some cases, the blocker oligonucleotide:oligonucleotide stoichiometric ratio may be at most about 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00, 1.05, 1.10, 1.15, 1.20, 1.25, 1.30, 1.35, 1.40, 1.45, 1.50, 1.55, 1.60, 1.65, 1.70, 1.75, 1.80, 1.85, 1.90, 1.95, 2.00, 2.10, 2.20, 2.30, 2.40, 2.50, 2.60, 2.70, 2.80, 2.90, 3.00, 3.50, 4.00, 4.50, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 10.0, 20, 30, 40, 50, or 100.

[00340] Moreover, incorporation of a blocker moiety (e.g., via a dideoxynucleotide (ddNTP), ddCTP, ddATP, ddGTP, ddTTP, etc. at the 3' or 5' end of the blocker oligonucleotide) to a blocker oligonucleotide and/or the inclusion of uracil containing nucleotides (e.g., substituted for all or a portion of thymine containing nucleotides) in a blocker oligonucleotide may prevent preferential binding of blocked portions of the blocked oligonucleotide sequence to the sample nucleic acid. Additional examples of blocker moieties include 3' phosphate, a blocked 3' end, 3'ddCTP, C3 Spacer (/3SpC3/), Dideoxy-C (/3ddC/). Blocker oligonucleotides may be cleaved from an oligonucleotide sequence by RNAse, RNAseH, an antisense DNA oligonucleotide, and/or alkaline phosphatase.

[00341] In some cases, an oligonucleotide sequence may be blocked with a blocker oligonucleotide such that the oligonucleotide sequence comprises a blocked 5' end, comprises a blocked 3' end, may be entirely blocked (e.g., may be entirely blocked, except for its random N-mer sequence), or may be blocked at another location (e.g., a partial sequence of the oligonucleotide, different from an oligonucleotide sequence's random N-mer). In some cases, an oligonucleotide sequence may comprise a plurality of blockers, such that multiple sites of the oligonucleotide are blocked. In some cases, an oligonucleotide sequence may comprise both a blocked 3' end and uracil containing nucleotides. In some cases, an oligonucleotide sequence comprising uracil containing nucleotides and a blocked 3' end may be adjacent to the N-mer sequence. In some cases, an oligonucleotide sequence may comprise a blocked 3' end. In some cases, an oligonucleotide sequence may comprise uracil containing nucleotides. In some cases, an oligonucleotide sequence may comprise both a blocked 5' end and uracil containing nucleotides.

[00342] In some cases, the oligonucleotide sequence comprising uracil containing nucleotides and a blocked 3' end may be adjacent to the N-mer sequence. In some cases, the oligonucleotide sequence comprising uracil containing nucleotides and a blocked 3' end may be adjacent to the barcode sequence. In some cases, the oligonucleotide sequence may comprise a blocked 3' end. In some cases, the oligonucleotide sequence may comprise uracil containing nucleotides. In some cases, the oligonucleotide sequence may comprise both the blocked 3' end and uracil containing nucleotides. Addition of a blocker oligonucleotide may prevent preferential binding to portions of the universal sequence, which may not be desired to be amplified.

[00343] In some cases, an oligonucleotide suitable for priming a sample nucleic acid via its random N-mer may also comprise a blocking sequence that can function in the same role as a blocker oligonucleotide. For example, an oligonucleotide may be arranged in a hairpin configuration with a blocking sequence that can function in the same role as a blocker oligonucleotide. An example oligonucleotide comprising a random N-mer, an R1c sequence, a P5 sequence, a barcode sequence, and an R1 sequence may be configured as follows:

5'-RandomNmer-R1c-P5-Barcode-R1-3'

The R1 sequence and R1c sequence of the oligonucleotide may hybridize to generate a hairpin with a hairpin loop comprising the P5 and Barcode sequences. The R1c sequence can function in the same role as a blocker oligonucleotide such that priming of sample nucleic acid with the oligonucleotide occurs via only the oligonucleotide's random N-mer. In some cases, one or more cleavage sites (e.g., a restriction site, a cleavage site, an abasic site, etc.) may be included in an oligonucleotide arranged as a hairpin with a blocking sequence, including an oligonucleotide's hairpin loop, to separate sequence components of the oligonucleotide downstream, if desired. Separation may occur, for example, via an enzymatic reaction, oxidation-reduction, radiation (e.g., UV-light), the addition of heat, or other suitable means.

[00344] An example uracil containing nucleotide-substituted oligonucleotide sequence linked to a random N-mer is depicted in **Fig 14B**. Specifically, a random primer (e.g., a random N-mer), of about 8N-12N in

length, **1404**, may be linked with an oligonucleotide sequence. The random N-mer may be used to randomly prime and extend from a sample nucleic acid, such as, genomic DNA (gDNA). The oligonucleotide sequence comprises: (1) sequences for compatibility with a sequencing device, such as, a flow cell (*e.g.* Illumina's P5, **1401**, and Read 1 Primer sites, **1402**) and (2) a barcode (BC), **1403**, (*e.g.*, 6-12 base sequences). Furthermore, the Read 1 Primer site **1402** of the oligonucleotide sequence may be hybridized with a blocking oligonucleotide comprising uracil containing nucleotides and a blocker moiety at its 3' end (*e.g.* 3'ddCTP, indicated by an "X"). The blocking oligonucleotide can be used to promote priming of a sample nucleic acid with only the random N-mer sequence and prevent preferential binding of the oligonucleotide sequence to portions of the sample nucleic acid that are complementary to the Read 1 Primer site, **1402**. Optionally, to further limit product lengths, a small percentage of terminating nucleotides (*e.g.*, 0.1-2% acyclonucleotides (acyNTPs)) (**Fig 16B**) may be included in oligonucleotide sequences to reduce undesired amplification products.

[00345] An example of partial hairpin amplification for attaching a uracil containing nucleotide-substituted oligonucleotide sequence comprising a random N-mer to a sample nucleic acid (e.g., genomic DNA (gDNA)) is depicted in **Fig 15**. First, initial denaturation of the sample nucleic acid may be achieved at a denaturation temperature (e.g., 98°C, for 2 minutes) followed by priming of a random portion of the sample nucleic acid with the random N-mer sequence at a priming temperature (e.g., 30 seconds at 4°C), **Fig 15A**. The oligonucleotide sequence is hybridized with a blocking oligonucleotide (black dumbbell in **Fig 15**), to ensure that only the random N-mer primes the sample nucleic acid and not another portion of the oligonucleotide sequence. Subsequently, sequence extension (e.g., via polymerase that does not accept or process a uracil containing nucleotide as a template) may follow as the temperature ramps to higher temperature (e.g., at 0.1°C/second to 45°C (held for 1 second)) (**Fig 15A**). Extension may then continue at elevated temperatures (e.g., 20 seconds at 70°C), continuing to displace upstream strands and create a first phase of redundancy (**Fig 15B**). Denaturation of the amplification product may then occur at a denaturing temperature (e.g., 98°C for 30 seconds) to release the sample nucleic acid and amplification product for additional priming.

[00346] After the first cycle, amplification products have a single 5' tag (**Fig 15C**) comprising the oligonucleotide sequence. These aforementioned steps are repeated to prime the amplification product and sample nucleic acid with the oligonucleotide sequence via its random N-mer. The black sequence indicates portions of the added 5' tags (added in cycle 1) that comprise uracil containing nucleotides and thus, will not be copied upon priming and amplification of the amplification product (**Fig 15D**). Following a second round of amplification, both 5' tagged products and 3' & 5' tagged products may be generated (**Fig 15E**). The 3' & 5' tagged products comprise a full oligonucleotide sequence at one end, the sample nucleic acid sequence, and a sequence partially complementary to the oligonucleotide sequence (e.g., complementary to regions of the oligonucleotide sequence not comprising uracil containing nucleotides) at the other end of the oligonucleotide. The oligonucleotide sequence may hybridize with its partially complementary sequence to generate a partial hairpin structure (**Fig 15F.**

Amplification can continue repeatedly for a desired number of cycles (e.g., up to 20 times), up until all oligonucleotide sequences have been exhausted (**Fig 15G**).

[00347] Partial hairpin formation may prevent generating a copy of a copy and may instead encourage only copies of the original template to be produced, thus reducing potential amplification bias, and other artifacts. Partial hairpin formation may encourage segregation of the desired product and may reduce production of copies.

[00348] Desirable properties for the uracil-non-reading polymerase to form the partial hairpin may include an exonuclease deficient polymerase (e.g., having low exonuclease activity, having substantially no exonuclease activity, having no exonuclease activity), strand displacing capabilities (e.g., a thermostable strand displacing polymerase enzyme), residual activity at temperatures < 50 °C, and discrimination against uracil containing nucleotides v thymine containing nucleotides. Examples of such polymerases may include 9 degrees North, modified (NEB), exo minus Pfu, Deep Vent exo minus, Vent exo minus, and homologs thereof. Moreover, a polymerase with low exonuclease activity may be a polymerase with less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or 0% exonuclease activity of a thermally stable polymerase with normal exonuclease activity (e.g., Taq polymerase). In some cases, a polymerase used for partial hairpin amplification may be capable of strand-displacement. In some cases, limiting the length of the amplified sequence may reduce undesired amplification products, wherein longer length products may include undesired upstream portions such as a barcode sequence. The amplified product length may be limited by inclusion of terminating nucleotides. An example of a terminating nucleotide may include an acyclonucleotide (acyNTPs). Terminating nucleotides may be present at about 0%, 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, 2%, 2.1%, 2.2%, 2.3%, 2.4%, or 2.5% of the amplified product length. In some cases, terminating nucleotides may be present at more than about 0%, 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, 2%, 2.1%, 2.2%, 2.3%, 2.4%, 2.5%, or more of the amplified product length. In some cases, terminating nucleotides may be present at less than about 0%, 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, 2%, 2.1%, 2.2%, 2.3%, 2.4%, or 2.5% of the amplified product length.

[00349] Amplification product length may also be controlled by pre-amplification of sample nucleic acid prior to initiation of PHASE amplification. For example, a random N-mer may be used for pre-amplification of the sample nucleic acid. A random N-mer may be used to prime a sample nucleic acid followed by extension of the primer using suitable thermal cycling conditions. Product length can be controlled by thermal cycling conditions (e.g., number of thermal cycles, temperatures utilized, cycle time, total run time, etc.) in addition to the random priming of the sample nucleic acid. In some cases, pre-amplification products smaller than the original sample nucleic acid can be obtained. Amplification products generated during pre-amplification may then be entered into a PHASE amplification and barcoded as described above.

[00350] As shown in **Fig 17**, addition of a blocking oligonucleotide may reduce start site bias by 50%. Incorporation of uracil containing nucleotides instead of thymine containing nucleotides into the universal sequence and using a polymerase that does not accept or process uracil-containing templates, may significantly reduce sequencing errors, as reported in **Fig 21** and **Fig 22**. For example, Q40 error may be reduced from about 0.002 to about 0.001, unmapped fraction ends may be reduced from about 0.996 to about 0.03, median insert size may be reduced from about 399 to about 310, IQR insert size may be reduced from about 413 to about 209, and zero coverage fraction may be reduced from about 0.9242 to about 0.0093.

[00351] Amplification schemes that do not involve the substitution of thymine containing nucleotides with uracil containing nucleotides are also envisioned for generating partial hairpin species. In some cases, other species unable to be recognized or be copied by a polymerase (e.g., methylated bases, abasic sites, bases linked to bulky side groups, etc.) may be used in place of uracil containing nucleotides to generate partial hairpin amplicons. In some cases, full hairpin amplicons may be generated and processed post-synthesis to generate partial hairpin species. In some cases, full hairpin amplicons may be generated and portions subsequently removed to generate partial hairpin species. For example, as shown in **Fig 34A**, full hairpin amplicons **3401** can be generated via the amplification scheme depicted in **Fig 15** when oligonucleotide primers comprising random N-mers do not comprise uracil containing nucleotides and/or a polymerase capable of accepting or processing a uracil containing template is used for amplification. Upon generation of the full hairpin amplicons **3401**, the full hairpin amplicons can be enzymatically (e.g., via a restriction enzyme or other site specific enzyme such as a nickase) or chemically nicked **3403** at one or more appropriate sites to generate partial hairpin species **3402**.

[00352] In some cases, full hairpin amplicons may be generated and portions added to the full hairpin amplicons to generate partial hairpin species. For example, a primer comprising a sequencing primer binding site (e.g., R1) coupled to a random N-mer and not comprising uracil containing nucleotides may be used to amplify sample nucleic acid and generate full hairpin amplicons (e.g., a full hairpin comprising the sequencing primer binding site (e.g., R1), the copied sample nucleic acid, and the complement to the sequencing primer binding site hybridized with the sequencing primer binding site (e.g., R1c) – **3404** in **Fig 34B**) via the amplification scheme depicted in **Fig 15**. Upon generation of the full hairpin amplicons **3404**, the full hairpin amplicons can have additional sequences (e.g., a sequence comprising a P5 sequence and a barcode sequence) **3405** added, for example, via ligation **3406**.

[00353] In some cases, primers (e.g., oligonucleotides comprising a random N-mer) used to generate full hairpin amplicons may be covalently modified to comprise an additional sequence via, for example, a linker (e.g., a linker not comprising nucleic acid or a linker comprising nucleic acid that does not participate in amplification). In some cases, the linker may be polyethylene glycol or a carbon-based linker. Full hairpin amplicons generated from the primers (e.g., via an amplification scheme depicted in **Fig 15**), thus, can also be covalently linked to the additional sequence via the linker. The attached sequence can then be ligated to the full hairpin amplicon to generate a partial hairpin species. An example of a full hairpin amplicon **3409** comprising an additional sequence **3408** via a linker **3407** is

shown in **Fig 34C**. Following full hairpin generation, the additional sequence **3408** can be ligated to the full hairpin amplicon **3409** such that a partial hairpin species (**3410**) comprising the additional sequence **3408** can be generated.

Targeted N-mers and Targeted Amplification

[00354] In addition to random amplification schemes, barcode constructs (e.g., oligonucleotides comprising a barcode sequence and an N-mer for priming a sample nucleic acid) comprising targeted priming sequences (e.g., a targeted N-mer) and targeted amplification schemes are also envisioned. Targeted amplification schemes may be useful, for example, in detecting a particular gene or sequence of interest via sequencing methods, may be useful in detecting a particular type of nucleic acid, may be useful in detecting the a particular strand of nucleic acid comprising a sequence, and combinations thereof. In general, targeted amplification schemes rely on targeted primers to complete amplification of a particular nucleic acid sequence. In some examples, PCR methods may be used for targeted amplification, via the use of primers targeted toward a particular gene sequence of interest or a particular sequence upstream of a particular gene sequence of interest, such that the particular gene sequence of interest is amplified during PCR.

[00355] The PHASE amplification reaction described above may also be modified such that target amplification of sample nucleic acid is achieved. Barcode constructs comprising a targeted priming sequence (e.g., a targeted N-mer), rather than a random sequence (e.g., a random N-mer), as described above, may be used to prime a specific sequence during PHASE amplification. The specific sequence, for example, may be a particular gene sequence of interest such that generation of amplicons is indicative of the sequence's presence. Or, the specific sequence may be a sequence known to be upstream from a particular gene sequence of interest. Such constructs may be generated, and, if desired, coupled to beads, using any of the methods described herein, including limiting dilution schemes depicted in **Fig 4** and the combinatorial plate schemes described elsewhere herein.

[00356] For example, as described previously with respect to **Fig 4**, a construct comprising a primer **403** (e.g., P5), a barcode sequence **408**, and a read primer binding site (e.g., R1) **415** can be generated (see **Fig 4A-4H**). As shown in **Fig 4I**, an additional sequence **413** can be added (optionally in bulk) to the construct via primer comprising a sequence **412** complementary to read primer binding site **415**. Sequence **413** may serve as a targeted sequence (e.g., a targeted N-mer) such that the targeted sequence corresponds to a particular target sequence of interest. The construct may also comprise an oligonucleotide blocker, as described elsewhere herein, in order to ensure that only the targeted sequence, and not other sequence portions of the construct, primes the sample nucleic acid. Upon entry of the completed construct into a PHASE reaction with sample nucleic acid, for example, the targeted construct may prime the sample nucleic acid (e.g., at the desired sequence site) and the amplification reaction can be initiated to generate partial hairpins from the sample nucleic acid as described above. In some cases, a combination of targeted N-mer primers and random N-mer primers are used to generate partial hairpin

amplicons. In some cases, targeted amplification may be useful in controlling the size (e.g., sequence length) of partial hairpin amplicons that are generated during amplification for a particular target.

[00357] In some cases, a plurality of constructs comprising a barcode sequence and a targeted N-mer may be coupled to a bead (e.g., a gel bead). In some cases, the plurality of constructs may comprise an identical barcode sequence and/or an identical targeted N-mer sequence. In some cases, the targeted N-mer sequence may vary amongst individual constructs of the plurality such that a plurality of target sequences on a sample nucleic acid may be primed via the various targeted N-mers. As described above, the beads may be partitioned (e.g., in fluidic droplets) with sample nucleic acid, the bead(s) in each partition degraded to release the coupled constructs into the partition, and the sample nucleic acid amplified via the targeted N-mer of the constructs. Post processing (e.g., addition of additional sequences (e.g., P7, R2), addition of a sample index, etc.) of the generated amplicons may be achieved with any method described herein, including bulk amplification methods (e.g., bulk PCR) and bulk ligation.

[00358] In a partition, constructs comprising a barcode sequence and a targeted N-mer may be coupled to a bead, may be free in solution (e.g., free in the aqueous interior of a fluidic droplet), or both. Moreover, a partition may comprise both targeted constructs (e.g., constructs comprising a targeted N-mer sequence) and non-targeted constructs (e.g., constructs comprising a random N-mer sequence). Each of the targeted and non-targeted constructs may be coupled to a bead, one of the two may be coupled to a bead, and either construct may also be in solution within a partition.

[00359] Where each type of construct is present in a partition, both targeted and non-targeted amplification of sample nucleic acids may take place. For example, with respect to a PHASE amplification reaction, a targeted barcode construct may be used to initially prime and extend a sample nucleic acid. In general, these steps correspond to the first cycle of PHASE amplification described above with respect to **Figs. 15A-C**, except that the targeted construct is used for initial priming. The extension products can then be primed with a barcode construct comprising a random N-mer such that a partial hairpin is generated, these steps corresponding to the second cycle of PHASE described above with respect to **Figs. 15D-F**. Amplification can continue for additional rounds (e.g., **Fig 15G**) until the desired number of rounds are complete. Post processing (e.g., addition of additional sequences (e.g., P7, R2), addition of a sample index, etc.) of the generated partial hairpin amplicons may be achieved with any method described herein, including bulk amplification methods (e.g., bulk PCR) and bulk ligation.

[00360] Moreover, targeted barcode constructs may be generated such that the construct's targeted N-mer is directed toward nucleic acid species other than DNA, such as, for example, an RNA species. In some cases, the targeted barcode construct's targeted N-mer may be directed toward a particular RNA sequence, such as, for example, a sequence corresponding to transcribed gene or other sequence on a messenger RNA (mRNA) transcript. In some cases, sequencing of barcoded products generated from RNA (e.g., an mRNA) may aid in determining the expression level of a gene transcribed by the RNA. In some cases, the targeted N-mer may be a poly-thymine (e.g., poly-T sequence) sequence capable of hybridizing with a poly-adenine (poly-A sequence) that can, for example, be found at the 3' end of an mRNA transcript. Upon priming of an mRNA with a targeted barcode construct comprising a poly-T sequence via

hybridization of the barcode construct's poly-T sequence with the mRNA's poly-A sequence, the targeted barcode construct can be extended via a reverse transcription reaction to generate a complementary DNA (cDNA) product comprising the barcode construct. In some cases, a targeted barcode construct comprising a poly-T targeted N-mer may also comprise an oligonucleotide blocker as described elsewhere herein, such that only the poly-T sequence hybridizes with RNA.

[00361] Targeted barcode constructs to RNA species may also be useful in generating partial hairpin amplicons via, for example, a PHASE amplification reaction. For example, a targeted barcode construct comprising a poly-T sequence can hybridize with an mRNA via its poly-A sequence. The targeted barcode construct can be extended via a reverse transcription reaction (e.g., via the action of a reverse transcriptase) such that a cDNA comprising the barcode construct is generated. These steps can correspond to the first cycle of PHASE amplification described above with respect to **Figs. 15A-C**, except that reverse transcription is used to generate the extension product. Following reverse transcription (e.g., a first PHASE cycle), a barcode construct comprising a random N-mer may prime the extension products such that a partial hairpin is generated as described above with respect to **Figs. 15D-F**. Amplification can continue for additional rounds (e.g., **Fig 15G**) until the desired number of rounds are complete.

[00362] In some cases, a plurality of targeted constructs comprising a barcode sequence and a targeted N-mer comprising a poly-T sequence may be coupled to a bead (e.g., a gel bead). In some cases, the plurality of constructs may comprise an identical barcode sequence. The beads may be partitioned (e.g., in fluidic droplets) with sample nucleic acid comprising RNA, the bead(s) in each partition degraded to release the coupled constructs into the partition, and the sample RNA captured via the targeted N-mer of the constructs. Partitions may also comprise barcode constructs (e.g., with barcode sequences identical to the targeted constructs) that comprise a random N-mer. In a first amplification cycle, extension of the targeted constructs can occur via reverse transcription within each partition, to generate extension products comprising the targeted construct. The extension products in each partition can then be primed with the barcode constructs comprising the random N-mer to generate partial hairpin amplicons as described above with respect to **Figs. 15-A-G**. Post processing (e.g., addition of additional sequences (e.g., P7, R2), addition of a sample index, etc.) of the generated amplicons may be achieved with any method described herein, including bulk amplification methods (e.g., bulk PCR) and bulk ligation.

[00363] In some cases, reverse transcription of RNA in a sample may also be used without the use of a targeted barcode construct. For example, sample nucleic acid comprising RNA may be first subject to a reverse transcription reaction with other types of reverse transcription primers such that cDNA is generated from the RNA. The cDNA that is generated may then undergo targeted or non-targeted amplification as described herein. For example, sample nucleic acid comprising RNA may be subject to a reverse transcription reaction such that cDNA is generated from the RNA. The cDNA may then enter a PHASE amplification reaction, using a barcode construct with a random N-mer as described above with respect to **Figs 15A-G**, to generate partial hairpin amplicons comprising the construct's barcode sequence. Post processing (e.g., addition of additional sequences (e.g., P7, R2), addition of a sample index, etc.) of

the generated partial hairpin amplicons may be achieved with any method described herein, including bulk amplification methods (e.g., bulk PCR) and bulk ligation.

[00364] Targeted barcode constructs may also be generated toward specific sequences (e.g., gene sequences) on specific strands of a nucleic acid such that strandedness information is retained for sequencer-ready products generated for each strand. For example, a sample nucleic may comprise double stranded nucleic acid (e.g., double-stranded DNA), such that each strand of nucleic acid comprises one or more different target gene sequences. Complementary DNA strands can comprise different gene sequences due to the opposite 5' to 3' directionalities and/or base composition of each strand. Targeted barcode constructs can be generated for each strand (based on 5' to 3' directionality of the strand) based on the targeted N-mer and configuration of the barcode construct. Example sets of targeted barcode constructs directed to forward and reverse strands of a double-stranded sample nucleic acid are shown in **Fig 28A**.

[00365] Example sets **2801** and **2802** of targeted barcode constructs each targeted to either of a forward (**2801**) strand and reverse (**2802**) strand of a double-stranded sample nucleic acid are shown in **Fig 28A**. Set **2801** comprises targeted barcode constructs **2803** and **2804** comprising a P5 sequence, a barcode sequence, and a targeted N-mer to either of a first target sequence (**2803**) or a second target sequence (**2804**). Set **2802** comprises targeted barcode constructs **2805** and **2806** comprising a P5 sequence, a barcode sequence, and a targeted N-mer to either of the first target sequence (**2805**) and the second target sequence (**2806**). Each construct can also comprise any additional sequences between the barcode and the targeted N-mer (indicated by an arrow in each construct shown in **Fig 28A**).

[00366] The barcode constructs in set **2801** are configured to prime their respective target sequences on the forward strand of the double-stranded sample nucleic acid. The barcode constructs of set **2802** are configured to prime their respective target sequences on the reverse strand of the double-stranded sample nucleic acid. As shown, the targeted barcode constructs in each set are configured in opposite directionality corresponding to the opposite directionality of forward and reverse strands of the double-stranded sample nucleic acid. Each barcode construct can prime its respective target sequence on its respective strand of sample nucleic acid to generate barcoded amplicons via an amplification reaction, such as any amplification reaction described herein.

[00367] Additional sequences can be added to barcoded amplicons using amplification methods described herein, including bulk amplification, bulk ligation, or a combination thereof. Example sets of primers that may be used to add a sample index and P7 sequence to amplicons generated from the targeted barcode constructs in **Fig 28A** are shown in **Fig 28B**. Primer set **2808** corresponds to targeted barcode construct set **2801** (e.g., targeted barcode construct **2803** corresponds to primer **2811**, targeted barcode construct **2804** corresponds to primer **2812**) and primer set **2808** corresponds to targeted barcode construct set **2801** (e.g., targeted barcode construct **2505** corresponds to primer **2809**, targeted barcode construct **2806** corresponds to primer **2810**). Each primer can prime its respective target sequence on its respective strand and bulk amplification (e.g., bulk PCR) initiated to generate sequencer-ready constructs that include the P7 and sample index sequences in analogous fashion to bulk amplification methods described

elsewhere herein. Based on the configuration and directionality of the various components of each sequencer-ready construct (e.g., P5, barcode, targeted N-mer, sample insert, etc.), the strand from which the sequencer-ready product is generated can be determined/is retained.

[00368] Libraries of barcode constructs (e.g., targeted barcode constructs) may be generated for both forward and reverse strands of a double stranded nucleic acid. For example, two libraries of beads (e.g., gel beads) comprising targeted barcode constructs may be generated using methods described herein, such that one library comprises targeted barcode constructs for forward strands of sample nucleic acids and the other library comprises targeted barcode constructs for reverse strands of sample nucleic acids. In some cases, each library may comprise beads each comprising an identical targeted N-mer. In some cases, each library may comprise two or more sets of beads, with each bead in a set comprising an identical targeted N-mer (e.g., a targeted N-mer targeted toward a particular gene) and different sets comprising different targeted N-mers. In some cases, the two libraries may be combined such that a library of forward strand and reverse strand beads is generated.

[00369] For example, a library can comprise two types of forward strand beads and two types of reverse strand beads, for a total of four types of beads. Each bead in the library may comprise a unique barcode sequence. One type of the forward strand beads and one type of the reverse strand beads may comprise targeted N-mers corresponding to a target sequence (e.g., a target gene sequence). For example, one type of forward strand beads may comprise a targeted barcode construct as shown in **2803** in **Fig 28A** and one type of reverse strand beads may comprise a targeted barcode construct as shown in **2805** in **Fig 28A**. Analogously, the second type of forward strand beads may comprise a targeted barcode construct as shown in **2804** in **Fig 28A** and one type of reverse strand beads may comprise a targeted barcode construct as shown in **2806** in **Fig 28A**.

[00370] A barcode library comprising forward strand and reverse strand beads (e.g., gel beads), with each bead comprising a unique barcode sequence may be partitioned to barcode sample nucleic acids as described elsewhere herein. For example, the mixed library of two types of forward strand and two types of reverse strand beads described above may be partitioned with a sample nucleic acid (e.g., genomic DNA) and any other desired reagents (e.g., reagents necessary for amplification of the sample nucleic acid, a reducing agent). The partitions may be, for example, fluidic droplets such as droplets of an emulsion. In general, each partition may comprise a bead (e.g., a forward strand bead or a reverse strand bead) coupled to a targeted barcode construct comprising a unique barcode sequence and a targeted N-mer. In some cases, though, one or more of the partitions may comprise multiple beads of the same type or of different types. The targeted barcode constructs may be released from the bead (e.g., via degradation of the bead – for example, via a reducing agent in cases where the bead is a gel bead comprising disulfide bonds) in the partition and allowed to prime their target sequence on their respective strand (e.g., forward strand or reverse strand) of sample nucleic acid.

[00371] A first product strand synthesis may take place in each partition via extension of the hybridized targeted barcode construct, via, for example, linear amplification of the sample nucleic acid. Additional rounds of linear amplification of the sample nucleic acid with the targeted barcode construct, for example,

may be used to generate additional copies of the first product strand. First product strands may then be removed from the partitions (e.g., in cases where the partitions are droplets of an emulsion, the emulsion may be broken to release first products) and pooled. The first products may be washed to remove targeted barcode constructs and any other waste products. In some cases, an optional double-stranded digestion may be completed to digest sample nucleic acid and remove it from the first product strands.

[00372] Next, the first product strands may be subject to bulk amplification to add additional sequences (e.g., P7, a sample index, etc.) to the first product strands, resulting in the generation of second product strands. The bulk amplification reaction mixture may comprise a plurality of primers, with each primer in the plurality corresponding to one of the bead types (and, thus, type of targeted barcode construct) used to generate the first products strands. For the example library comprising two types of forward strand beads and two types of reverse strand beads described above, primers shown as **2809**, **2810**, **2811**, and **2812** in **Fig 28B** may be used to add additional sample index and P7 sequences to first product strands generated from targeted barcode constructs **2803**, **2804**, **2805**, and **2806** respectively via bulk amplification. Second product strands may then be washed to remove primers from the reaction mixture. Fresh primers (e.g., primers comprising P5 and P7 for the example described above) may then be added one or more additional rounds of amplification (e.g., via PCR) to generate final, sequencer-ready products. Thus, final products can comprise the original targeted barcode construct, the strand of sample nucleic acid amplified, and the additional sequences (e.g., P7, sample index) added to first product strands.

[00373] Methods described herein may be useful in whole genome amplification. In some embodiments of whole genome amplification, a random primer (e.g., a random N-mer sequence) can be hybridized to a genomic nucleic acid. The random primer can be a component of a larger oligonucleotide that may also include a universal nucleic acid sequence (including any type of universal nucleic acid sequence described herein) and a nucleic acid barcode sequence. In some cases, the universal nucleic acid sequence may comprise one or more uracil containing nucleotides. Moreover, in some cases, the universal nucleic acid sequence may comprise a segment of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more nucleotides that do not comprise uracil. The random primer can be extended (e.g., in a primer extension reaction or any other suitable type of nucleic acid amplification reaction) to form an amplified product.

[00374] As described elsewhere herein, the amplified product may undergo an intramolecular hybridization reaction to form a hairpin molecule such as, for example, a partial hairpin molecule. In some cases, whole genome amplification may occur in the presence of an oligonucleotide blocker (also referred to as a blocker oligonucleotide elsewhere herein) that may or may not comprise a blocker moiety (e.g., C3 spacer (/3SpC3/), Dideoxy-C (/3ddC/), 3' phosphate, or any other type of blocker moiety described elsewhere herein). Furthermore, the oligonucleotide blocker may be capable of hybridizing to at least a portion of the universal nucleic acid sequence or any other part of an oligonucleotide comprising the random primer.

[00375] In some embodiments of whole genome amplification, a genomic component (e.g., a chromosome, genomic nucleic acid such as genomic DNA, a whole genome of an organism, or any other

type of genomic component described herein) may be fragmented in a plurality of first fragments. The first fragments can be co-partitioned into a plurality of partitions with a plurality of oligonucleotides. The oligonucleotides in each of the partitions may comprise a primer sequence (including a type of primer sequence described elsewhere herein) and a common sequence (e.g., a barcode sequence). Primer sequences in each partition can then be annealed to a plurality of different regions of the first fragments within each partition. The primer sequences can then be extended along the first fragments to produce amplified first fragments within each partition of the plurality of partitions. The amplified first fragments within the partitions may comprise any suitable coverage (as described elsewhere herein) of the genomic component. In some cases, the amplified first fragments within the partitions may comprise at least 1X coverage, at least 2Xcoverage, at least 5X coverage, at least 10X coverage, at least 20X coverage, at least 40X coverage, or greater coverage of the genomic component.

VII. Digital Processor

[00376] The methods, compositions, devices, and kits of this disclosure may be used with any suitable processor, digital processor or computer. The digital processor may be programmed, for example, to operate any component of a device and/or execute methods described herein. The digital processor may be capable of transmitting or receiving electronic signals through a computer network, such as for example, the Internet and/or communicating with a remote computer. One or more peripheral devices such as screen display, printer, memory, data storage, and/or electronic display adaptors may be in communication with the digital processor. One or more input devices such as keyboard, mouse, or joystick may be in communication with the digital processor. The digital processor may also communicate with detector such that the detector performs measurements at desired or otherwise predetermined time points or at time points determined from feedback received from pre-processing unit or other devices.

[00377] A conceptual schematic for an example control assembly is shown in **Fig 18**. A computer, serves as the central hub for control assembly. The computer is in communication with a display, one or more input devices (*e.g.*, a mouse, keyboard, camera, etc.), and optionally a printer. The control assembly, via its computer, is in communication with one or more devices: optionally a sample pre-processing unit, one or more sample processing units (such as a sequence, thermocycler, or microfluidic device), and optionally a detector. The control assembly may be networked, for example, via an Ethernet connection. A user may provide inputs (*e.g.*, the parameters necessary for a desired set of nucleic acid amplification reactions or flow rates for a microfluidic device) into the computer, using an input device. The inputs are interpreted by the computer, to generate instructions. The computer communicates such instructions to the optional sample pre-processing unit, the one or more sample processing units, and/or the optional detector for execution.

[00378] Moreover, during operation of the optional sample pre-processing unit, one or more sample processing units, and/or the optional detector, each device may communicate signals back to computer. Such signals may be interpreted and used by computer to determine if any of the devices require further

instruction. The computer may also modulate the sample pre-processing unit such that the components of a sample are mixed appropriately and fed, at a desired or otherwise predetermined rate, into the sample processing unit (such as the microfluidic device).

[00379] The computer may also communicate with a detector such that the detector performs measurements at desired or otherwise predetermined time points or at time points determined from feedback received from pre-processing unit or sample processing unit. The detector may also communicate raw data obtained during measurements back to the computer for further analysis and interpretation.

[00380] Analysis may be summarized in formats useful to an end user via a display and/or printouts generated by a printer. Instructions or programs used to control the sample pre-processing unit, the sample processing unit, and/or the detector; data acquired by executing any of the methods described herein; or data analyzed and/or interpreted may be transmitted to or received from one or more remote computers, via a network, which, for example, could be the Internet.

[00381] In some embodiments, the method of bead formation may be executed with the aid of a digital processor in communication with a droplet generator. The digital processor may control the speed at which droplets are formed or control the total number of droplets that are generated. In some embodiments, the method of attaching samples to barcoded beads may be executed with the aid of a digital processor in communication with the microfluidic device. Specifically, the digital processor may control the volumetric amount of sample and/or beads injected into the input channels and may also control the flow rates within the channels. In some embodiments, the method of attaching oligonucleotides, primers, and the like may be executed with the aid of a digital processor in communication with a thermocycler or other programmable heating element. Specifically, the digital processor may control the time and temperature of cycles during ligation or amplification. In some embodiments, the method of sequencing a sample may be executed with the aid of a digital processor in communication with a sequencing device.

VIII. Kits

[00382] In some cases, this disclosure provides a kit comprising a microfluidic device, a plurality of barcoded beads, and instructions for utilizing the microfluidic device and combining barcoded beads with customer sample to create fluidic droplets containing both. As specified throughout this disclosure, any suitable sample may be incorporated into the fluidic droplets. As described throughout this disclosure, a bead may be designed to be degradable or non-degradable. In this case, the kit may or may not include a reducing agent for bead degradation.

[00383] In some cases, this disclosure provides a kit comprising a plurality of barcoded beads, suitable amplification reagents, e.g., optionally including one or more of polymerase enzymes, nucleoside triphosphates or their analogues, primer sequences, buffers, and the like, and instructions for combining barcoded beads with customer sample. As specified throughout this disclosure, any suitable sample may be used. As specified throughout this disclosure, the amplification reagents may include a polymerase that

will not accept or process uracil-containing templates. A kit of this disclosure may also provide agents to form an emulsion, including an oil and surfactant.

IX. Applications

Barcoding Sample Materials

[00384] The methods, compositions and systems described herein are particularly useful for attaching barcodes, and particularly barcode nucleic acid sequences, to sample materials and components of those sample materials. In general, this is accomplished by partitioning sample material components into separate partitions or reaction volumes in which are co-partitioned a plurality of barcodes, which are then attached to sample components within the same partition.

[00385] In an exemplary process, a first partition is provided that includes a plurality of oligonucleotides (e.g., nucleic acid barcode molecules) that each comprise a common nucleic acid barcode sequence. The first partition may comprise any of a variety of portable partitions, e.g., a bead (e.g., a degradable bead, a gel bead), a droplet (e.g., an aqueous droplet in an emulsion), a microcapsule, or the like, to which the oligonucleotides are releasably attached, releasably coupled, or are releasably associated. Moreover, any suitable number of oligonucleotides may be included in the first partition, including numbers of oligonucleotides per partition described elsewhere herein. For example, the oligonucleotides may be releasably attached to, releasably coupled to, or releasably associated with the first partition via a cleavable linkage such as, for example, a chemically cleavable linkage (e.g., a disulfide linkage, or any other type of chemically cleavable linkage described herein), a photocleavable linkage, and/or a thermally cleavable linkage. In some cases, the first partition may be a bead and the bead may be a degradable bead (e.g., a photodegradable bead, a chemically degradable bead, a thermally degradable bead, or any other type of degradable bead described elsewhere herein). Moreover, the bead may comprise chemically-cleavable cross-linking (e.g., disulfide cross-linking) as described elsewhere herein.

[00386] The first partition is then co-partitioned into a second partition, with a sample material, sample material component, fragment of a sample material, or a fragment of a sample material component. The sample material (or component or fragment thereof) may be any appropriate sample type, including the example sample types described elsewhere herein. In cases where a sample material or component of a sample material comprises one or more nucleic acid fragments, the one or more nucleic acid fragments may be of any suitable length, including, for example, nucleic acid fragment lengths described elsewhere herein. The second partition may include any of a variety of partitions, including for example, wells, microwells, nanowells, tubes or containers, or in preferred cases droplets (e.g., aqueous droplets in an emulsion) or microcapsules in which the first partition may be co-partitioned. In some cases, the first partition may be provided in a first aqueous fluid and the sample material, sample material component, or fragment of a sample material component may be provided in a second aqueous fluid. During co-partitioning, the first aqueous fluid and second aqueous fluid may be combined within a droplet within an immiscible fluid. In some cases, the second partition may comprise no more than one first partition. In

other cases, the second partition may comprise no more than one, two, three, four, five, six, seven, eight, nine, or ten first partitions. In other cases, the second partition may comprise at least one, two, three, four, five, six, seven, eight, nine, ten, or more first partitions.

[00387] Once co-partitioned, the oligonucleotides comprising the barcode sequences may be released from the first partition (e.g., via degradation of the first partition, cleaving a chemical linkgage between the oligonucleotides and the first partition, or any other suitable type of release, including types of release described elsewhere herein) into the second partition, and attached to the sample components co-partitioned therewith. In some cases, the first partition may comprise a bead and the crosslinking of the bead may comprise a disulfide linkage. In addition, or as an alternative, the oligonucleotides may be linked to the bead via a disulfide linkage. In either case, the oligonucleotides may be released from the first partition by exposing the first partition to a reducing agent (e.g., DTT, TCEP, or any other exemplary reducing agent described elsewhere herein).

[00388] As noted elsewhere herein, attachment of the barcodes to sample components includes the direct attachment of the barcode oligonucleotides to sample materials, e.g. through ligation, hybridization, or other associations. Additionally, in many cases, for example, in barcoding of nucleic acid sample materials (e.g., template nucleic acid sequences, template nucleic acid molecules), components or fragments thereof, such attachment may additionally comprise use of the barcode containing oligonucleotides that also comprise as priming sequences. The priming sequence can be complementary to at least a portion of a nucleic acid sample material and can be extended along the nucleic acid sample materials to create complements to such sample materials, as well as at least partial amplification products of those sequences or their complements.

[00389] In another exemplary process, a plurality of first partitions can be provided that comprise a plurality of different nucleic acid barcode sequences. Each of the first partitions can comprise a plurality of nucleic acid barcode molecules having the same nucleic acid barcode sequence associated therewith. Any suitable number of nucleic acid barcode molecules may be associated with each of the first partitions, including numbers of nucleic acid barcode molecules per partition described elsewhere herein. The first partitions may comprise any suitable number of different nucleic acid barcode sequences, including, for example, at least about 2, 10, 100, 500, 1000, 5000, 10000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, or 1000000000, or more different nucleic acid barcode sequences.

[00390] In some cases, the plurality of first partitions may comprise a plurality of different first partitions where each of the different first partitions comprises a plurality of releasably attached, releasably coupled, or releasably associated oligonucleotides comprising a common barcode sequence, with the oligonucleotides associated with each different first partitions comprising a different barcode sequence. The number of different first partitions may be, for example, at least about 2, 10, 100, 500, 1000, 5000, 10000, 50000, 100000, 500000, 1000000, 5000000, 10000000, 50000000, or 1000000000, or more different first partitions.

[00391] The first partitions may be co-partitioned with sample materials, fragments of a sample material, components of a sample material, or fragments of a component(s) of a sample material into a plurality of

second partitions. In some cases, a subset of the second partitions may comprise the same nucleic acid barcode sequence. For example, at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more of the second partitions may comprise the same nucleic acid barcode sequence. Moreover, the distribution of first partitions per second partition may also vary according to, for example, occupancy rates described elsewhere herein. In cases where the plurality of first partitions comprises a plurality of different first partitions, each different first partition may be disposed within a separate second partition.

[00392] Following co-partitioning, the nucleic acid barcode molecules associated with the first partitions can be released into the plurality of second partitions. The released nucleic acid barcode molecules can then be attached to the sample materials, sample material components, fragments of a sample material, or fragments of sample material components, within the second partitions. In the case of barcoded nucleic acid species (e.g., barcoded sample nucleic acid, barcoded template nucleic acid, barcoded fragments of one or more template nucleic acid sequences, etc.), the barcoded nucleic acid species may be sequenced as described elsewhere herein.

[00393] In another exemplary process, an activatable nucleic acid barcode sequence may be provided and partitioned with one or more sample materials, components of a sample material, fragments of a sample material, or fragments of a component(s) of a sample material into a first partition. With the first partition, the activatable nucleic acid barcode sequence may be activated to produce an active nucleic acid barcode sequence. The active nucleic acid barcode sequence can then be attached to the one or more sample materials, components of a sample material, fragments of a sample material, or fragments of a component(s) of a sample material.

[00394] In some cases, the activatable nucleic acid barcode sequence may be coupled to a second partition that is also partitioned in the first partition with the activatable nucleic acid barcode sequence. As described elsewhere herein, an activatable nucleic acid barcode sequence may be activated by releasing the activatable nucleic acid barcode sequence from an associated partition (e.g., a bead). Thus, in cases where an activatable nucleic acid barcode sequence is associated with a second partition (e.g., a bead) that is partitioned in a first partition (e.g., a fluidic droplet), the activatable nucleic acid barcode sequence may be activated by releasing the activatable nucleic acid barcode sequence from its associated second partition. In addition, or as an alternative, an activatable barcode may also be activated by removing a removable blocking or protecting group from the activatable nucleic acid barcode sequence.

[00395] In another exemplary process, a sample of nucleic acids may be combined with a library of barcoded beads (including types of beads described elsewhere herein) to form a mixture. In some cases, the barcodes of the beads may, in addition to a barcode sequence, each comprise one or more additional sequences such as, for example, a universal sequence and/or a functional sequence (e.g., a random N-mer or a targeted N-mer, as described elsewhere herein). The mixture may be partitioned into a plurality of partitions, with at least a subset of the partitions comprising at most one barcoded bead. Within the partitions, the barcodes may be released from the beads, using any suitable route, including types of release described herein. A library of barcoded beads may be generated via any suitable route, including

the use of methods and compositions described elsewhere herein. In some cases, the sample of nucleic acids may be combined with the library of barcoded beads and/or the resulting mixture partitioned with the aid of a microfluidic device, as described elsewhere herein. In cases where the released barcodes also comprise a primer sequence (e.g., such as a targeted N-mer or a random N-mer as described elsewhere herein), the primer sequences of the barcodes may be hybridize with the sample nucleic acids and, if desired, an amplification reaction can be completed in the partitions.

Polynucleotide Sequencing

[00396] Generally, the methods and compositions provided herein are useful for preparation of oligonucleotide fragments for downstream applications such as sequencing. In particular, these methods, compositions and systems are useful in the preparation of sequencing libraries. Sequencing may be performed by any available technique. For example, sequencing may be performed by the classic Sanger sequencing method. Sequencing methods may also include: high-throughput sequencing, pyrosequencing, sequencing-by-ligation, sequencing by synthesis, sequencing-by-hybridization, RNA-Seq (Illumina), Digital Gene Expression (Helicos), next generation sequencing, single molecule sequencing by synthesis (SMSS) (Helicos), massively-parallel sequencing, clonal single molecule Array (Solexa), shotgun sequencing, Maxim-Gilbert sequencing, primer walking, and any other sequencing methods known in the art.

[00397] For example, a plurality of target nucleic acid sequences may be sequenced by providing a plurality of target nucleic sequences and separating the target nucleic acid sequences into a plurality of separate partitions. Each of the separate partitions can comprise one or more target nucleic acid sequences and a plurality of oligonucleotides. The separate partitions may comprise any suitable number of different barcode sequences (e.g., at least 1,000 different barcode sequences, at least 10,000 different barcode sequences, at least 100,000 different barcode sequences, at least 1,000,000 different barcode sequences, at least 10,000,000 different barcode sequences, or any other number of different barcode sequences as described elsewhere herein). Moreover, the oligonucleotides in a given partition can comprise a common barcode sequence. The oligonucleotides and associated common barcode sequence in a given partition can be attached to fragments of the one or more target nucleic acids or to copies of portions of the target nucleic acid sequences within the given partition. Following attachment, the separate partitions can then be pooled. The fragments of the target nucleic acids or the copies of the portions of the target nucleic acids and attached barcode sequences can then be sequenced.

[00398] In another example, a plurality of target nucleic acid sequences may be sequenced by providing the target nucleic acid sequences and separating them into a plurality of separate partitions. Each partition of the plurality of separate partitions can include one or more of the target nucleic acid sequences and a bead having a plurality of attached oligonucleotides. The oligonucleotides attached to a given bead may comprise a common barcode sequence. The oligonucleotides associated with a bead can be attached to fragments of the target nucleic acid sequences or to copies of portions of the target nucleic acid sequences within a given partition, such that the fragments or copies of the given partition are also

attached to the common barcode sequence associated with the bead. Following attachment of the oligonucleotides to the fragments of the target nucleic acid sequences or the copies of the portions of the target nucleic acid sequences, the separate partitions can then be pooled. The fragments of the target nucleic acid sequences or the copies of the portions of the target nucleic acid sequences and any attached barcode sequences can then be sequenced (e.g., using any suitable sequencing method, including those described elsewhere herein) to provide barcoded fragment sequences or barcoded copy sequences. The barcoded fragment sequences or barcoded copy sequences can be assembled into one or more contiguous nucleic acid sequence based, in part, upon a barcode portion of the barcoded fragment sequences or barcoded copy sequences.

[00399] In some cases, varying numbers of barcoded-oligonucleotides are sequenced. For example, in some cases about 30%-90% of the barcoded-oligonucleotides are sequenced. In some cases, about 35%-85%, 40%-80%, 45%-75%, 55%-65%, or 50%-60% of the barcoded-oligonucleotides s are sequenced. In some cases, at least about 30%, 40%, 50%, 60%, 70%, 80%, or 90% of barcoded-oligonucleotides are sequenced. In some cases, less than about 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the barcoded-oligonucleotides are sequenced.

[00400] In some cases, sequences from fragments are assembled to provide sequence information for a contiguous region of the original target polynucleotide that may be longer than the individual sequence reads. Individual sequence reads may be about 10-50, 50-100, 100-200, 200-300, 300-400, or more nucleotides in length. Examples of sequence assembly methods include those set forth in U.S. Provisional Patent Application No. _____ (Attorney Docket No. 43487-729.101), filed of even date herewith.

[00401] The identities of the barcodes may serve to order the sequence reads from individual fragments as well as to differentiate between haplotypes. For example, when combining individual sample fragments and barcoded beads within fluidic droplets, parental polynucleotide fragments may be separated into different droplets. With an increase in the number of fluidic droplets and beads within a droplet, the likelihood of a fragment from both a maternal and paternal haplotype contained within the same fluidic droplet associated with the same bead may become negligibly small. Thus, sequence reads from fragments in the same fluidic droplet and associated with the same bead may be assembled and ordered.

[00402] In at least one example, the present disclosure provides nucleic acid sequencing methods, systems compositions, and combinations of these that are useful in providing myriad benefits in both sequence assembly and read-length equivalent, but do so with very high throughput and reduced sample preparation time and cost.

[00403] In general, the sequencing methods described herein provide for the localized tagging or barcoding of fragments of genetic sequences. By tagging fragments that derive from the same location within a larger genetic sequence, one can utilize the presence of the tag or barcode to inform the assembly process as alluded to above. In addition, the methods described herein can be used to generate and barcode shorter fragments from a single, long nucleic acid molecule. Sequencing and assembly of these

shorter fragments provides a long read equivalent sequence, but without the need for low throughput longer read-length sequencing technologies.

[00404] **Fig 39** provides a schematic illustration of an example sequencing method. As shown, a first genetic component **3902** that may comprise, for example, a chromosome or other large nucleic acid molecule, is fragmented into a set of large first nucleic acid fragments, e.g., including fragments **3904** and **3906**. The fragments of the large genetic component may be non-overlapping or overlapping, and in some cases, may include multifold overlapping fragments, in order to provide for high confidence assembly of the sequence of the larger component. In some cases, the fragments of the larger genetic component provide 1X, 2X, 5X, 10X, 20X, 40X or greater coverage of the larger component.

[00405] One or more of the first fragments **3904** is then processed to separately provide overlapping set of second fragments of the first fragment(s), e.g., second fragment sets **3908** and **3910**. This processing also provides the second fragments with a barcode sequence that is the same for each of the second fragments derived from a particular first fragment. As shown, the barcode sequence for second fragment set **3908** is denoted by "1" while the barcode sequence for fragment set **3910** is denoted by "2". A diverse library of barcodes may be used to differentially barcode large numbers of different fragment sets. However, it is not necessary for every second fragment set from a different first fragment to be barcoded with different barcode sequences. In fact, in many cases, multiple different first fragments may be processed concurrently to include the same barcode sequence. Diverse barcode libraries are described in detail elsewhere herein.

[00406] The barcoded fragments, e.g., from fragment sets **3908** and **3910**, may then be pooled for sequencing. Once sequenced, the sequence reads **3912** can be attributed to their respective fragment set, e.g., as shown in aggregated reads **3914** and **3916**, at least in part based upon the included barcodes, and optionally, and preferably, in part based upon the sequence of the fragment itself. The attributed sequence reads for each fragment set are then assembled to provide the assembled sequence for the first fragments, e.g., fragment sequences **3918** and **3920**, which in turn, may be assembled into the sequence **3922** of the larger genetic component.

[00407] In accordance with the foregoing, a large genetic component, such as a long nucleic acid fragment, e.g., 1, 10, 20, 40, 50, 75, 100, 1000 or more kb in length, a chromosomal fragment or whole chromosome, or part of or an entire genome (e.g., genomic DNA) is fragmented into smaller first fragments. Typically, these fragments may be anywhere from about 1000 to about 100000 bases in length. In certain preferred aspects, the fragments will be between about 1 kb and about 100 kb, or between about 5 kb and about 50 kb, or from about 10kb to about 30kb, and in some cases, between about 15 kb and about 25 kb. Fragmentation of these larger genetic components may be carried out by any of a variety of convenient available processes, including commercially available shear based fragmenting systems, e.g., Covaris fragmentation systems, size targeted fragmentation systems, e.g., Blue Pippin (Sage Sciences), enzymatic fragmentation processes, e.g., using restriction endonucleases, or the like. As noted above, the first fragments of the larger genetic component may comprise overlapping or non-overlapping first fragments. Although described here as being fragmented prior to partitioning, it will be appreciated

that fragmentation may optionally and/or additionally be performed later in the process, e.g., following one or more amplification steps, to yield fragments of a desired size for sequencing applications.

[00408] In preferred aspects, the first fragments are generated from multiple copies of the larger genetic component or portions thereof, so that overlapping first fragments are produced. In preferred aspects, the overlapping fragments will constitute greater than 1X coverage, greater than 2X coverage, greater than 5X coverage, greater than 10X coverage, greater than 20X coverage, greater than 40 X coverage, or even greater coverage of the underlying larger genetic component or portion thereof. The first fragments are then segregated to different reaction volumes. In some cases, the first fragments may be separated so that reaction volumes contain one or fewer first fragments. This is typically accomplished by providing the fragments in a limiting dilution in solution, such that allocation of the solution to different reaction volumes results in a very low probability of more than one fragment being deposited into a given reaction volume. However, in most cases, a given reaction volume may include multiple different first fragments, and can even have 2, 5, 10, 100, 100 or even up to 10,000 or more different first fragments in a given reaction volume. Again, achieving a desired range of fragment numbers within individual reaction volumes is typically accomplished through the appropriate dilution of the solution from which the first fragments originate, based upon an understanding of the concentration of nucleic acids in that starting material.

[00409] The reaction volumes may include any of variety of different types of vessels or partitions. For example, the reaction volumes may include conventional reaction vessels, such as test tubes, reaction wells, microwells, nanowells, or they may include less conventional reaction volumes, such as droplets within a stabilized emulsion, e.g., a water in oil emulsion system. In preferred aspects, droplets are preferred as the reaction volumes for their extremely high multiplex capability, e.g., allowing the use of hundreds of thousands, millions, tens of millions or even more discrete droplet/reaction volumes within a single container. Within each reaction volume, the fragments that are contained therein are then subjected to processing that both derives sets of overlapping second fragments of each of the first fragments, and also provides these second fragments with attached barcode sequences. As will be appreciated, in preferred aspects, the first fragments are partitioned into droplets that also contain one or more microcapsules or beads that include the members of the barcode library used to generate and barcode the second fragments.

[00410] In preferred aspects, the generation of these second fragments is carried out through the introduction of primer sequences that include the barcode sequences and that are capable of hybridizing to portions of the first fragment and be extended along the first fragment to provide a second fragment including the barcode sequence. These primers may comprise targeted primer sequences, e.g., to derive fragments that overlap specific portions of the first fragment, or they may comprise universal priming sequences, e.g., random primers, that will prime multiple different regions of the first fragments to create large and diverse sets of second fragments that span the first fragment and provide multifold overlapping coverage. These extended primer sequences may be used as the second fragments, or they may be further replicated or amplified. For example, iterative priming against the extended sequences, e.g., using the

same primer containing barcoded oligonucleotides. In certain preferred aspects, the generation of the second sets of fragments generates the partial hairpin replicates of portions of the first fragment, as described elsewhere herein that each include barcode sequences, e.g., for PHASE amplification as described herein. As noted elsewhere herein, the formation of the partial hairpin is generally desired to prevent repriming of the replicated strand, e.g., making a copy of a copy. As such, the partial hairpin is typically preferentially formed from the amplification product during annealing as compared to a primer annealing to the amplification product, e.g., the hairpin will have a higher Tm than the primer product pair.

[00411] The second fragments are generally selected to be of a length that is suitable for subsequent sequencing. For short read sequencing technologies, such fragments will typically be from about 50 bases to about 1000 bases in sequenceable length, from about 50 bases to about 900 bases in sequenceable length, from about 50 bases to about 800 bases in sequenceable length, from about 50 bases to about 700 bases in sequenceable length, from about 50 bases to about 600 bases in sequenceable length, from about 50 bases to about 500 bases in sequenceable length, from about 50 bases to about 400 bases in sequenceable length, from about 50 bases to about 300 bases in sequenceable length, from about 50 bases to about 250 bases in sequenceable length, from about 50 bases to about 200 bases in sequenceable length, or from about 50 bases to about 100 bases in sequenceable length, including the barcode sequence segments, and functional sequences that are subjected to the sequencing process.

[00412] Once the overlapping, barcoded second fragment sets are generated, they may be pooled for subsequent processing and ultimately, sequencing. For example, in some cases, the barcoded fragments may be subsequently subjected to additional amplification, e.g., PCR amplification, as described elsewhere herein. Likewise, these fragments may additionally, or concurrently, be provided with sample index sequences to identify the sample from which collections of barcoded fragments have derived, as well as providing additional functional sequences for use in sequencing processes.

[00413] In addition, clean up steps may also optionally be performed, e.g., to purify nucleic acid components from other impurities, to size select fragment sets for sequencing, or the like. Such clean up steps may include purification and/or size selection upon SPRI beads (such as Ampure® beads, available from Beckman Coulter, Inc.). In some cases, multiple process steps may be carried out in an integrated process while the fragments are associated with SPRI beads, e.g., as described in Fisher et al., Genome Biol. 2011:12(1):R1 (E-pub Jan 4, 2011), which is incorporated herein by reference in its entirety for all purposes.

[00414] As noted previously, in many cases, short read sequencing technologies are used to provide the sequence information for the second fragment sets. Accordingly, in preferred aspects, second fragment sets will typically comprise fragments that, when including the barcode sequences, will be within the read length of the sequencing system used. For example, for Illumina HiSeq® sequencing, such fragments may be between generally range from about 100 bases to about 200 bases in length, when carrying out paired end sequencing. In some cases, longer second fragments may be sequenced when accessing only the terminal portions of the fragments by the sequencing process.

[00415] As noted above with reference to **Fig 39**, the sequence reads for the various second fragments are then attributed to their respective starting nucleic acid segment based in part upon the presence of a particular barcode sequence, and in some cases, based in part on the actual sequence of the fragment, i.e., a non-barcode portion of the fragment sequence. As will be appreciated, despite being based upon short sequence data, one can infer that two sequences sharing the same barcode likely originated from the same longer first fragment sequence, especially where such sequences are otherwise assemble-able into a contiguous sequence segment, e.g., using other overlapping sequences bearing the common barcode. Once the first fragments are assembled, they may be assembled into larger sequence segments, e.g., the full length genetic component.

[00416] In one exemplary process, one or more fragments of one or more template nucleic acid sequences may be barcoded using a method described herein. A fragment of the one or more fragments may be characterized based at least in part upon a nucleic acid barcode sequence attached thereto. Characterization of the fragment may also include mapping the fragment to its respective template nucleic acid sequence or a genome from which the template nucleic acid sequence was derived. Moreover, characterization may also include identifying an individual nucleic acid barcode sequence and a sequence of a fragment of a template nucleic acid sequence attached thereto.

[00417] In some cases, sequencing methods described herein may be useful in characterizing a nucleic acid segment or target nucleic acid. In some example methods, a nucleic acid segment may be characterized by co-partitioning the nucleic acid segment and a bead (e.g., including any suitable type of bead described herein) comprising a plurality of oligonucleotides that include a common nucleic acid barcode sequence, into a partition (including any suitable type of partition described herein, such as, for example, a droplet). The oligonucleotides may be releasably attached to the bead (e.g., releasable from the bead upon application of a stimulus to the bead, such as, for example, a thermal stimulus, a photo stimulus, and a chemical stimulus) as described elsewhere herein, and/or may comprise one or more functional sequences (e.g., a primer sequence, a primer annealing sequence, an immobilization sequence, any other suitable functional sequence described elsewhere herein, etc.) and/or one or more sequencing primer sequences as described elsewhere herein. Moreover, any suitable number of oligonucleotides may be attached to the bead, including numbers of oligonucleotides attached to beads described elsewhere herein.

[00418] Within the partition, the oligonucleotides may be attached to fragments of the nucleic segment or to copies of portions of the nucleic acid segment, such that the fragmentsor copies are also attached to the common nucleic barcode sequence. The fragments may be overlapping fragments of the nucleic acid segment and may, for example, provide greater than 2X coverage, greater than 5X coverage, greater than 10X coverage, greater than 20X coverage, greater than 40X coverage, or even greater coverage of the nucleic acid segment. In some cases, the oligonucleotides may comprise a primer sequence capable of annealing with a portion of the nucleic acid segment or a complement thereof. In some cases, the oligonucleotides may be attached by extending the primer sequences of the oligonucleotides to replicate at least a portion of the nucleic acid segment or complement thereof, to produce a copy of at least a

portion of the nucleic acid segment comprising the oligonucleotide, and, thus, the common nucleic acid barcode sequence.

[00419] Following attachment of the oligonucleotides to the fragments of the nucleic acid segment or to the copies of the portions of the nucleic acid segment, the fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment and the attached oligonucleotides (including the oligonucleotide's barcode sequence) may be sequenced via any suitable sequencing method, including any type of sequencing method described herein, to provide a plurality of barcoded fragment sequences or barcoded copy sequences. Following sequencing, the fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment can be characterized as being linked within the nucleic acid segment at least in part, upon their attachment to the common nucleic acid barcode sequence. As will be appreciated, such characterization may include sequences that are characterized as being linked and contiguous, as well as sequences that may be linked within the same fragment, but not as contiguous sequences. Moreover, the barcoded fragment sequences or barcoded copy sequences generated during sequencing can be assembled into one or more contiguous nucleic acid sequences based at least in part on the common nucleic acid barcode sequence and/or a non-barcode portion of the barcoded fragment sequences or barcoded copy sequences.

[00420] In some cases, a plurality of nucleic acid segments (e.g., fragments of at least a portion of a genome, as described elsewhere herein) may be co-partitioned with a plurality of different beads in a plurality of separate partitions, such that each partition of a plurality of different partitions of the separate partitions contains a single bead. The plurality of different beads may comprise a plurality of different barcode sequences (e.g., at least 1,000 different barcode sequences, at least 10,000 different barcode sequences, at least 100,000 different barcode sequences, at least 1,000,000 different barcodes sequences, or any other number of different barcode sequences as described elsewhere herein). In some cases, two or more, three or more, four or more, five or more, six or more, seven or more of the plurality of separate partitions may comprise beads that comprise the same barcode sequence. In some cases, at least 0.01%, 0.1%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% of the separate partitions may comprise beads having the same barcode sequence. Moreover, each bead may comprise a plurality of attached oligonucleotides that include a common nucleic acid barcode sequence.

[00421] Following co-partitioning, barcode sequences can be attached to fragments of the nucleic acid segments or to copies of portions of the nucleic acid segments in each partition. The fragments of the nucleic acid segments or the copies of the portions of the nucleic acid segments can then be pooled from the separate partitions. After pooling, the fragments of the nucleic acid segments or copies of the portions of the nucleic acid segments and any associated barcode sequences can be sequenced (e.g., using any suitable sequencing method, including those described herein) to provide sequenced fragment or sequenced copies. The sequenced fragments or sequenced copies can be characterized as deriving from a common nucleic acid segment, based at least in part upon the sequenced fragments or sequenced copies comprising a common barcode sequence. Moreover, sequences obtained from the sequenced fragments

or sequenced copies may be assembled to provide a contiguous sequence of a sequence (e.g., at least a portion of a genome) from which the sequenced fragments or sequenced copies originated. Sequence assembly from the sequenced fragments or sequenced copies may be completed based, at least in part, upon each of a nucleotide sequence of the sequenced fragments and a common barcode sequence of the sequenced fragments.

[00422] In another example method, a target nucleic acid may be characterized by partitioning fragments of the target nucleic acid into a plurality of droplets. Each droplet can comprise a bead attached to a plurality of oligonucleotides comprising a common barcode sequence. The common barcode sequence can be attached to fragments of the fragments of the target nucleic acid in the droplets. The droplets can then be pooled and the fragments and associated barcode sequences of the pooled droplets sequenced using any suitable sequencing method, including sequencing methods described herein. Following sequencing, the fragments of the fragments of the target nucleic acid may be mapped to the fragments of the target nucleic acid based, at least in part, upon the fragments of the fragments of the target nucleic acid comprising a common barcode sequence.

[00423] The application of the methods, compositions and systems described herein in sequencing may generally be applicable to any of a variety of different sequencing technologies, including NGS sequencing technologies such as Illumina MiSeq, HiSeq and X10 Sequencing systems, as well as sequencing systems available from Life Technologies, Inc., such as the Ion Torrent line of sequencing systems. While discussed in terms of barcode sequences, it will be appreciated that the sequenced barcode sequences may not include the entire barcode sequence that is included, e.g., accounting for sequencing errors. As such, when referring to characterization of two barcode sequences as being the same barcode sequence, it will be appreciated that this may be based upon recognition of a substantial portion of a barcode sequence, e.g., varying by fewer than 5, 4, 3, 2 or even a single base.

Sequencing from Small Numbers of Cells

[00424] Methods provided herein may also be used to prepare polynucleotides contained within cells in a manner that enables cell-specific information to be obtained. The methods enable detection of genetic variations from very small samples, such as from samples comprising about 10-100 cells. In some cases, about 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 cells may be used in the methods described herein. In some cases, at least about 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 cells may be used in the methods described herein. In other cases, at most about 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 cells may be used in the methods described herein.

[00425] In an example, a method may comprise partitioning a cellular sample (or crude cell extract) such that at most one cell (or extract of one cell) is present within a partition, e.g., fluidic droplet, and is co-partitioned with the barcode oligonucleotides, e.g., as described above. Processing then involves lysing the cells, fragmenting the polynucleotides contained within the cells, attaching the fragmented polynucleotides to barcoded beads, pooling the barcoded beads, and sequencing the resulting barcoded nucleic acid fragments.

[00426] As described elsewhere herein, the barcodes and other reagents may be encapsulated within, coated on, associated with, or dispersed within a bead (*e.g.* gel bead). The bead may be loaded into a fluidic droplet contemporaneously with loading of a sample (*e.g.* a cell), such that each cell is contacted with a different bead. This technique may be used to attach a unique barcode to oligonucleotides obtained from each cell. The resulting tagged oligonucleotides may then be pooled and sequenced, and the barcodes may be used to trace the origin of the oligonucleotides. For example, oligonucleotides with identical barcodes may be determined to originate from the same cell, while oligonucleotides with different barcodes may be determined to originate from different cells.

[00427] The methods described herein may be used to detect a specific gene mutation that may indicate the presence of a disease, such as cancer. For example, detecting the presence of a V600 mutation in the BRAF gene of a colon tissue sample may indicate the presence of colon cancer. In other cases, prognostic applications may include the detection of a mutation in a specific gene or genes that may serve as increased risk factors for developing a specific disease. For example, detecting the presence of a BRCA1 mutation in a mammary tissue sample may indicate a higher level of risk to developing breast cancer than a person without this mutation. In some examples, this disclosure provides methods of identifying mutations in two different oncogenes (*e.g.*, KRAS and EGRF). If the same cell comprises genes with both mutations, this may indicate a more aggressive form of cancer. In contrast, if the mutations are located in two different cells, this may indicate that the cancer may be more benign, or less advanced.

Analysis of Gene Expression

[00428] Methods of the disclosure may be applicable to processing samples for the detection of changes in gene expression. A sample may comprise a cell, mRNA, or cDNA reverse transcribed from mRNA. The sample may be a pooled sample, comprising extracts from several different cells or tissues, or a sample comprising extracts from a single cell or tissue.

[00429] Cells may be placed directly into a fluidic droplet and lysed. After lysis, the methods of the disclosure may be used to fragment and barcode the oligonucleotides of the cell for sequencing. Oligonucleotides may also be extracted from cells prior to introducing them into a fluidic droplet used in a method of the disclosure. Reverse transcription of mRNA may be performed in a fluidic droplet described herein, or outside of such a fluidic droplet. Sequencing cDNA may provide an indication of the abundance of a particular transcript in a particular cell over time, or after exposure to a particular condition.

Partitioning Polynucleotides from Cells or Proteins

[00430] In one example the compositions, methods, devices, and kits provided in this disclosure may be used to encapsulate cells or proteins within the fluidic droplets. In one example, a single cell or a plurality of cells (*e.g.*, 2, 10, 50, 100, 1000, 10000, 25000, 50000, 10000, 50000, 1000000, or more cells) may be loaded onto, into, or within a bead along with a lysis buffer within a fluidic droplet and incubated for a specified period of time. The bead may be porous, to allow washing of the contents of the bead, and

introduction of reagents into the bead, while maintaining the polynucleotides of the one or more cells (*e.g.* chromosomes) within the fluidic droplets. The encapsulated polynucleotides of the one or more cells (*e.g.* chromosomes) may then be processed according to any of the methods provided in this disclosure, or known in the art. This method can also be applied to any other cellular component, such as proteins.

Epigenetic Applications

[00431] Compositions, methods, devices, and kits of this disclosure may be useful in epigenetic applications. For example, DNA methylation can be in indicator of epigenetic inheritance, including single nucleotide polymorphisms (SNPs). Accordingly, samples comprising nucleic acid may be treated in order to determine bases that are methylated during sequencing. In some cases, a sample comprising nucleic acid to be barcoded may be split into two aliquots. One aliquot of the sample may be treated with bisulfite in order to convert unmethylated cytosine containing nucleotides to uracil containing nucleotides. In some cases, bisulfite treatment can occur prior to sample partitioning or may occur after sample partitioning. Each aliquot may then be partitioned (if not already partitioned), barcoded in the partitions, and additional sequences added in bulk as described herein to generate sequencer-ready products. Comparison of sequencing data obtained for each aliquot (e.g., bisulfite-treated sample vs. untreated sample) can be used to determine which bases in the sample nucleic acid are methylated.

[00432] In some cases, one aliquot of a split sample may be treated with methylation-sensitive restriction enzymes (MSREs). Methylation specific enzymes can process sample nucleic acid such that the sample nucleic acid is cleaved as methylation sites. Treatment of the sample aliquot can occur prior to sample partitioning or may occur after sample partitioning and each aliquot may be partitioned used to generate barcoded, sequencer-ready products. Comparison of sequencing data obtained for each aliquot (e.g., MSRE-treated sample vs. untreated sample) can be used to determine which bases in the sample nucleic acid are methylated.

Low Input DNA Applications

[00433] Compositions and methods described herein may be useful in the analysis and sequencing of low polynucleotide input applications. Methods described herein, such as PHASE, may aid in obtaining good data quality in low polynucleotide input applications and/or aid in filtering out amplification errors. These low input DNA applications include the analysis of samples to sequence and identify a particular nucleic acid sequence of interest in a mixture of irrelevant or less relevant nucleic acids in which the sequence of interest is only a minority component, to be able to individually sequence and identify multiple different nucleic acids that are present in an aggregation of different nucleic acids, as well as analyses in which the sheer amount of input DNA is extremely low. Specific examples include the sequencing and identification of somatic mutations from tissue samples, or from circulating cells, where the vast majority of the sample will be contributed by normal healthy cells, while a small minority may derive from tumor or other cancer cells. Other examples include the characterization of multiple individual population components, e.g., in microbiome analysis applications, where the contributions of individual population

members may not otherwise be readily identified amidst a large and diverse population of microbial elements. In a further example, being able to individually sequence and identify different strands of the same region from different chromosomes, e.g., maternal and paternal chromosomes, allows for the identification of unique variants on each chromosome. Additional examples of low polynucleotide input applications of the compositions, methods, and systems described herein are set forth in U.S. Provisional Patent Application No. _____ (Attorney Docket No. 43487-727.101), filed of even date herewith.

[00434] The advantages of the methods and systems described herein are clearer upon a discussion of the problems confronted in the present state of the art. In analyzing the genetic makeup of sample materials, e.g., cell or tissue samples, most sequencing technologies rely upon the broad amplification of target nucleic acids in a sample in order to create enough material for the sequencing process. Unfortunately, during these amplification processes, majority present materials will preferentially overwhelm portions of the samples that are present at lower levels. For example, where a genetic material from a sample is comprised of 95% normal tissue DNA, and 5% of DNA from tumor cells, typical amplification processes, e.g., PCR based amplification, will quickly amplify the majority present material to the exclusion of the minority present material. Furthermore, because these amplification reactions are typically carried out in a pooled context, the origin of an amplified sequence, in terms of the specific chromosome, polynucleotide or organism will typically not be preserved during the process.

[00435] In contrast, the methods and systems described herein partition individual or small numbers of nucleic acids into separate reaction volumes, e.g., in droplets, in which those nucleic acid components may be initially amplified. During this initial amplification, a unique identifier may be coupled to the components to the components that are in those separate reaction volumes. Separate, partitioned amplification of the different components, as well as application of a unique identifier, e.g., a barcode sequence, allows for the preservation of the contributions of each sample component, as well as attribution of its origin, through the sequencing process, including subsequent amplification processes, e.g., PCR amplification.

[00436] The term "about," as used herein and throughout the disclosure, generally refers to a range that may be 15% greater than or 15% less than the stated numerical value within the context of the particular usage. For example, "about 10" would include a range from 8.5 to 11.5.

[00437] As will be appreciated, the instant disclosure provides for the use of any of the compositions, libraries, methods, devices, and kits described herein for a particular use or purpose, including the various applications, uses, and purposes described herein. For example, the disclosure provides for the use of the compositions, methods, libraries, devices, and kits described herein in partitioning species, in partitioning oligonucleotides, in stimulus-selective release of species from partitions, in performing reactions (e.g., ligation and amplification reactions) in partitions, in performing nucleic acid synthesis reactions, in barcoding nucleic acid, in preparing polynucleotides for sequencing, in sequencing polynucleotides, in polynucleotide phasing (see e.g., U.S. Provisional Patent Application No._____ , (Attorney Docket No. 43487-726.101) filed of even date herewith), in sequencing polynucleotides from small numbers of cells, in analyzing gene expression, in partitioning polynucleotides from cells, in mutation

detection, in neurologic disorder diagnostics, in diabetes diagnostics, in fetal aneuploidy diagnostics, in cancer mutation detection and forensics, in disease detection, in medical diagnostics, in low input nucleic acid applications, such as circulating tumor cell (CTC) sequencing, in a combination thereof, and in any other application, method, process or use described herein.

[00438] Any concentration values provided herein are provided as admixture concentration values, without regard to any in situ conversion, modification, reaction, sequestration or the like. Moreover, where appropriate, the sensitivity and/or specificity of methods (e.g., sequencing methods, barcoding methods, amplification methods, targeted amplification methods, methods of analyzing barcoded samples, etc.) described herein may vary. For example, a method described herein may have specificity of greater than 50%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% and/or a sensitivity of greater than 50%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5%.

X. Examples

Example 1: Creation of Gel Beads Functionalized with Acrydite Primer

[00439] Gel beads are produced according to the method illustrated in **Fig 2**. In nuclease free water, 1 mL stock solutions are prepared at the following concentrations: an acrylamide precursor (Compound A) = 40% (v/v) stock solution, a crosslinker (Bis-acryloyl cystamine - Compound B) = 3.19mg/mL in 50:50 mix of acetonitrile:water, an initiator (Compound C) = 20 mg/mL, and di-sulfide acrydite primer (Compound D) = 1mM. From these stock solutions, 1 mL of an aqueous Gel Bead (GB) working solution is prepared by mixing the following volumes: nuclease free water = 648 µL, Compound A = 150 µL, Compound B = 100 µL, Compound C = 100 µL, and Compound D = 2 µL. Stock solutions of Compound A and B and GB working solutions are prepared daily.

[00440] The Gel Bead (GB) working solution, **201,** is an aqueous fluid that contains the crosslinker, BAC, and a polymer precursor solution with di-sulfide-modified acrydite oligonucleotides at a concentration of between about 0.1 and about 100µm. The second fluid, **202,** is a fluorinated oil containing the surfactant, Krytox FSH 1.8% w/w HFE 7500. The accelerator, tetramethylethylenediamine (TEMED) is added a) to the oil prior to droplet generation, **203**, b) in the line after droplet generation, **205**, and/or c) to the outlet reservoir after droplet generation, **206** to give a final concentration of 1% (v/v). TEMED is made fresh daily. Gel beads are generated by sending the aqueous and oil phase fluids to a droplet generator, **204**. Polymerization is initiated immediately after droplet generation and continues to the outlet well. Gelation is considered complete after 15-20 minutes. After gelation, generated gel beads are subjected to continuous phase exchange by washing in HFE 7500, **207**, to remove excess oil, and re-suspending the beads in aqueous solution. In some cases, the resulting beads may be present in an agglomeration. The agglomeration of gel beads are separated into individual gel beads with vortexing. Gel beads are visualized under a microscope.

Example 2: Creation of Barcoded Gel Beads by Limiting Dilution

[00441] Functionalized gel beads are produced by limiting dilution according to the method illustrated in **Fig 3A** and **Fig 4**. Gel beads with acrydite oligonucleotides (with or without a di-sulfide modification), **301**, **401**, are mixed with barcode-containing template sequences, **302**, at a limiting dilution. PCR reagents, **303**, including a biotin labeled read primer, **406**, are mixed with the gel beads and template sequences, **304**. The beads, barcode template, and PCR reagents are emulsified into a gel-water-oil emulsion by shaking/agitation, flow focusing, or microsieve, **305,** preferably such that at most one barcode template is present in a partition (*e.g.*, droplet) within the emulsion. The emulsion is exposed to one or more thermal cycles, **306**. The first thermocycle incorporates the complement barcode sequence, **408**, and immobilizes it onto the gel bead.

[00442] Continued thermal cycling is performed to clonally amplify the barcode throughout the gel bead and to incorporate the 5' biotin labeled primer into the complementary strand for downstream sorting of beads which contain barcode sequences from those that do not. The emulsion is broken, **307**, by adding perfluorodecanol, removing the oil, washing with HFE-7500, adding aqueous buffer, centrifuging, removing supernatant, removing undesired products (*e.g.* primer dimers, starting materials, deoxynucleotide triphosphates (dNTPs), enzymes, etc.) and recovering degradable gel beads into an aqueous suspension. The functionalized gel beads are re-suspended in high salt buffer, **308**. Streptavidin-labeled magnetic beads are added to the re-suspension, which is then incubated to allow binding to gel beads attached to biotinylated barcodes **308**, **410**. A magnetic device is then used to separate positive barcoded gel beads from beads that are not attached to barcode, **308**. Denaturation conditions, **309**, (*e.g.* heat or chemical denaturant) are applied to the gel beads in order to separate the biotinylated complementary strand from the barcoded beads. The magnetic beads are subsequently removed from the solution; and the resulting solution of partially-functionalized barcoded beads is pooled for further processing.

Example 3: Further Functionalization of Barcoded beads

[00443] As shown in **Fig 3B**, the barcoded gel beads, **311**, from **Example 2**, are further functionalized as follows. The beads are combined with an additional template oligonucleotide, **310**, (such as an oligonucleotide containing a random N-mer sequence, **413**, as shown in **Fig 4)**, and PCR reagents, **312**, **313,** and subjected to conditions to enable hybridization of the template oligonucleotide with the read primer attached to the gel bead. An extension reaction is performed so that the barcode strands are extended, **314,** thereby incorporating the complementary sequence of the template oligonucleotide. Resulting functionalized gel beads are re-suspended in aqueous buffer, **315**, and exposed to heating conditions to remove complement strands, **316**, and placed into aqueous storage buffer, **317**.

Example 4: Step-by-Step Description of Bead Functionalization

[00444] Fig 4 provides a step-by-step description of an example process of functionalizing the gel beads with barcodes and random N-mers. As shown in **Fig 4A**, the process begins with gel beads, **401**, that are

attached to a universal primer, such as a P5 primer (or its complement), **403**. The beads may be linked to the primer via a di-sulfide bond, **402.** The gel beads are provided in an aqueous solution (g/w). Using a limiting dilution and partitioning, unique barcode sequence templates, **405**, are combined with the beads such that at most one unique barcode sequence occupies the same partition as a gel bead. Generally, the partitions are aqueous droplets within a gel/water/oil (g/w/o) emulsion. As shown in **Fig 4B**, the barcode sequence template, **405**, is contained within a larger nucleotide strand that contains a sequence, **404**, that is complementary to the universal primer **403**, as well as a sequence, **407**, that is identical in sequence to a biotin labeled read primer, **406**.

[00445] As shown in **Fig 4C**, an amplification reaction is then conducted to incorporate the complement, **408**, of the barcode template, **405**, onto the strand that is attached to the bead**.** The amplification reaction also results in incorporation of a sequence, **415**, that is complementary to sequence, **407.** Additional amplification cycles result in hybridization of the biotin labeled read primer, **406**, to sequence, **415** (**Fig 4D**), and the biotin labeled read primer is then extended (**Fig 4E**). The emulsion may then be broken, and the gel beads may then be pooled into a gel/water common solution.

[00446] In the gel/water solution, magnetic capture beads, **409,** are then used to capture the biotinylated nucleic acids attached to the gel beads, which are then isolated from beads that only contain the original primer (**Fig 4F** and **Fig 4G**). The biotinylated strand is then removed from the strand attached to the gel bead (**Fig 4H**). Random N-mer sequences, **414**, may then be attached to the strands attached to the gel bead. For each gel bead, an identical barcode sequence, **408**, is attached to each primer throughout the gel bead; each barcode sequence is then functionalized with a random N-mer sequence, **414**, such that multiple different random N-mer sequences are attached to each bead. For this process, a random N-mer template sequence, **413**, linked to a sequence, **412**, complementary to sequence, **415**, is introduced to the solution containing the pooled beads (**Fig 4I**). The solution is subjected to conditions to enable hybridization of the template to the strand attached to the bead and sequence **415** is extended to include the random N-mer, **414**. (**Fig 4J**). The fully functionalized beads (**Fig 4K**) are then combined with a sample nucleic acid and a reducing agent (*e.g.*, dithiothreitol (DTT) at a concentration of 1mM) and partitioned within droplets of a gel/water/oil emulsion (**Fig 4L**). This combining step may be conducted with a microfluidic device (**Fig 5A**). The gel beads are then degraded within each partition (*e.g.*, droplet) such as by the action of a reducing agent, and the barcoded sequence is released from the droplet (**Fig 4M** and **Fig 4N**). The random N-mer within the barcoded sequence may serve as a primer for amplification of the sample nucleic acid.

Example 5: Use of a Microfluidic Chip to Combine the Gel-Beads-in Emulsions (GEMs) with Sample

[00447] The functionalized gel beads may be combined with sample using a double-cross microfluidic device illustrated in **Fig 5**. Degradable gel beads are introduced to the fluidic input, **501**, in a fluid stream, which contains about 7% glycerol. The experimental sample of interest is introduced to the fluidic input, **502**, in a fluid stream, which is aqueous phase. The reducing agent, dithiothreitol (DTT) at a concentration of about 1mM is introduced to the fluidic input, **503**, in a fluid stream, which contains

about 7% glycerol. Fluidic inputs **501**, **502**, and **503** mix at a microfluidic cross junction, **504**, and enter a second microfluidic cross junction, **506**. The second microfluidic cross junction can be used to produce emulsified (w/o) droplets containing the el beads. Fluidic input, **505**, is used to introduce oil with 2% (w/w) bis krytox peg (BKP). Individual droplets exiting from the second microfluidic cross junction, **507**, are added into microplate wells, **Fig 5C**, for further downstream applications. **Fig 5D** is an image of droplets generated in the absence of DTT (and therefore containing gel beads). **Fig 5E** is an image of droplets generated with DTT that caused the internal gel beads to degrade.

Example 6: Fluorescent Identification of Positive Gel Beads

[00448] **Fig 6** depicts images of gel beads containing amplified nucleic acids that have been labeled with a fluorescent label. Functionalization of the gel beads is first performed using a limiting dilution so that only a portion of the gel beads are functionalized with barcodes. Gel beads suspended in a bis krytox peg (BKP) emulsion are imaged at 4X magnification following PCR thermocycling but before washing. The bright field image, **Fig 6A**, shows all emulsion-generated droplets, and the fluorescent image, **Fig 6B**, shows only positive functionalized gel beads. Many non-fluorescent droplets are generated indicating empty droplets, which do not contain either gel bead and/or oligonucleotide. Empty droplets are washed away by multiple re-suspensions and washing in HFE-7500. **Fig 6C** and **Fig 6D** show positive gel bead enrichment following emulsion breaking and further wash steps. The bright field images (4X), **Fig 6C**, and (10X) **Fig 6E**, show all gel beads. The fluorescent images (4X), **Fig 6D**, and (10X), **Fig 6F**, show 30% positive beads from SYBR staining. The 30% positive bead result matches predicted value from gDNA input.

[00449] **Fig 7** shows images of gel beads containing single stranded (ss) DNA, double-stranded (ds) DNA, and denatured, ssDNA. Gel beads stained with 1X EvaGreen are brighter in the presence of dsDNA as evident from the fluorescent images taken at step 1: Make (ssDNA), **Fig 7A**, step 2: Extension (dsDNA), **Fig 7B**, and step 3: Denature (ssDNA), **Fig 7C**. Fluorescent images show that beads become brighter after extension and become dimmer after denaturation.

Example 7: Enrichment of Positive Gel Beads Using Streptavidin-Coated Magnetic Beads

[00450] Enrichment of positive gel beads using streptavidin-coated magnetic beads is depicted in **Fig 8**. **Fig 8A** (bright field) and **Fig B** (fluorescent) provides images of SYBR-stained gel beads 24 hours following the addition of magnetic beads. Magnetic coated positive gel beads are brighter due to SYBR staining. Bright field images before, **Fig 8C,** and after sorting, **Fig 8D,** at a magnetic bead concentration of 40mg/mL, show positive gel bead enrichment, where coated beads are optically brighter. Bright field images before, **Fig 8E**, and after sorting, **Fig 8F**, at a magnetic bead concentration of 60 mg/mL, show positive gel bead enrichment, where coated beads are optically brighter. At each magnetic bead working concentration, a single gel bead is coated by about 100-1000 magnetic beads.

Example 8: Dissolution of Gel Beads

[00451] Heating gel beads in basic solution degrades the gel beads as evident in **Fig 9**. Gel beads are heated in basic solution at 95°C and monitored at 5 minute heating intervals: t = 0min, **Fig 9A**, t = 5min, **Fig 9B**, t = 10min, **Fig 9C**, t = 15min, **Fig 9D**. Following 15 minutes, gel beads are completely degraded. Gel beads more than double in size while they are degrading. **Fig 10** depicts dissolution of the gel beads using *tris*(2-carboxyethyl)phosphine (TCEP), which is an effective and irreversible di-sulfide bond reducing agent. Functionalized gel beads, **Fig 10A**, are placed into basic solution, pH = 8, with 1mM TCEP and monitored at 2 minute intervals: t = 0min, **Fig 10B**, t = 2min, **Fig 10C**, t = 4min, **Fig 10D**, t = 6min, **Fig 10E**, t = 8min, **Fig 10F**, t = 10min, **Fig 10G**. Between about 6 and about 10 minutes, the functionalized gel beads are completely degraded.

Example 9: Analysis of Content After Dissolution Gel Beads (GB)

[00452] An analysis of content attached to gel beads is provided in **Fig 11**, and **Fig 12**. Gel beads are functionalized, **1101**, with barcode or barcode complement (N12C) and a random N-mer (8mer) that is 8 nucleotides in length, **1102**. The random N-mer is attached by performing a primer extension reaction using a template construct containing R1C and a random N-mer **1102**. The length of the entire oligonucleotide strand (including the bar code and random N-mer) is 82 bp, **1101**. The strand length of the random N-mer and the R1C is 42 base pairs (bp), **1102**. The extension reaction is performed using a KAPA HIFI RM Master Mix under high primer concentration (10μm) at 65°C for one hour. Increasing the number of wash steps before the step of degrading the gel beads results in a reduction in the amount of primer dimers within the sample. When no washes are performed, **1103**, both 42 bp products, **1106**, and 80 bp products, **1107**, can be observed. After three washes, the level of primer dimer, **1104**, is reduced relative to the no-wash experiment. After six washes, **1105**, 80 bp products, **1107**, are observed, but no primer dimers are observed.

[00453] The six-wash experiment can also be performed using six different temperatures (65°C, 67°C, 69°C, 71°C, 73°C, 75°C, **Fig 11C**) for the extension step. In this specific example, a high primer concentration (10μm) is used and the extension step lasts one hour. It appears that 67°C is the optimal temperature for both optimizing the level of 80bp products and minimizing the number of 42 bp products, **1109**.

[00454] The temperature, 67°C, is chosen for subsequent denaturation studies. Heat denaturation of the complementary strand, wherein the sample is heated to 95°C six times and washed to remove complementary strand, results in an 84 bp peak, **1202**, before denaturation, and shows a reduced peak, **1201**, following denaturation. The control value measured from step 1 is shown at **1203**.

Example 10: Creation of Barcoded Gel Beads by Partitioning in Wells

[00455] Functionalized beads are produced by partitioning in wells according to the method illustrated in **Fig 13A** and **13B**. The first functionalization step is outlined in **Fig 13A**, the second functionalization step is outlined in **Fig 13B**. An example multiplex adaptor creation scheme is outlined in **Fig 13C** and

described in Example 11. As shown in **Fig 13A**, functionalized beads, **1301** (e.g., beads with acrydite oligos and primer (e.g., 5'-AAUGAUACGGCGACCACCGAGA-3'), the template with barcode sequence, **1302** (e.g., 5'-XXXXXXTCTCGGTGGTCGCCGTATCATT-3'), and appropriate PCR reagents, **1303**, are mixed together, **1304/1305** and divided into 384 wells of a multi-well plate. Each well comprises multiple copies of a unique barcode sequence and multiple beads. Thermocycling, **1306**, with an extension reaction is performed in each individual well to form beads with attached barcodes. All wells are pooled together and cleaned up in bulk, **1307/1308**.

[00456] To add a random N-mer, the partially functionalized beads, **1310**, the template random N-mer oligonucleotides, **1309**, and the appropriate PCR reagents, **1311**, are mixed together, **1312**, and the functionalized beads **1310** subjected to extension reactions **1313** to add a random N-mer sequence complementary to the random N-mer template, to the beads. Following thermal cycling, the beads are cleaned up in bulk, **1314-1316**.

Example 11: Combinatorial Plate Technique

[00457] As shown in **Fig 13C**, beads **1317** attached to primers (*e.g.*, P5 oligomers, 5'-AAUGAUACGGCGACCACCGAGA-3') **1318** are partitioned into wells of a multi-well plate (such as a 5X-1 384-well plate **1319**) with multiple copies of a template **1321** comprising a unique template partial barcode sequence (e.g., 5'-XXXXXXTCTCGGTGGTCGCCGTATCATT-3). Extension reactions (e.g., extension of primer **1318** via template **1321**) are performed to generate Bead-P5-[5X-1], **1320** comprising an extension product (e.g., an oligonucleotide comprising primer **1318** and a partial barcode sequence complementary to the template partial barcode sequence) in each well. The beads are removed from the wells are pooled together and a clean-up step is performed in bulk.

[00458] The pooled mixture is then re-divided into wells of a second multiwell plate such as a 384-well plate with 5X-2, **1322**, with each well also comprising an oligonucleotide comprising a second unique partial barcode sequence and a random N-mer (e.g., 5'P-YYYYYYCGCACACUCUUUCCCUACACGACGCUCUUCCGAUCUNNNNNNNN-BLOCK). The oligonucleotide may have a blocker oligonucleotide attached (e.g., via hybridization) (e.g., "BLOCK"). Single-stranded ligation reactions **1324** are performed between the extension product bound to the bead and the oligonucleotide comprising the second partial barcode sequence and random N-mer. Following the ligation reaction, beads comprising a full barcode sequence (e.g., XXXXXXYYYYYY) and a random N-mer are generated, **1323** (e.g., Bead-P5-[5X-1][5X-2]R1[8N-Blocker]). The beads also comprise the blocker oligonucleotide. All wells are then pooled together, the blocking groups are cleaved, and the bead products are cleaned up in bulk. Beads comprising a large diversity of barcode sequences are obtained.

Example 12: Partial Hairpin Amplification for Sequencing (PHASE) Reaction

[00459] Partial Hairpin Amplification for Sequencing (PHASE) reaction is a technique that can be used to mitigate undesirable amplification products according to the method outlined in **Fig 14** and **Fig 15** by

forming partial hairpin structures. Specifically, random primers, of about 8N-12N in length, **1404**, tagged with a universal sequence portion, **1401/1402/1403**, may be used to randomly prime and extend from a nucleic acid, such as, genomic DNA (gDNA). The universal sequence comprises: (1) sequences for compatibility with a sequencing device, such as, a flow cell (*e.g.* Illumina's P5, **1401**, and Read 1 Primer sites, **1402**) and (2) a barcode (BC), **1403**, (*e.g.*, 6 base sequences). In order to mitigate undesirable consequences of such a long universal sequence portion, uracil containing nucleotides are substituted for thymine containing nucleotides for all but the last 10-20 nucleotides of the universal sequence portion, and a polymerase that will not accept or process uracil-containing templates is used for amplification of the nucleic acid, resulting in significant improvement of key sequencing metrics, **Fig 16A**, **Fig 21**, and **Fig 22**. Furthermore, a blocking oligonucleotide comprising uracil containing nucleotides and a blocked 3' end (*e.g.* 3'ddCTP) are used to promote priming of the nucleic acid by the random N-mer sequence and prevent preferential binding to portions of the nucleic acid that are complementary to the Read 1 Primer site, **1402**. Additionally, product lengths are further limited by inclusion of a small percentage of terminating nucleotides (*e.g.*, 0.1-2% acyclonucleotides (acyNTPs)) (**Fig 16B**) to reduce undesired amplification products.

[00460] An example of partial hairpin formation to prevent amplification of undesired products is provided here. First, initial denaturation is achieved at 98°C for 2 minutes followed by priming a random portion of the genomic DNA sequence by the random N-mer sequence acting as a primer for 30 seconds at 4°C (**Fig 15A**). Subsequently, sequence extension follows as the temperature ramps at 0.1°C/second to 45°C (held for 1 second) (**Fig 15A**). Extension continues at elevated temperatures (20 seconds at 70°C), continuing to displace upstream strands and creating a first phase of redundancy (**Fig 15B**). Denaturation occurs at 98°C for 30 seconds to release genomic DNA for additional priming. After the first cycle, amplification products have a single 5' tag (**Fig 15C**). These aforementioned steps are repeated up to 20 times, for example by beginning cycle 2 at 4°C and using the random N-mer sequence to again prime the genomic DNA where the black sequence indicates portions of the added 5' tags (added in cycle 1) that cannot be copied (**Fig 15D**). Denaturation occurs at 98°C to again release genomic DNA and the amplification product from the first cycle for additional priming. After a second round of thermocycling, both 5' tagged products and 3' & 5' tagged products exist (**Fig 15E).** Partial hairpin structures form from the 3' & 5' tagged products preventing amplification of undesired products **(Fig 15F).** A new random priming of the genomic DNA sequence begins again at 4°C (**Fig 15G**).

Example 13: Adding Additional Sequences by Amplification

[00461] For the completion of sequencer-ready libraries, an additional amplification (*e.g.*, polymerase chain reaction (PCR) step) is completed to add additional sequences, **Fig 14C**. In order to out-compete hairpin formation, a primer containing locked nucleic acid (LNAs) or locked nucleic acid nucleotides, is used. Furthermore, in cases where the inclusion of uracil containing nucleotides is used in a previous step, a polymerase that does not discriminate against template uracil containing nucleotides is used for this step. The results presented in **Fig 17** show that a blocking oligonucleotide reduces start site bias, as measured

by sequencing on an Illumina MiSeq sequencer. The nucleic acid template in this case is yeast gDNA.

Example 14: Digital Processor

[00462] A conceptual schematic for an example control assembly, **1801**, is shown in **Fig 18**. A computer, **1802**, serves as the central hub for control assembly, **1801**. Computer, **1802**, is in communication with a display, **1803**, one or more input devices (*e.g.*, a mouse, keyboard, camera, etc.) **1804**, and optionally a printer, **1805**. Control assembly, **1801,** via its computer, **1802**, is in communication with one or more devices: optionally a sample pre-processing unit, **1806**, one or more sample processing units (such as a sequence, thermocycler, or microfluidic device) **1807**, and optionally a detector, **1808**. The control assembly may be networked, for example, via an Ethernet connection. A user may provide inputs (*e.g.*, the parameters necessary for a desired set of nucleic acid amplification reactions or flow rates for a microfluidic device) into computer, **1802**, using an input device, **1804.** The inputs are interpreted by computer, **1802**, to generate instructions. The computer, **1802**, communicates such instructions to the optional sample pre-processing unit, **1806**, the one or more sample processing units, **1807**, and/or the optional detector, **1808**, for execution. Moreover, during operation of the optional sample pre-processing unit, **1806**, one or more sample processing units, **1807**, and/or the optional detector, **1808**, each device may communicate signals back to computer, **1802**. Such signals may be interpreted and used by computer, **1802**, to determine if any of the devices require further instruction. Computer, **1802**, may also modulate sample pre-processing unit, **1806**, such that the components of a sample are mixed appropriately and fed, at a desired or otherwise predetermined rate, into the sample processing unit (such as the microfluidic device), **1807**. Computer, **1802**, may also communicate with detector, **1808**, such that the detector performs measurements at desired or otherwise predetermined time points or at time points determined from feedback received from pre-processing unit, **1806**, or sample processing unit, **1807**. Detector, **1808**, may also communicate raw data obtained during measurements back to computer, **1802**, for further analysis and interpretation. Analysis may be summarized in formats useful to an end user via display, **1803**, and/or printouts generated by printer, **1805**. Instructions or programs used to control the sample pre-processing unit, **1806**, the sample processing unit, **1807**, and/or detector, **1808**; data acquired by executing any of the methods described herein; or data analyzed and/or interpreted may be transmitted to or received from one or more remote computers, **1809**, via a network, **1810**, which, for example, could be the Internet.

Example 15: Combinatorial Technique via Ligation

[00463] As shown in **Fig 23A**, beads **2301** are generated and covalently linked (e.g., via an acrydite moiety) to a partial P5 sequence **2302**. Separately, in 50 μL of each well of 4 96 well plates, an oligonucleotide **2303**, comprising the remaining P5 sequence and a unique partial barcode sequence (indicated by bases "DDDDDD" in oligonucleotide **2303**), is hybridized to an oligonucleotide **2304** that comprises the reverse complement to oligonucleotide **2303** and additional bases that overhang each end

of oligonucleotide **2303**. Splint **2306** is generated. Each overhang is blocked (indicated with an "X" in **Fig 23**) with 3' C3 Spacer, 3' Inverted dT, or dideoxy-C (ddC) to prevent side product formation.

[00464] As shown in **Fig 23B,** splints **2306** are each added to 4 96 deep well plates, with each well comprising 2 mL beads **2301** and a splint comprising a unique partial barcode sequence. In each well, the splint **2306** hybridizes with the partial P5 sequence **2302** of beads **2301**, via the corresponding overhang of oligonucleotide **2304**. Following hybridization, partial P5 sequence **2302** is ligated to oligonucleotide **2303** (which will typically have been 5' phosphorylated) via the action of a ligase, e.g., a T4 ligase, at 16°C for 1 hour. Following ligation, the products are pooled and the beads washed to remove unligated oligonucleotides.

[00465] As shown in **Fig 23C**, the washed products are then redistributed into wells of 4 new 96 well plates, with each well of the plate comprising 2 mL of beads **2301** and an oligonucleotide **2305** that has a unique partial barcode sequence (indicated by "DDDDDD" in oligonucleotide **2305**) and an adjacent short sequence (e.g., "CC" adjacent to the partial barcode sequence and at the terminus of oligonucleotide **2305**) complementary to the remaining overhang of oligonucleotide **2304**. Oligonucleotide **2305** also comprises a random N-mer (indicated by "NNNNNNNNNN" in oligonucleotide **2305**). Via the adjacent short sequence, oligonucleotide **2305** is hybridized with oligonucleotide **2304** via the remaining overhang of oligonucleotide **2304**. Oligonucleotide **2305** is then ligated to oligonucleotide **2303** via the action of a ligase at 16°C for 1 hour. Ligation of oligonucleotide **2305** to oligonucleotide **2303** results in the generation of a full barcode sequence. As shown in **Fig 23D**, the products are then pooled, the oligonucleotide **2304** is denatured from the products, and the unbound oligonucleotides are then washed away. Following washing, a diverse library of barcoded beads is obtained, with each bead bound to an oligonucleotide comprising a P5 sequence, a full barcode sequence, and a random N-mer. The generated library comprises approximately 147,000 different barcode sequences.

Example 16: Substitution of Uracil Containing Nucleotides for Thymine Containing Nucleotides in Barcode Primers

[00466] As shown in **Fig 33A**, two barcode primers **3301** and **3302** suitable for PHASE amplification were used to amplify sample nucleic acid obtained from a yeast genome. Following PHASE amplification, additional sequences were added (e.g., via bulk PCR) to generate sequencer-ready products. Barcode primers **3301** (also shown as U.2 in **Fig 33A**) and **3302** (also shown at U.1 in **Fig 33A**) comprised an identical sequence except that barcode primer **3301** comprised an additional uracil containing nucleotide-for-thymine containing nucleotide substitution at position **3306**. Sets of amplification experiments were run for each barcode primer, with each set corresponding to a particular blocker oligonucleotide mixed with the respective barcode primer at various stoichiometries. For barcode primer **3302**, sets of amplification experiments corresponding to a standard blocker oligonucleotide **3303**, a full blocker oligonucleotide comprising bridged nucleic acid (BNAs) **3304** (also shown as BNA blocker in **Fig 33A**), or a full blocker oligonucleotide **3305** were conducted. Blocker oligonucleotides **3303** and **3305** comprised uracil containing nucleotide-for-thymine containing nucleotide substitutions at all thymine

containing nucleotide positions and a ddC blocked end. In each set, the blocker oligonucleotide:barcode primer stoichiometry was either 0, 0.4, 0.8, or 1.2. For barcode primer **3301**, each type of blocker oligonucleotide **3303**, **3304**, and **3305** was tested at a 0.8 blocker oligonucleotide:barcode primer stoichiometry.

[00467] The size results of PHASE amplification products are depicted in **Fig 33B**. As shown, barcode primer **3302** (e.g., comprising the extra uracil containing nucleotide-for-thymine containing nucleotide substitution) coupled to blocker oligonucleotide **3303** generally produced the smallest amplification products across the stoichiometries tested. Results for barcode primer **3302** with respect to blocker oligonucleotides **3304** and **3305** varied, with sizes generally larger than results for blocker oligonucleotide **3303**. For barcode primer **3301**, amplification product sizes were also generally larger than those obtained for barcode primer **3301** coupled to blocker oligonucleotide **3303** across the blocker oligonucleotides tested. The size results of sequencer-ready products are depicted in **Fig 33C**.

[00468] Key sequencing metrics obtained from the amplification products are depicted in **Fig 33D**. As shown, the fraction of unmapped reads (panel I in **Fig 33D**) was generally lower for sequencing runs for amplification products generated from barcode primer **3302**. For example, the fraction of unmapped reads for amplification products generated from barcode primer **3302** and blocker oligonucleotide **3303** at 0.8 blocker oligonucleotide:barcode primer stoichiometry was approximately 7-8%, whereas results obtained using barcode primer **3301** at the same conditions was approximately 17-18%. Moreover, Q40 error rates (panel II in **Fig 33D**) were also lower for barcode primer **3302**. For example, Q40 error rate for amplification products generated from barcode primer **3302** and blocker oligonucleotide **3303** at 0.8 blocker oligonucleotide:barcode primer stoichiometry was approximately 0.105%, whereas results obtained using barcode primer **3301** at the same conditions was approximately 0.142%. Read 1start site (panel III) and Read 2 start site (panel IV) relative entropies determined during sequencing are shown in **Fig 33E**.

Example 17: Post-Synthesis Functionalization of Gel Beads via Disulfide Exchange

[00469] Gel beads comprising disulfide bonds were generated according to one or more methods described herein. The gel beads were then reacted with TCEP at ratios of molecules of TCEP to gel beads (TCEP:GB). The tested ratios were 0, 2.5 billion, and 10.0 billion. The TCEP functions as a reducing agent to generate free thiols within the gel beads. Following reduction, the gel beads were washed once to remove the TCEP. Next, the generated free thiols of the gel beads were reacted with an acrydite-S-S-P5 species (e.g., **3505** in **Fig 35A**) to link the acrydite-S-S-P5 to the gel beads via Michael addition chemistry as shown in **Fig 35A**. Different ratios of acrydite-S-S-P5 to each type (e.g., ratio of TCEP:GB used to generate free thiols on the gel beads) of the activated gel beads were tested. The tested ratios of acrydite-S-S-P5 species to activated gel beads (P5:GB) were 50 million, 500 million, and 5 billion.

[00470] Following syntheses, the gel beads from each reaction were washed and treated with DTT in a reaction mixture to degrade the gel beads and release any bound acrydite-S-S-P5 species. An aliquot of

each reaction mixture was entered into a lane of a gel and free oligonucleotides subject to gel electrophoresis as shown in **Fig 36** (e.g., lanes 3-11 in **Fig 36**). A 50 picomole acrydite-S-S-P5 standard was also run (e.g., lane 1 in **Fig 36**) along with a 25 base pair ladder (e.g., lane 2 in **Fig 36**). Bands corresponding to loaded acrydite-S-S-P5 were generated in lanes 5 and 8 (indicated by arrows in **Fig 36**). Lane 5 corresponds to gel beads treated at a TCEP:GB ratio of 2.5 billion and the TCEP treated gel beads reacted with acrydite-S-S-P5 at a P5:GB ratio of 5 billion. Lane 8 corresponds to gel beads treated at a TCEP:GB ratio of 10.0 billion and the TCEP treated gel beads reacted with acrydite-S-S-P5 at a P5:GB ratio of 5 billion.

Example 18: Post-Synthesis Functionalization of Gel Beads via Disulfide Exchange

[00471] Gel beads comprising disulfide bonds were generated according to one or more methods described herein. The gel beads were then reacted with TCEP in 0.1M phosphate buffer at a concentration of 4 μg TCEP/100,000 gel beads. The TCEP can function as a reducing agent to generate gel beads with free thiol groups. Following reduction, the gel beads were washed once to separate the gel beads from the TCEP. Next, the free thiols of the gel beads were reacted with 2,2'-dithiopyridine (e.g., **3507** in **Fig 35B**) in a saturated solution (~0.2 mM) of 2,2'-dithiopyridine to link pyridine groups to the gel beads via disulfide exchange chemistry as shown in **Fig 35B**. Following synthesis, the gel beads were washed three times to remove excess 2,2'-dithiopyridine.

[00472] The washed gel beads were then reacted with an oligonucleotide **3702** comprising a full construct barcode (FCBC – e.g., an oligonucleotide comprising P5, a barcode sequence, R1, and a random N-mer) sequence at one end and a free thiol group at its other end. Two reactions were completed at two different ratios of molecules of FCBC to gel beads (e.g., FCBC:GB) and the reactions were allowed to proceed overnight. The tested FCBC:GB ratios were 400 million and 1.6 billion. Oligonucleotide **3702** was initially supplied with its free thiol group protected in a disulfide bond, shown as **3701** in **Fig 37A**. To generate the free thiol as in oligonucleotide **3702**, oligonucleotide **3701** was treated with 0.1 M DTT in 1X Tris-EDTA buffer (TE) buffer for 30 minutes. Salt exchange on a Sephadex (NAP-5) column was used to remove DTT after reduction and purify oligonucleotide **3702**. For each reaction, purified oligonucleotides **3702** were then reacted with the dithio-pyridine species of the gel beads via thiol-disulfide exchange (e.g., see **Fig 35B**) to generate gel beads comprising oligonucleotide **3702**. Following the reaction, the gel beads were purified by washing the beads three times.

[00473] For comparison purposes, gel beads comprising disulfide bonds and the FCBC sequence were also generated via polymerization of monomers as described elsewhere herein. The FCBC was linked to a monomer comprising an acrydite species that was capable of participating in a polymerization with acrylamide and bis(acryloyl)cystamine to generate the gel beads. The FCBC sequence was linked to the gel beads via the acrydite moiety.

[00474] Following syntheses, the gel beads from each reaction were washed and treated with DTT in a reaction mixture to degrade the gel beads and release any bound oligonucleotide **3702**. Gel beads comprising the FCBC sequence that were synthesized via polymerization were also treated with DTT in a

reaction mixture. An aliquot of each reaction mixture was entered into a lane of a gel and free oligonucleotides subject to gel electrophoresis as shown in **Fig 37B**. As shown in the gel photograph depicted in **Fig 37B**, lane 1 corresponds to a 50 base pair ladder; lane 2 corresponds to gel beads functionalized via disulfide exchange chemistry at an FCBC:GB ratio of 400 million; lane 3 corresponds to gel beads functionalized via disulfide exchange chemistry at an FCBC:GB ratio of 1.6 billion; and lane 4 corresponds to functionalized gel beads generated via polymerization of acrydite species. Bands corresponding to loaded oligonucleotides were generated for functionalized gel beads generated at both FCBC:GB ratios and were at a similar position to the band generated for functionalized gel beads generated via polymerization of acrydite species.

[00475] Following syntheses, gel beads from each reaction were also washed and stained with SYBR Gold fluorescent stain. Gel beads comprising the FCBC sequence that were synthesized via polymerization were also stained with SYBR Gold. SYBR Gold can stain functionalized beads by intercalating any bound oligonucleotides. Following staining, the beads were pooled and imaged using fluorescence microscopy, as shown in the micrograph depicted in **Fig 37C**. Brighter beads (**3704**) in **Fig 37C** correspond to beads functionalized during polymerization of the beads and dim beads (still showing SYBR gold signal) (**3705**) correspond to beads functionalized with disulfide exchange chemistry after gel bead generation. Loading of oligonucleotides via disulfide-exchange was approximately 30% of that achieved with functionalization of beads during gel bead polymerization.

[00476] It should be understood from the foregoing that, while particular implementations have been illustrated and described, various modifications may be made thereto and are contemplated herein. It is also not intended that the invention be limited by the specific examples provided within the specification. While the invention has been described with reference to the aforementioned specification, the descriptions and illustrations of the preferable embodiments herein are not meant to be construed in a limiting sense. Furthermore, it shall be understood that all aspects of the invention are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. Various modifications in form and detail of the embodiments of the invention will be apparent to a person skilled in the art. It is therefore contemplated that the invention shall also cover any such modifications, variations and equivalents. 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

1. A method of barcoding sample materials, comprising:

providing a first partition comprising a plurality of nucleic acid barcode molecules associated therewith, the plurality of nucleic acid barcode molecules comprising the same nucleic acid barcode sequence;

co-partitioning the first partition with components of a sample material into a second partition;

releasing the barcode molecules from the first partition into the second partition; and

attaching the barcode molecules to one or more of the components of the sample material or fragments thereof within the second partition.

2. The method of claim 1, wherein the first partition comprises at least 1,000 barcode molecules associated therewith, the at least 1,000 barcode molecules having the same barcode sequence.

3. The method of claim 1, wherein the first partition comprises at least 10,000 barcode molecules associated therewith, the at least 10,000 barcode molecules having the same barcode sequence.

4. The method of claim 1, wherein the first partition comprises at least 100,000 barcode molecules associated therewith, the at least 100,000 barcode molecules having the same barcode sequence.

5. The method of claim 1, wherein the first partition comprises at least 1,000,000 barcode molecules associated therewith, the at least 1,000,000 barcode molecules having the same barcode sequence.

6. The method of claim 1, wherein the first partition is selected from a bead, a microcapsule and a droplet.

7. The method of claim 1, wherein the first partition comprises a bead, and the barcode molecules are releasably coupled to the bead.

8. The method of claim 7, wherein the bead comprises a gel bead.

9. The method of claim 1, wherein the second partition comprises a droplet.

10. The method of claim 7, wherein the co-partitioning step comprises combining a first aqueous fluid comprising beads with a second aqueous fluid comprising the sample components in a droplet within an immiscible fluid.

11. The method of claim 1, wherein the releasing step comprises degrading the first partition.

12. The method of claim 7, wherein the releasing step comprises degrading the bead to release the barcode molecules into the second partition.

13. The method of claim 7, wherein the releasing step comprises cleaving a chemical linkage between the barcode molecules and the bead.

14. The method of claim 7, wherein at least one of crosslinking of the bead and a linkage between the bead and the barcode molecules comprises a disulfide linkage, and the releasing step comprises exposing the first partition to a reducing agent.

15. The method of claim 14, wherein the reducing agent comprises a reducing agent selected from the group consisting of DTT and TCEP.

16. The method of claim 1, wherein the sample materials comprise one or more template nucleic acid molecules, and the attaching step comprises attaching the barcode molecules to one or more fragments of the template nucleic acid molecules.

17. The method of claim 1, wherein the sample materials comprise one or more template nucleic acid molecules, the barcode molecules further comprise a primer sequence complementary to at least a portion of the template nucleic acid molecules, and the attaching step comprises extending the barcode molecules to replicate at least a portion of the template nucleic acid molecules.

18. The method of claim 1, wherein the sample materials comprise contents of a single cell.

19. The method of claim 18, wherein the single cell comprises a cancer cell.

20. The method of claim 18, wherein the single cell comprises a bacterial cell.

21. The method of claim 20, wherein the bacterial cell is isolated from a human microbiome sample.

22. The method of claim 1, wherein the second partition comprises no more than one first partition.

23. The method of claim 1, wherein:

the providing step comprises providing a plurality of first partitions comprising a plurality of different nucleic acid barcode sequences, wherein each of the plurality of first partitions comprises a plurality of at least 1000 nucleic acid barcode molecules having the same nucleic acid barcode sequence associated therewith;

the co-partitioning step comprises co partitioning the plurality of first partitions with the components of the sample material into a plurality of second partitions;

the releasing step comprises releasing the nucleic acid barcode molecules from the plurality of first partitions into the plurality of second partitions; and

the attaching step comprises attaching the nucleic acid barcode molecules to the components of the sample material or fragments thereof within the plurality of second partitions.

24. The method of claim 23, wherein the plurality of different nucleic acid barcode sequences comprises at least about 1000 different barcode sequences.

25. The method of claim 23, wherein the plurality of different nucleic acid barcode sequences comprises at least about 10,000 different barcode sequences

26. The method of claim 23, wherein the plurality of different nucleic acid barcode sequences comprises at least about 100,000 different barcode sequences

27. The method of claim 23, wherein the plurality of different nucleic acid barcode sequences comprises at least about 500,000 different barcode sequences

28. The method of claim 23, wherein a subset of the plurality of second partitions comprise the same nucleic acid barcode sequence.

29. The method of claim 28, wherein at least about 1% the plurality of second partitions comprise the same nucleic acid barcode sequence.

30. The method of claim 28, wherein at least about 2% of the plurality of second partitions comprise the same nucleic acid barcode sequence.

31. The method of claim 28, wherein at least about 5% of the plurality of second partitions comprise the same nucleic acid barcode sequence.

32. The method of claim 28, wherein the fragments of the components of the sample material comprise one or more fragments of one or more template nucleic acid sequences and the method further

comprises sequencing the one or more fragments of the one or more template nucleic acid sequences and characterizing the one or more fragments of the template nucleic acid sequences based at least in part upon a nucleic acid barcode sequence attached thereto.

33. The method of claim 32, wherein the characterizing the one or more fragments of the one or more template nucleic acid sequences comprises mapping a fragment of an individual template nucleic acid sequence of the one or more template nucleic acid sequences to an individual template nucleic acid sequence of the one or more template nucleic acid sequences or a genome from which the individual template nucleic acid sequence was derived.

34. The method of claim 32, wherein the characterizing the one or more fragments of the template nucleic acid sequences comprises at least identifying an individual nucleic acid barcode sequence of said plurality of different nucleic acid barcode sequences, and identifying a sequence of an individual fragment of the one or more fragments of the one or more template nucleic acid sequences attached to the individual nucleic acid barcode sequence.

35. The method of claim 23, wherein at least 50% of the plurality of second partitions contain no more than one first partition.

36. The method of claim 23, wherein at least 70% of the plurality of second partitions contain no more than one first partition.

37. The method of claim 23, wherein at least 90% of the plurality of second partitions contain no more than one first partition.

38. The method of claim 23, wherein at least 50% of the plurality of second partitions comprises exactly one first partition.

39. The method of claim 23, wherein at least 70% of the plurality of second partitions comprises exactly one first partition.

40. The method of claim 23, wherein at least 90% of the plurality of second partitions comprises exactly one first partition.

41. A method of barcoding sample materials, comprising:

providing a plurality of first partitions comprising a plurality of different nucleic acid barcode sequences, wherein each of the plurality of first partitions comprises a plurality of nucleic acid barcode molecules having the same nucleic acid barcode sequence associated therewith;

co-partitioning the plurality of first partitions with components of a sample material into a plurality of second partitions;

releasing the barcode molecules from the plurality of first partitions into the plurality of second partitions; and

attaching the barcode molecules to the components of the sample material within the plurality of second partitions.

42. A method of barcoding sample materials, comprising:

providing an activatable nucleic acid barcode sequence;

partitioning the activatable nucleic acid barcode sequence with one or more components of a sample material into a first partition;

activating the activatable nucleic acid barcode sequence to produce an active nucleic acid barcode sequence in the first partition; and

attaching the active nucleic acid barcode sequence to the one or more components of the sample material.

43. The method of claim 42, wherein the step of activating the activatable nucleic acid barcode sequence comprises releasing the activatable nucleic acid barcode sequence from a second partition within the first partition.

44. The method of claim 42, wherein the step of activating the activatable nucleic acid barcode sequence comprises removing a removable protecting group from the activatable nucleic acid barcode sequence.

45. A composition, comprising:

a first partition, comprising one or more sample components; and

a second partition contained within the first partition, the second partition having a plurality of oligonucleotides releasably associated therewith, the plurality of oligonucleotides comprising a common barcode sequence.

46. The composition of claim 45, wherein the first partition comprises an aqueous droplet in an emulsion and the second partition comprises a microcapsule or bead.

47. The composition of claim 45, wherein the plurality of oligonucleotides are releasably associated with the second partition by a cleavable linkage.

48. The composition of claim 47, wherein the cleavable linkage comprises a linkage selected from a chemically cleavable linkage, a photocleavable linkage and a thermally cleavable linkage.

49. The composition of claim 47, wherein the cleavable linkage comprises a disulfide linkage.

50. The composition of claim 45, wherein the second partition comprises a degradable bead.

51. The composition of claim 50, wherein the degradable bead is selected from a photodegradable bead, a chemically degradable bead, and a thermally degradable bead.

52. The composition of claim 50, wherein the degradable bead comprises chemically cleavable cross-linking.

53. The composition of claim 52, wherein the chemically cleavable cross-linking comprises disulfide cross-linking.

54. The composition of claim 45, wherein the one or more sample components comprise nucleic acids or fragments thereof.

55. The composition of claim 54, wherein the nucleic acids comprise genomic DNA.

56. The composition of claim 54, wherein the nucleic acids comprise nucleic acid fragments having a length of between about 1 kb and about 100 kb.

57. The composition of claim 54, wherein the nucleic acids comprise nucleic acid fragments having a length of between about 5 kb and about 50 kb.

58. The composition of claim 54, wherein the nucleic acids comprise nucleic acid fragments having a length of between about 10 kb and about 30 kb.

59. The composition of claim 45, comprising:

a plurality of first partitions; and

a plurality of different second partitions, each of the plurality of different second partitions being disposed within a separate one of the plurality of first partitions and comprising a plurality of oligonucleotides releasably associated therewith, the plurality of oligonucleotides associated with each of the plurality of second partitions comprising a common barcode sequence, and the oligonucleotides associated with different second partitions comprising different barcode sequences.

60. The composition of claim 59, wherein the plurality of different second partitions comprises at least 1000 different second partitions.

61. The composition of claim 59, wherein the plurality of different second partitions comprises at least 10,000 different second partitions.

62. The composition of claim 59, wherein the plurality of different second partitions comprises at least 100,000 different second partitions.

63. The composition of claim 59, wherein the plurality of different second partitions comprises at least 500,000 different second partitions.

64. A method, comprising:

combining a sample of nucleic acids with a library of barcoded beads to form a mixture;

partitioning the mixture into a plurality of partitions, wherein at least a subset of the partitions comprises at most one barcoded bead; and

within the partitions, releasing barcodes from the barcoded beads.

65. The method of claim 64, wherein the barcoded beads are capable of being dissolved by a reducing agent.

66. The method of claim 65, wherein the barcoded beads comprise disulfide bonds.

67. The method of claim 64, wherein the sample of nucleic acids is genomic DNA.

68. The method of claim 67, wherein the genomic DNA is fragmented prior to being combined with the barcoded beads.

69. The method of claim 64, further comprising using a microfluidic device to combine the sample of nucleic acids with the library of barcoded beads.

70. The method of claim 64, further comprising using a microfluidic device to partition the mixture into the plurality of partitions.

71. The method of claim 64, wherein the partitions are aqueous droplets within a water-in-oil emulsion.

72. The method of claim 71, further comprising using a microfluidic device to partition the mixture into the plurality of partitions.

73. The method of claim 72, wherein the microfluidic device is a droplet generator.

74. The method of claim 71, wherein the microfluidic device comprises a first input channel and a second input channel that meet at a junction that is fluidly connected to an output channel.

75. The method of claim 74, wherein the method further comprises introducing the sample of nucleic acids into the first input channel and introducing the library of barcoded beads into the second input channel, thereby generating the mixture of the sample of nucleic acids and the library of barcoded beads in the output channel.

76. The method of claim 75, wherein the output channel is fluidly connected to a third input channel at a junction.

77. The method of claim 76, further comprising introducing an oil into the third input channel, such that the aqueous droplets within a water-in-oil emulsion are formed, individual droplets comprising, on average, at most ten barcoded beads.

78. The method of claim 77, further comprising introducing oil into a fourth input channel, wherein the fourth input channel intersects the third input channel and the output channel at a junction.

79. The method of claim 74, further comprising introducing a reducing agent into the first or second input channel.

80. The method of claim 74, wherein the microfluidic device further comprises a third input channel that intersects the first input channel, the second input channel, or the junction of the first input channel and the second input channel, and the method further comprises introducing a reducing agent into the third input channel.

81. The method of claim 74, wherein the first input channel and the second input channel form a substantially perpendicular angle between one another.

82. The method of claim 64, wherein the barcodes are released from the barcoded beads by the action of a reducing agent.

83. The method of claim 64, wherein the barcoded beads comprise a matrix that is crosslinked with disulfide bonds and the barcodes are released from the barcoded beads by the action of a reducing agent that dissolves the barcoded beads.

84. The method of claim 64, further comprising releasing the barcodes from the barcoded beads by heating the partitions.

85. The method of claim 64, wherein the barcodes are pre-synthesized with known sequences.

86. The method of claim 64, wherein the barcodes further comprise a plurality of random Nmers.

87. The method of claim 86, further comprising hybridizing the plurality of random Nmers to the sample of nucleic acids.

88. The method of claim 86, further comprising performing a nucleic acid amplification reaction within the partitions.

89. A composition, comprising a bead that is covalently linked to a plurality of oligonucleotides comprising an identical barcode sequence and a variable domain.

90. A composition, comprising a bead, wherein:

the bead comprises a plurality of more than 1,000,000 oligonucleotides, wherein each of the plurality of oligonucleotides comprises a constant region and a variable region; and

the bead is capable of being substantially dissolved with a reducing agent.

91. A composition, comprising:

a plurality of beads;

wherein each of the plurality of beads comprises a plurality of oligonucleotides releasably coupled thereto;

wherein the plurality of oligonucleotides associated with an individual bead from the plurality of beads comprises a common barcode domain and a variable domain; and

wherein the common barcode domain is different between two or more beads in the plurality of beads.

92. A method of generating functionalized beads comprising:

a. mixing a plurality of polymers or monomers with one or more oligonucleotides;
b. crosslinking the plurality of polymers or monomers such that disulfide bonds form between the plurality of polymers or monomers, thereby forming hardened beads; and
c. causing covalent linkages to form between the one or more oligonucleotides and the plurality of polymers or monomers.

93. A method of preparing a barcode library, comprising:

providing a plurality of separate first bead populations;

attaching a first oligonucleotide comprising a first barcode sequence segment to the separate first bead populations, wherein each separate first bead population comprises a different first barcode sequence segment attached thereto;

pooling the separate bead populations to provide a first pooled bead population;

separating the first pooled bead population into a plurality of second bead populations;

attaching a second oligonucleotide comprising a second barcode sequence segment to the first oligonucleotide already attached to the second bead populations, wherein each of the plurality of second bead populations comprises a different second barcode sequence segment; and

pooling the plurality of second bead populations to provide a second pooled bead population comprising a barcode library.

94. A method of preparing a barcode library, comprising:

providing a first pooled bead population comprising a plurality of different first bead populations, each different first bead population having a different first oligonucleotide attached thereto each different first oligonucleotide comprising a different first barcode sequence segment;

separating the first pooled bead population into a plurality of second bead populations;

attaching a second oligonucleotide comprising a second barcode sequence segment to the first oligonucleotide already attached to the second bead populations, wherein each second bead population comprises a different second barcode sequence segment; and

pooling the second bead populations to provide a second pooled bead population comprising a barcode library.

95. A barcode library, comprising a plurality of different oligonucleotides, each different oligonucleotide comprising:

a first barcode sequence segment selected from a first set of barcode sequence segments;

a second barcode sequence segment selected from a second set of barcode sequence segments; and

a linking sequence joining the first barcode sequence segment and the second barcode sequence segment, the linking sequence having from about 2 to about 6 nucleotides in length and being selected from a set of linking sequences.

96. A method of amplifying a template nucleic acid sequence, comprising:

(a) co-partitioning a template nucleic acid sequence and a bead comprising a plurality of oligonucleotides releasably attached thereto into a partition, the plurality of oligonucleotides comprising a primer sequence complementary to one or more regions of the template nucleic acid sequence and a common sequence;

(b) annealing the primer sequence to the template nucleic acid sequence; and

(c) extending the primer sequence to produce one or more first copies of at least a portion of the template nucleic acid sequence, the one or more first copies comprising the primer sequence and the common sequence.

97. A method of amplifying a plurality of different nucleic acids, comprising:

partitioning different nucleic acids into separate first partitions, each first partition comprising a second partition having a plurality of oligonucleotides releasably associated therewith, the plurality of oligonucleotides associated with a given second partition comprising a variable primer sequence and a barcode sequence, and the plurality of oligonucleotides associated with different second partitions comprising different barcode sequences;

releasing oligonucleotides associated with the plurality of second partitions into the first partitions;

annealing variable primer sequences within the first partitions to nucleic acids within the first partitions; and

extending the variable primer sequences to produce one or more copies of at least a portion of the nucleic acids within the first partitions, the one or more copies comprising the oligonucleotides released into the first partitions.

98. A method of whole genome amplification comprising:

(a) hybridizing a random primer to a genomic nucleic acid, wherein the random primer is attached to a universal nucleic acid sequence and a nucleic acid barcode sequence, wherein the universal nucleic acid sequence comprises one or more uracil containing nucleotides;

(b) extending the random primer to form an amplified product; and

(c) causing the amplified product to undergo an intramolecular hybridization reaction, thereby forming a partial hairpin molecule.

99. A method of amplifying nucleic acids, comprising:

fragmenting a genomic component into a plurality of first fragments;

co-partitioning the first fragments into a plurality of partitions with a plurality of oligonucleotides, the plurality of oligonucleotides within each of the plurality of partitions comprising a primer sequence and a common sequence;

annealing primer sequences in each partition of the plurality of partitions to a plurality of different regions of the first fragments within each partition of the plurality of partitions; and

extending the primer sequences along the plurality of first fragments to produce amplified first fragments within each partition of the plurality of partitions.

100. A method of characterizing a nucleic acid segment, comprising:

co-partitioning the nucleic acid segment and a bead comprising a plurality of oligonucleotides that comprise a common nucleic acid barcode sequence into a partition;

attaching the oligonucleotides to fragments of the nucleic acid segment or to copies of portions of the nucleic acid segment, thereby attaching the common nucleic acid barcode sequence to the fragments of the nucleic acid segment or to the copies of the portions of the nucleic acid segment;

sequencing the fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment and the attached common nucleic acid barcode sequence; and

characterizing the fragments or the copies of the portions of the nucleic acid segment as being linked within the nucleic acid segment based at least in part, upon their attachment to the common nucleic acid barcode sequence.

101. A method of characterizing a target nucleic acid, comprising:

partitioning first fragments of the target nucleic acid into a plurality of droplets, each droplet of the plurality of droplets comprising a bead having a plurality of oligonucleotides attached thereto, the oligonucleotides attached to a given bead comprising a common barcode sequence;

attaching the common barcode sequence to second fragments of the first fragments;

pooling the plurality of droplets;

sequencing the second fragments and attached barcode sequences; and

mapping a plurality of the second fragments to one or more of the first fragments based, at least in part, upon the plurality of second fragments comprising a common barcode sequence.

102. A method of sequencing nucleic acids, comprising:

providing a plurality of target nucleic acid sequences;

separating the target nucleic acid sequences into a plurality of separate partitions, each partition of the plurality of separate partitions comprising one or more target nucleic acid sequences, and a bead having a plurality of oligonucleotides attached thereto, the oligonucleotides attached to a given bead comprising a common barcode sequence;

attaching the oligonucleotides to fragments of the one or more target nucleic acid sequences or to copies of portions of the one or more target nucleic acid sequences within a partition, thereby attaching the common barcode sequence to the fragments of the one or more target nucleic acid sequences or to the copies of the portions of the one or more target nucleic acid sequences;

pooling the plurality of separate partitions; and

sequencing the fragments of the one or more target nucleic acid sequences or the copies of the portions of the one or more target nucleic acid sequences and attached barcode sequences to provide barcoded fragment sequences or barcoded copy sequences.

103. A method of characterizing a nucleic acid segment, comprising:

co-partitioning the nucleic acid segment, and a bead comprising a plurality of oligonucleotides that comprise a common nucleic acid barcode sequence, into a first droplet;

attaching the oligonucleotides to fragments of the nucleic acid segment or to copies of portions of the nucleic acid segment, thereby attaching the common nucleic acid barcode sequence to the fragments of the nucleic acid segment or to the copies of the portions of the nucleic acid segment;

sequencing the fragments of the nucleic acid segment or the copies of the portions of the nucleic acid segment and attached common nucleic acid barcode sequence to provide a plurality of barcoded fragment sequences or a plurality of barcoded copy sequences; and

assembling the plurality of the barcoded fragment sequences or the plurality barcoded copy sequences into one or more contiguous nucleic acid sequences based at least in part on the common nucleic acid barcode sequence.

104. A method of sequencing nucleic acids, comprising:

providing a plurality of target nucleic acid sequences;

separating the target nucleic acid sequences into a plurality of separate partitions, each partition of the plurality of separate partitions comprising one or more target nucleic acid sequences, and a plurality of oligonucleotides, the oligonucleotides in a given partition comprising a common barcode sequence, and wherein the plurality of separate partitions comprise at least 10,000 different barcode sequences;

attaching the common barcode sequence in each partition to fragments of the one or more target nucleic acid sequences or to copies of portions of the one or more target nucleic acid sequences within the partition;

pooling the plurality of separate partitions; and

sequencing the fragments of the one or more target nucleic acid sequences or the copies of the portions of the one or more target nucleic acid sequences and attached barcode sequences.

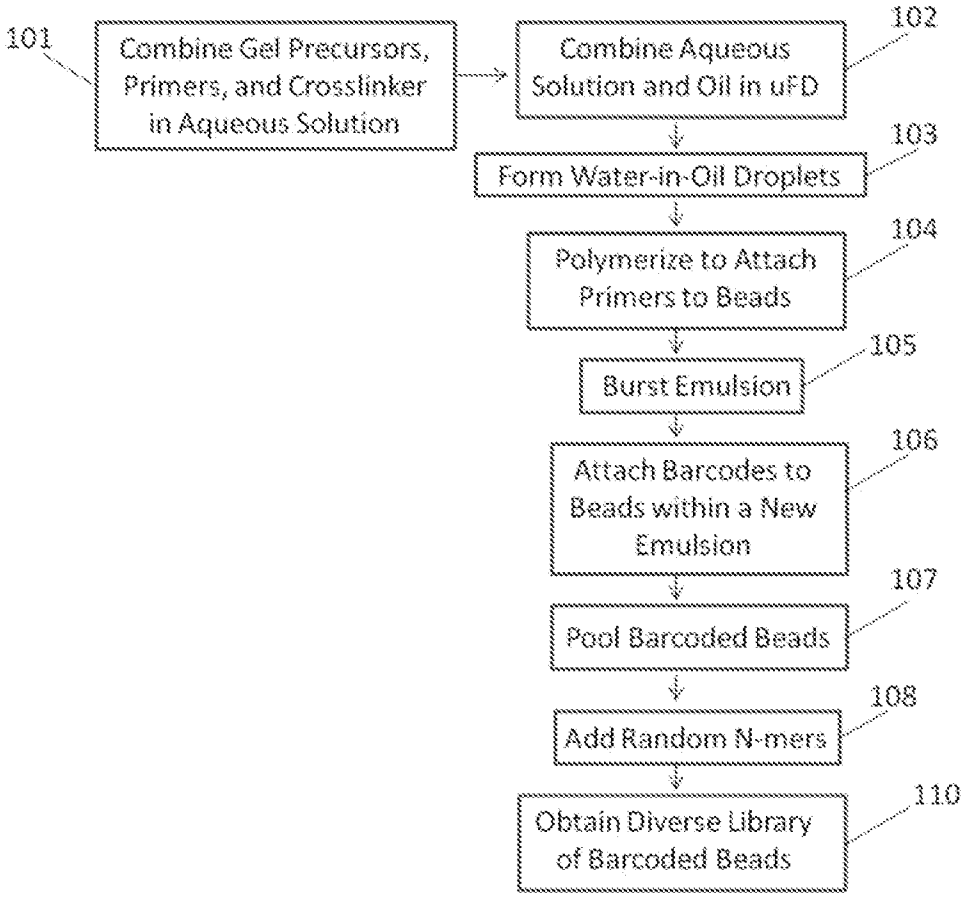

101
Combine Gel Precursors, Primers, and Crosslinker in Aqueous Solution
102
Combine Aqueous Solution and Oil in uFD
103
Form Water-in-Oil Droplets
104
Polymerize to Attach Primers to Beads
105
Burst Emulsion
106
Attach Barcodes to Beads within a New Emulsion
107
Pool Barcoded Beads
108
Add Random N-mers
110
Obtain Diverse Library of Barcoded Beads

Figure 1A.

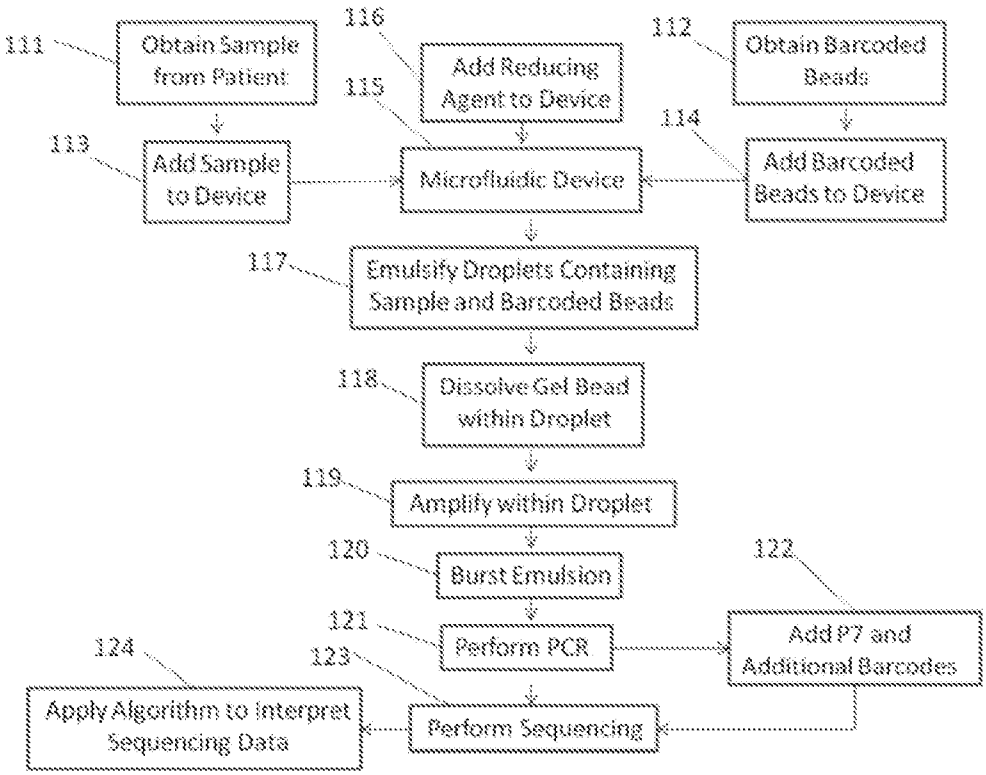

111
Obtain Sample from Patient
116
Add Reducing Agent to Device
112
Obtain Barcoded Beads
115
114
113
Add Sample to Device
Microfluidic Device
Add Barcoded Beads to Device
117
Emulsify Droplets Containing Sample and Barcoded Beads
118
Dissolve Gel Bead within Droplet
119
Amplify within Droplet
120
Burst Emulsion
122
121
Perform PCR
Add P7 and Additional Barcodes
124
123
Apply Algorithm to Interpret Sequencing Data
Perform Sequencing

Figure 1B.

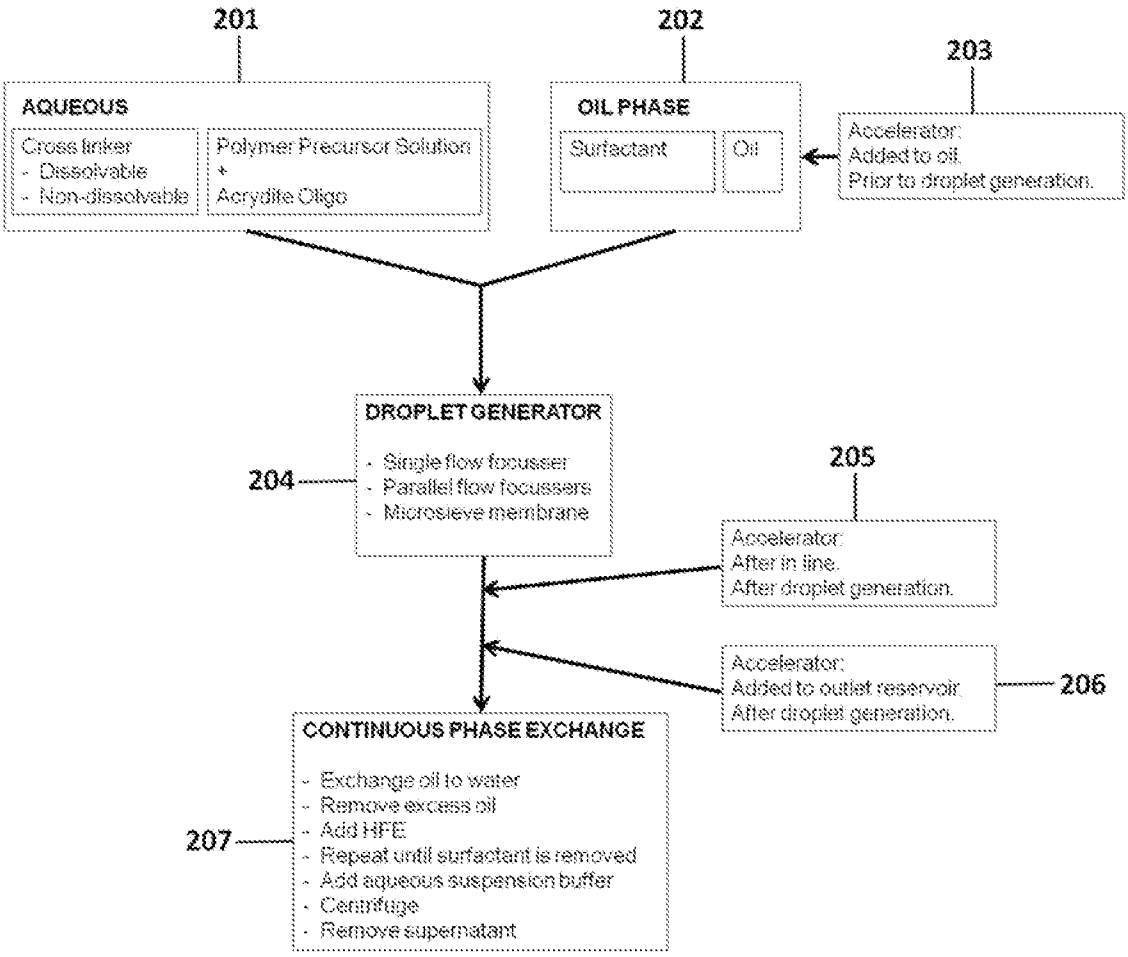

201
202
203
AQUEOUS
Cross linker
- Dissolvable
- Non-dissolvable
Polymer Precursor Solution
+
Acrydite Oligo
OIL PHASE
Surfactant
Oil
Accelerator:
Added to oil.
Prior to droplet generation.
DROPLET GENERATOR
- Single flow focusser
- Parallel flow focussers
- Microsieve membrane
204
205
Accelerator:
After in line.
After droplet generation.
Accelerator:
Added to outlet reservoir
After droplet generation.
206
CONTINUOUS PHASE EXCHANGE
- Exchange oil to water
- Remove excess oil
- Add HFE
- Repeat until surfactant is removed
- Add aqueous suspension buffer
- Centrifuge
- Remove supernatant
207

Figure 2.

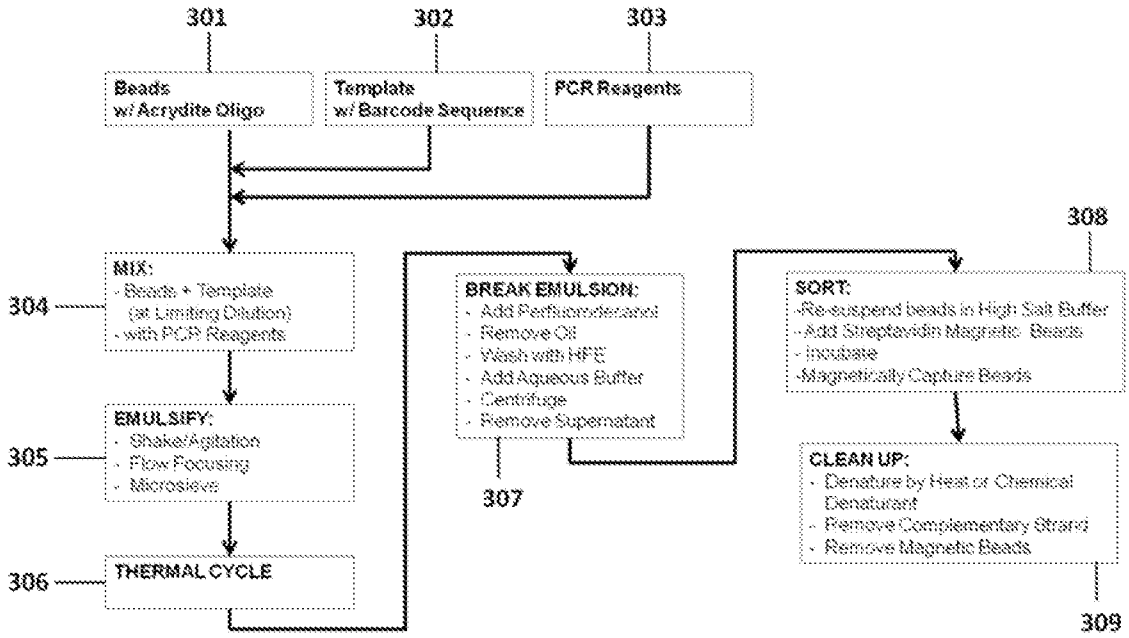

301
302
303
Beads
w/ Acrydite Oligo
Template
w/ Barcode Sequence
PCR Reagents
MIX:
- Beads + Template
(at Limiting Dilution)
- with PCR Reagents
304
EMULSIFY:
- Shake/Agitation
- Flow Focusing
- Microsieve
305
THERMAL CYCLE
306
BREAK EMULSION:
- Add Perfluorodecanol
- Remove Oil
- Wash with HFE
- Add Aqueous Buffer
- Centrifuge
- Remove Supernatant
307
308
SORT:
-Re-suspend beads in High Salt Buffer
- Add Streptavidin Magnetic Beads
- Incubate
-Magnetically Capture Beads
CLEAN UP:
- Denature by Heat or Chemical Denaturant
- Remove Complementary Strand
- Remove Magnetic Beads
309

Figure 3A.

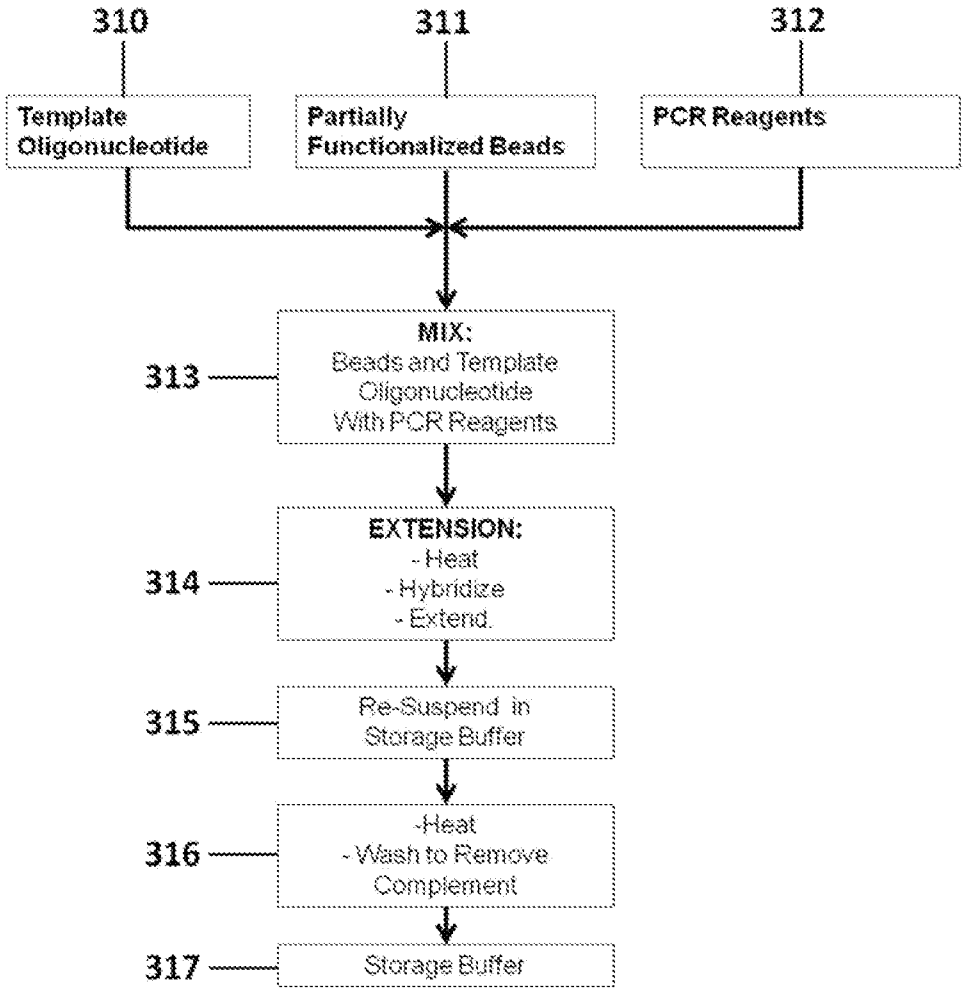

310
311
312
Template Oligonucleotide
Partially Functionalized Beads
PCR Reagents
313
MIX:
Beads and Template Oligonucleotide With PCR Reagents
314
EXTENSION:
- Heat
- Hybridize
- Extend.
315
Re-Suspend in Storage Buffer
316
-Heat
- Wash to Remove Complement
317
Storage Buffer

Figure 3B.

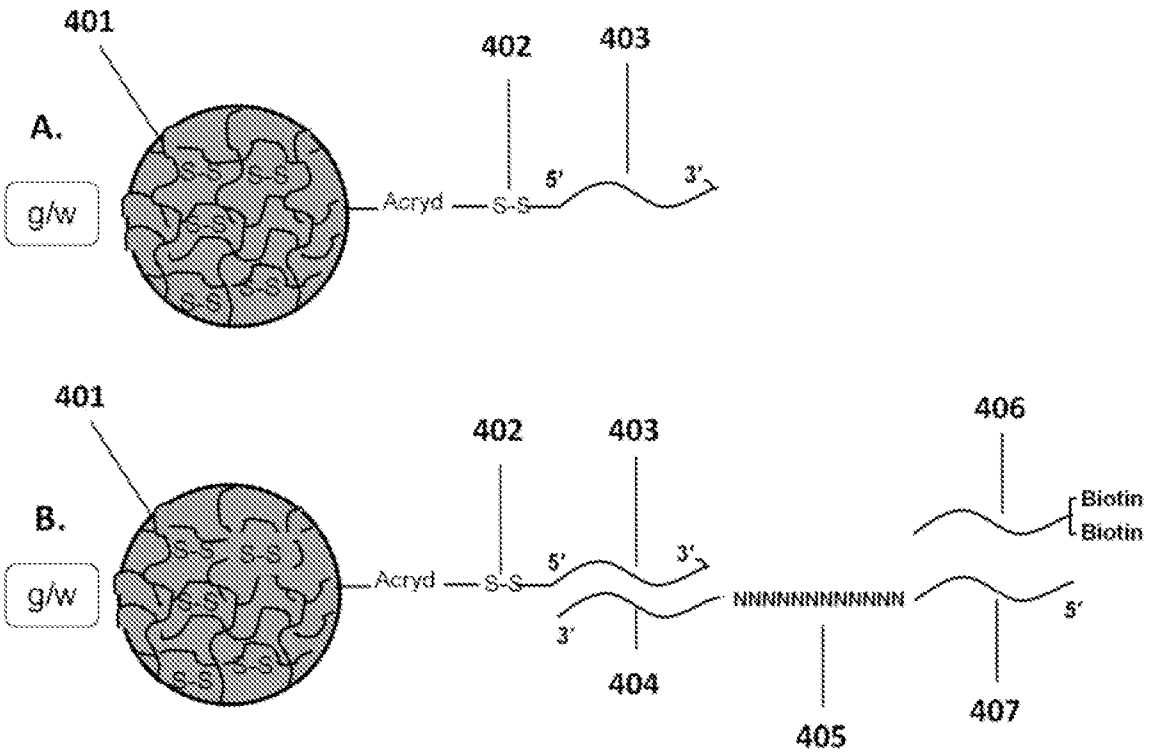

401
402
403
A.
g/w
Acryd
S-S
5'
3'
401
406
402
403
B.
Biotin
Biotin
g/w
Acryd
S-S
5'
3'
NNNNNNNNNNNN
5'
3'
404
405
407

Figure 4.

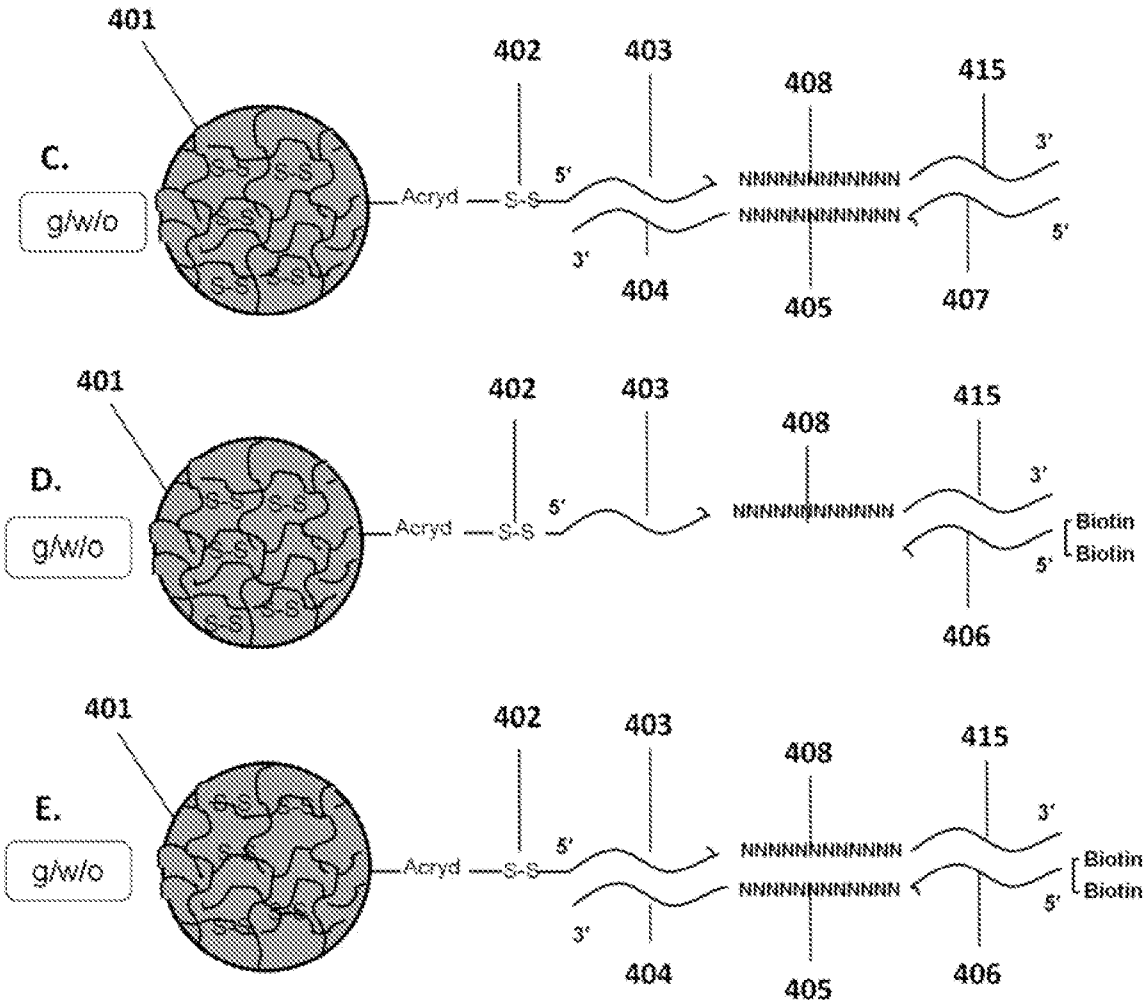

C.
g/w/o
401
402
403
404
405
407
408
415
Acryd
S-S
5'
3'
NNNNNNNNNNN
D.
406
Biotin
E.

Figure 4.

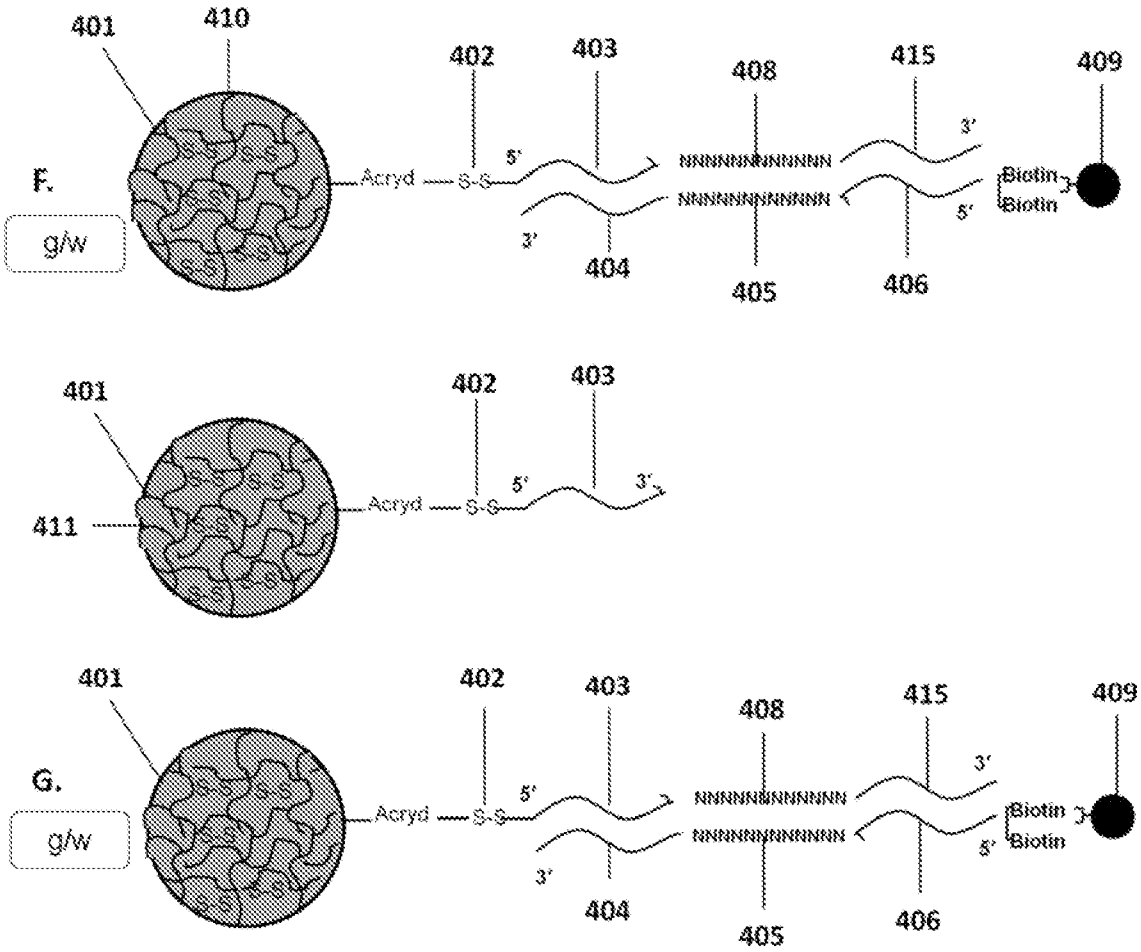

F.
g/w
401
410
402
403
408
415
409
Acryd
S-S
5'
3'
NNNNNNNNNNNN
NNNNNNNNNNNN
Biotin
Biotin
404
405
406
411
G.

Figure 4.

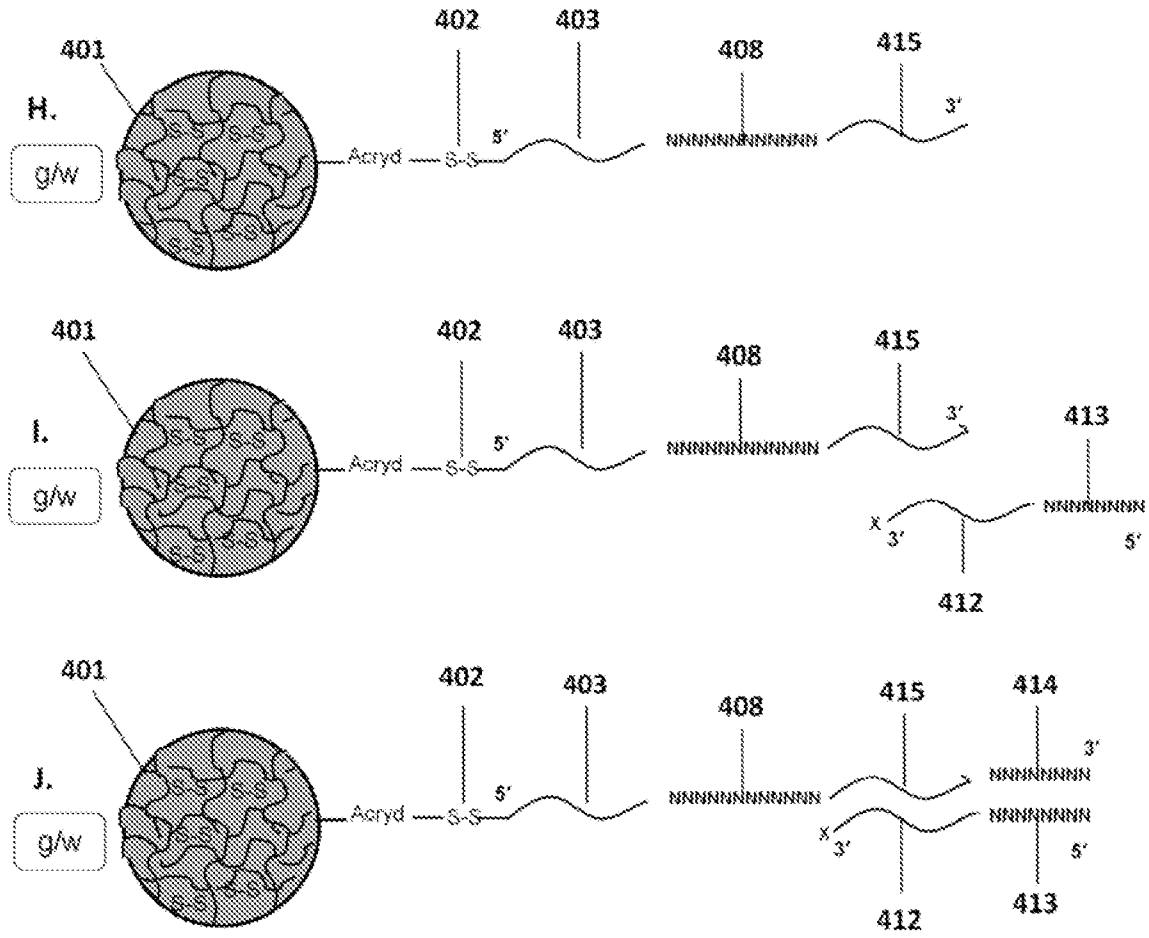

H.
401
g/w
Acryd
S-S
5'
402
403
408
415
3'
NNNNNNNNNNNN
I.
401
g/w
Acryd
S-S
5'
402
403
408
415
3'
413
NNNNNNNNN
5'
X
3'
412
J.
401
g/w
Acryd
S-S
5'
402
403
408
415
414
3'
NNNNNNNNN
X
3'
5'
412
413

Figure 4.

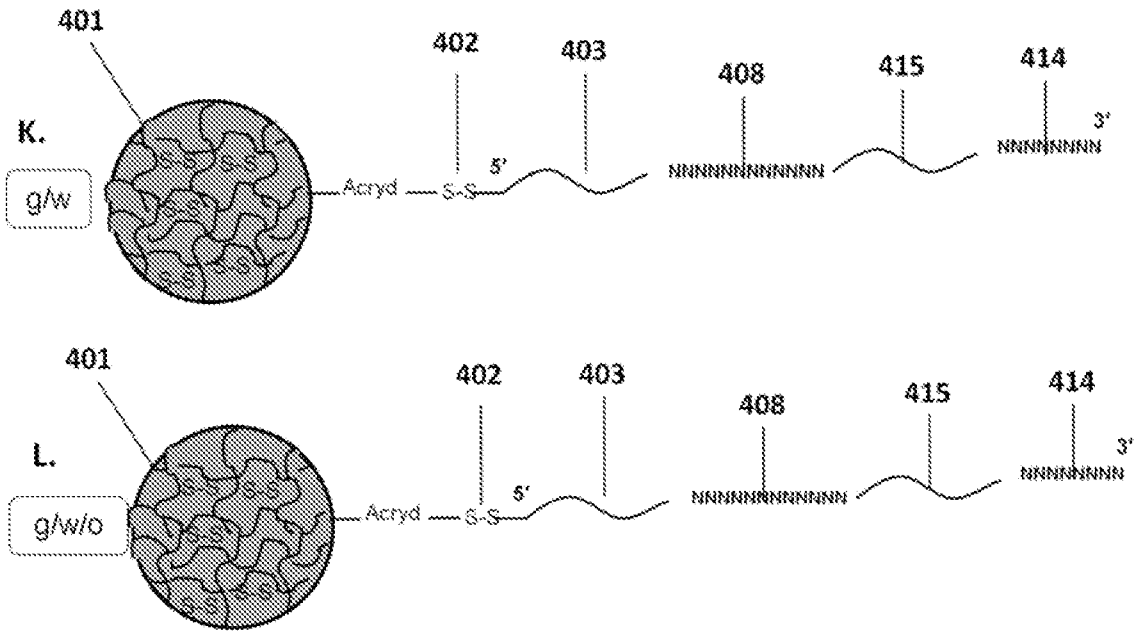

401
402
403
408
415
414
K.
g/w
Acryd
S-S
5'
NNNNNNNNNNN
NNNNNNNN
3'
401
402
403
408
415
414
L.
g/w/o
Acryd
S-S
5'
NNNNNNNNNNN
NNNNNNNN
3'

Figure 4.

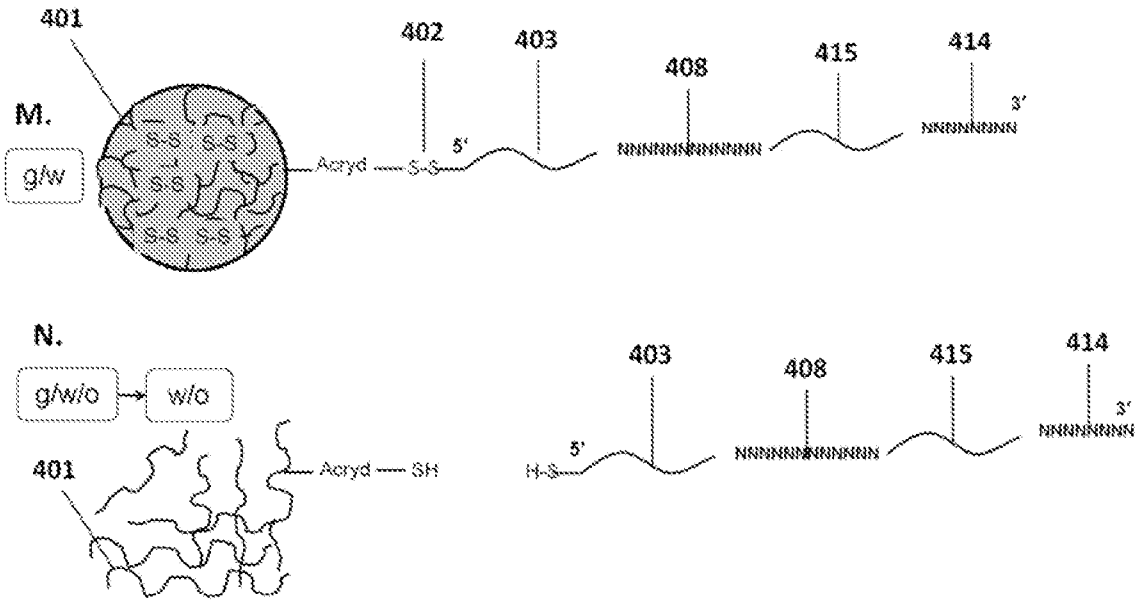

401
402
403
408
415
414
M.
g/w
Acryd
5'
3'
N.
g/w/o
w/o
401
Acryd — SH
403
408
415
414
5'
3'

Figure 4.

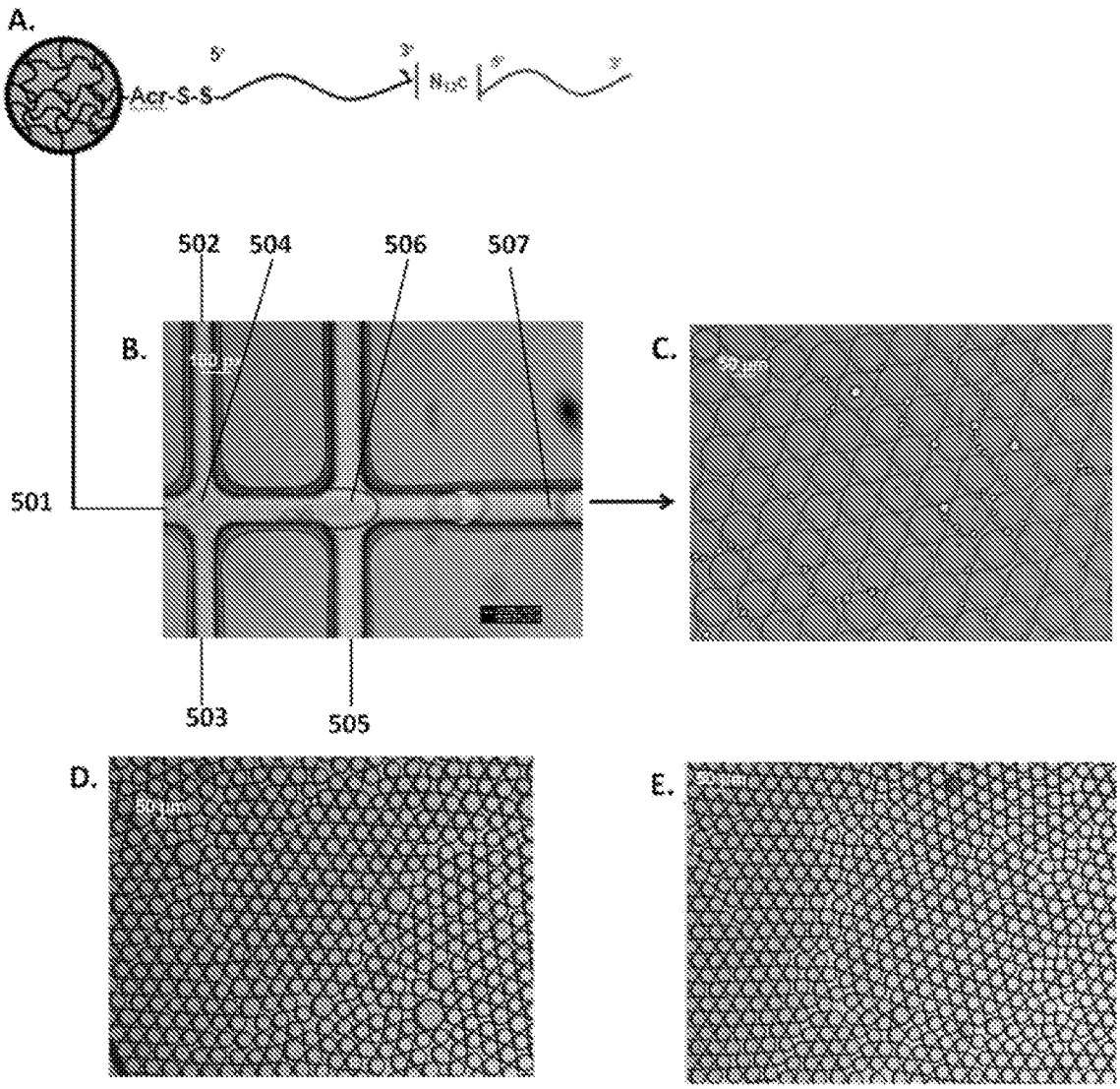


Figure 5.

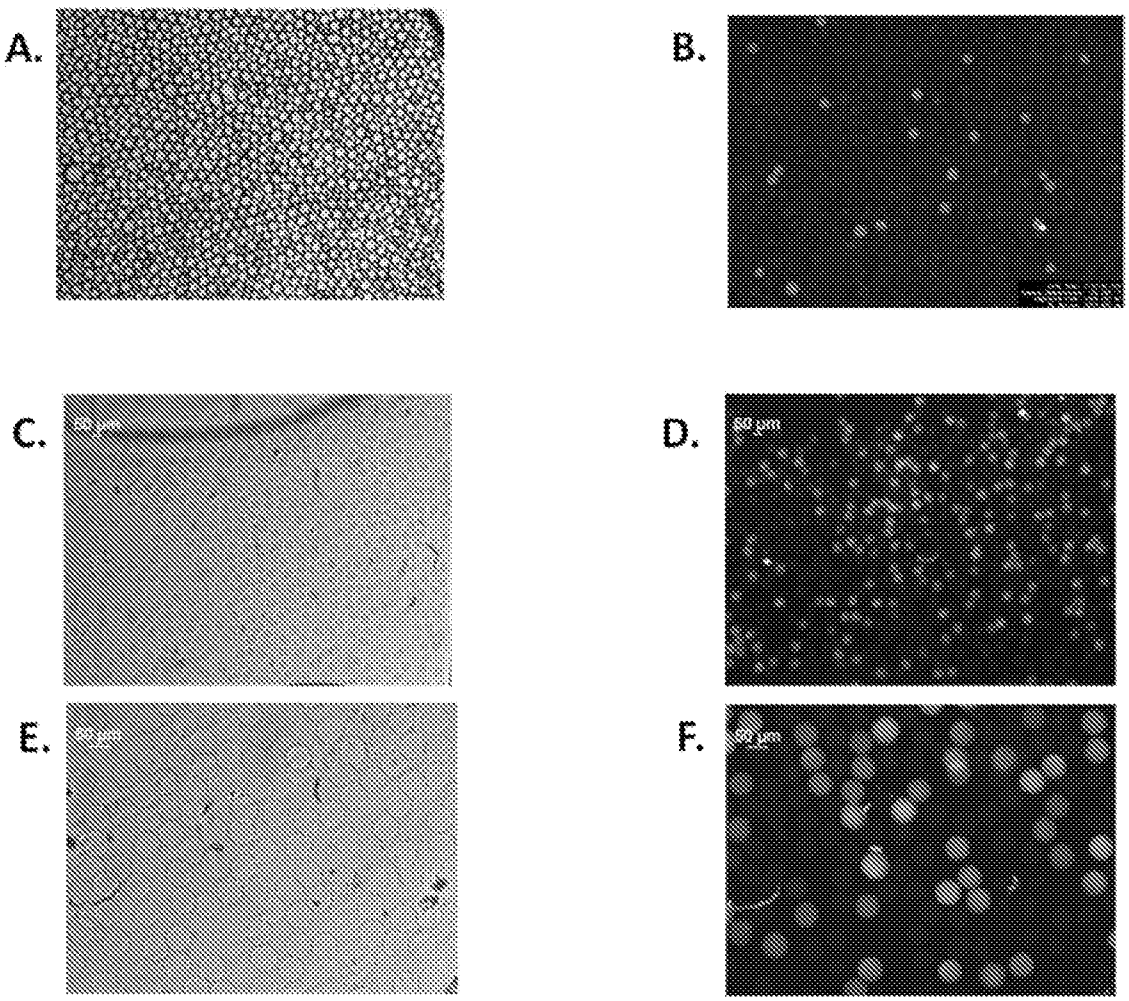


Figure 6.

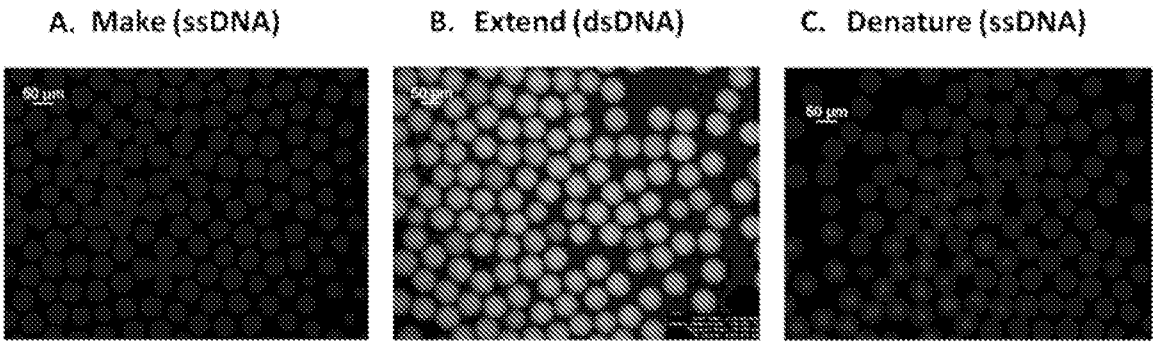


Figure 7.

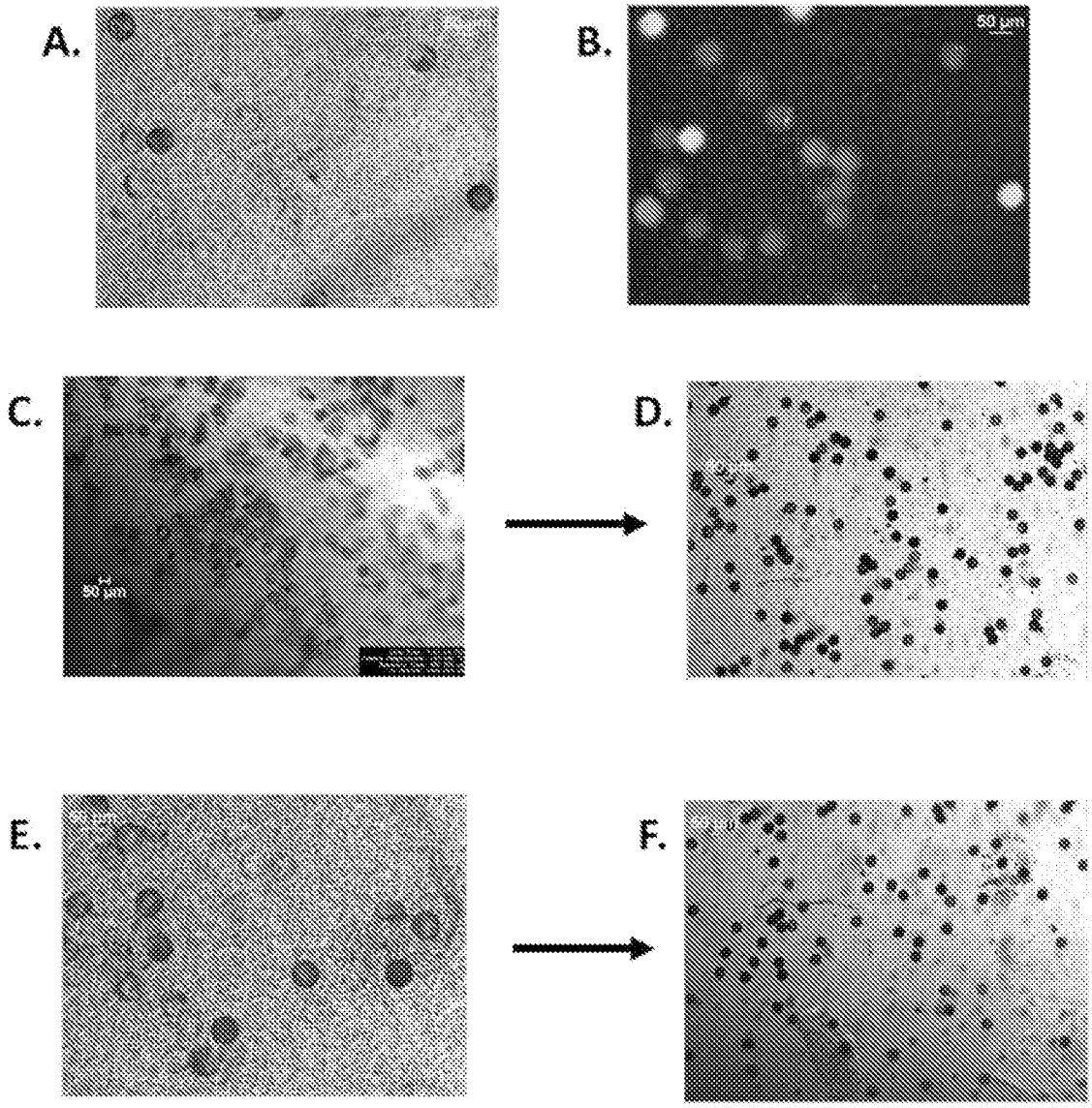


Figure 8.

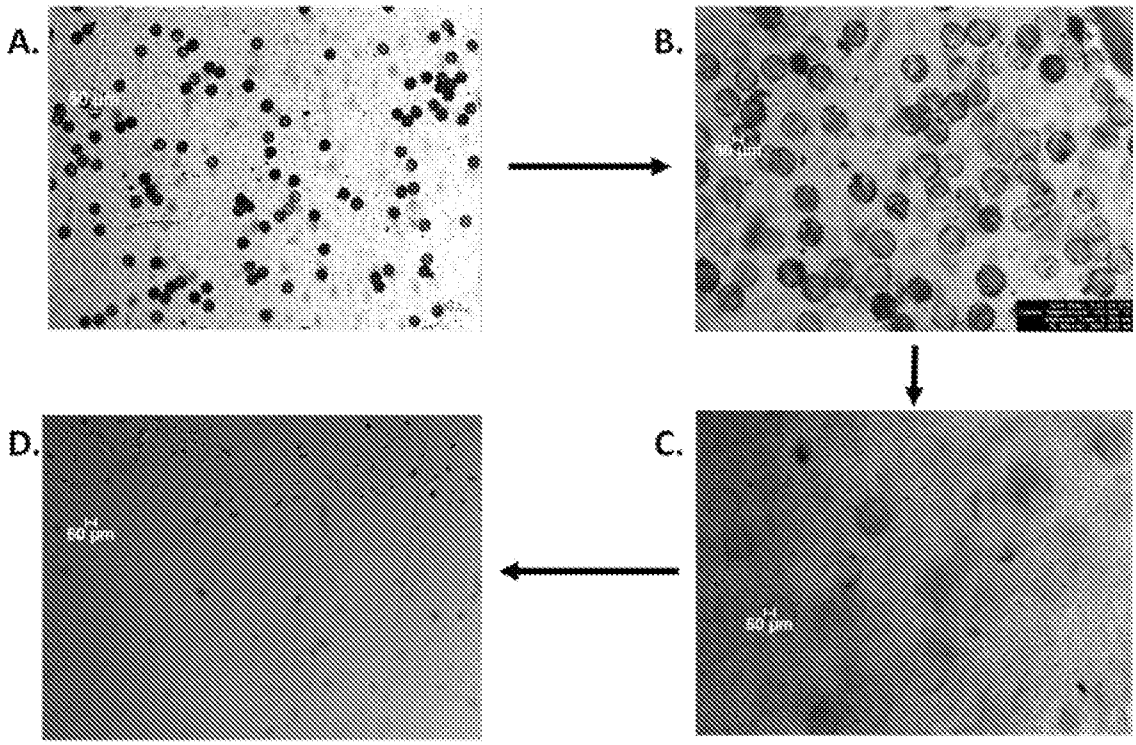


Figure 9.

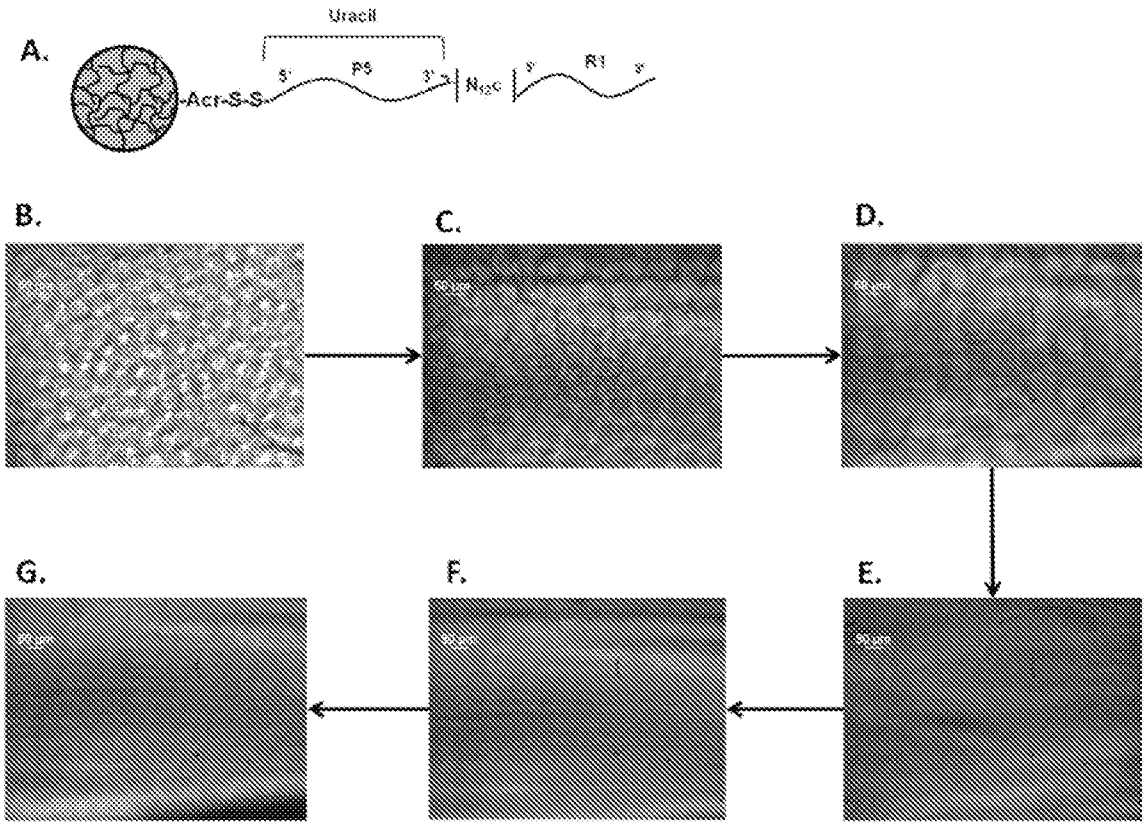


Figure 10.

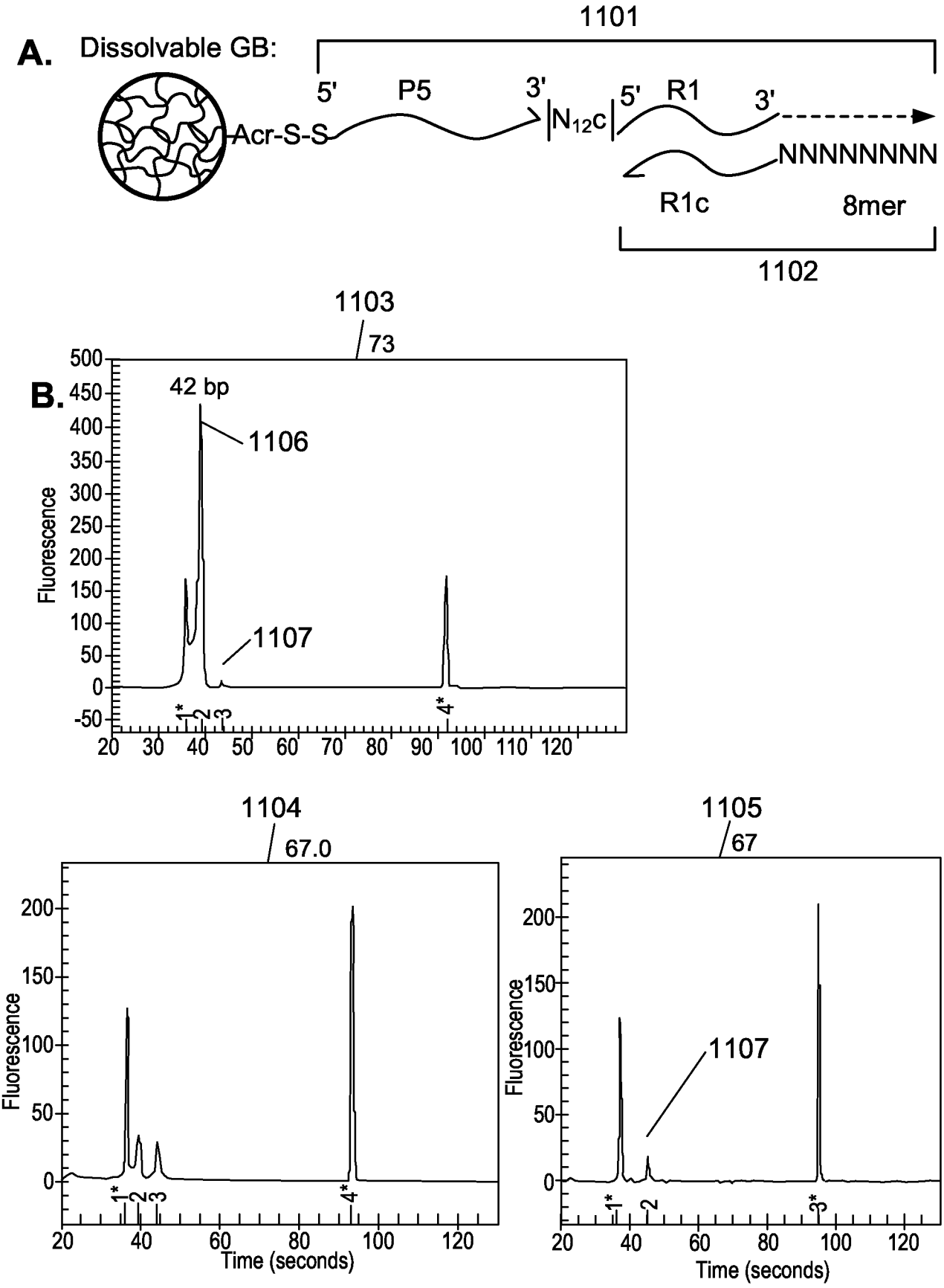
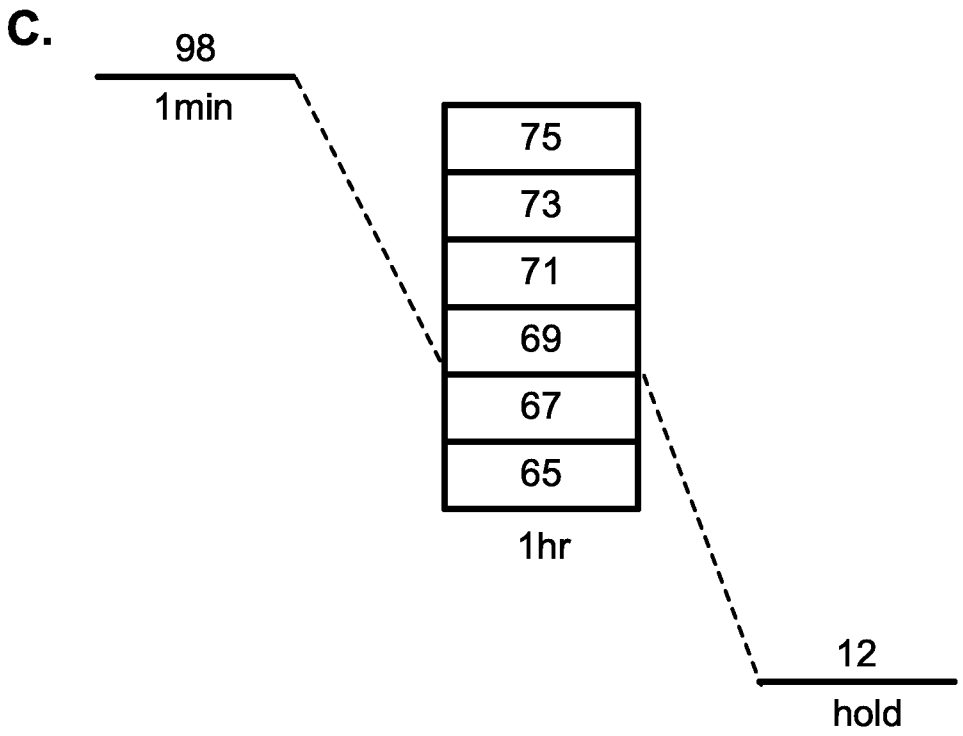


Figure 11. (continued)


C.
98
1min
75
73
71
69
67
65
1hr
12
hold

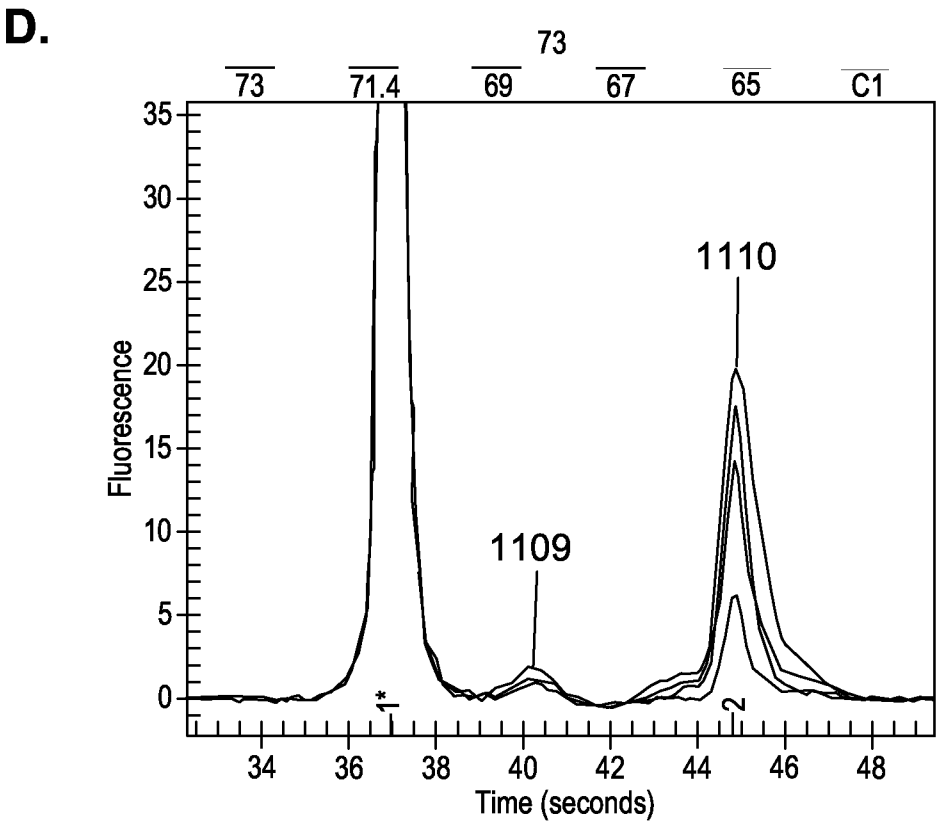

D.
73
73
71.4
69
67
65
C1
35
30
25
20
15
10
5
0
Fluorescence
1110
1109
1*
2
34
36
38
40
42
44
46
48
Time (seconds)

Figure 11.

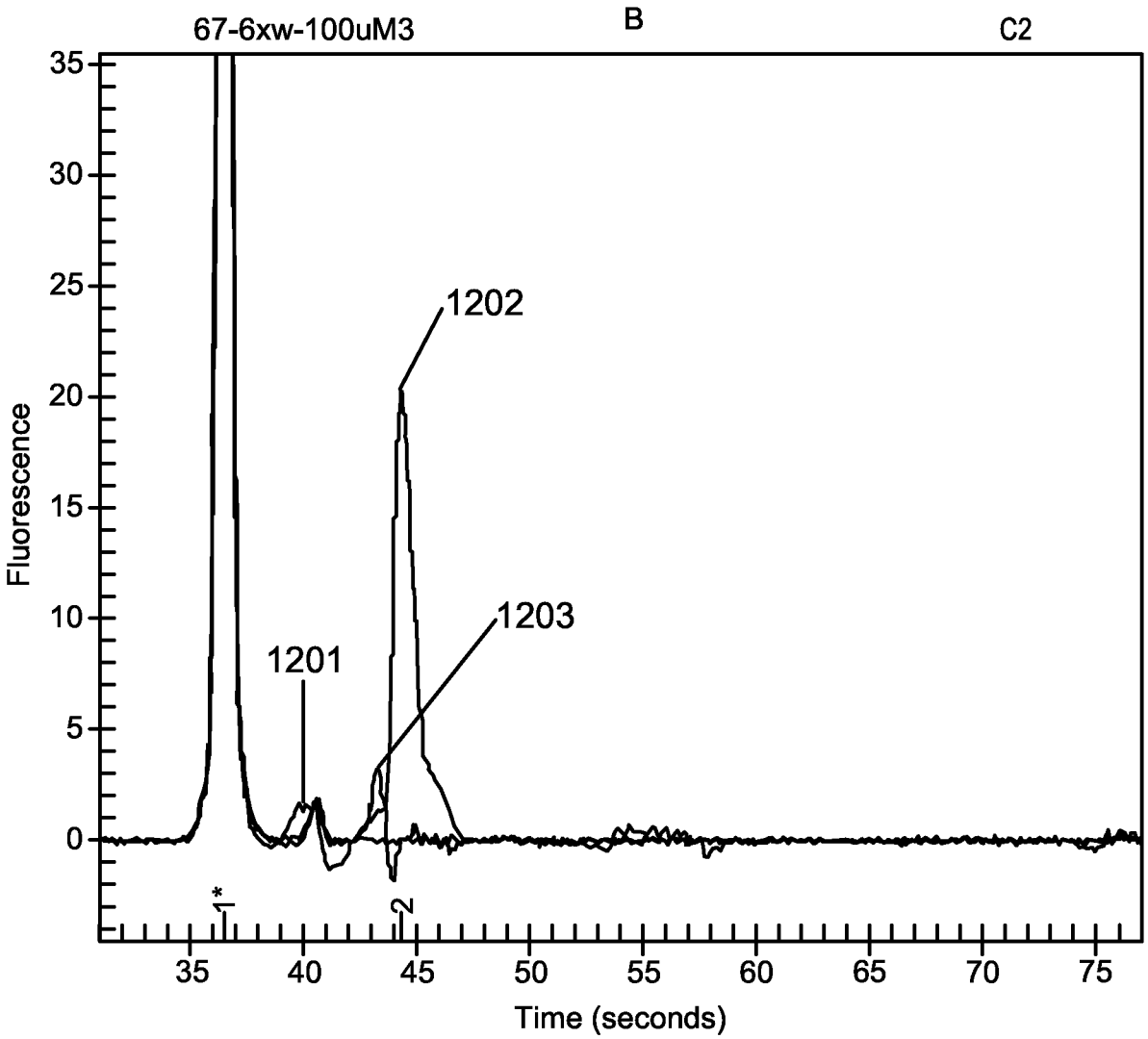

67-6xw-100uM3
B
C2
Fluorescence
35
30
25
20
15
10
5
0
1202
1203
1201
*1
2
35
40
45
50
55
60
65
70
75
Time (seconds)

Figure 12.

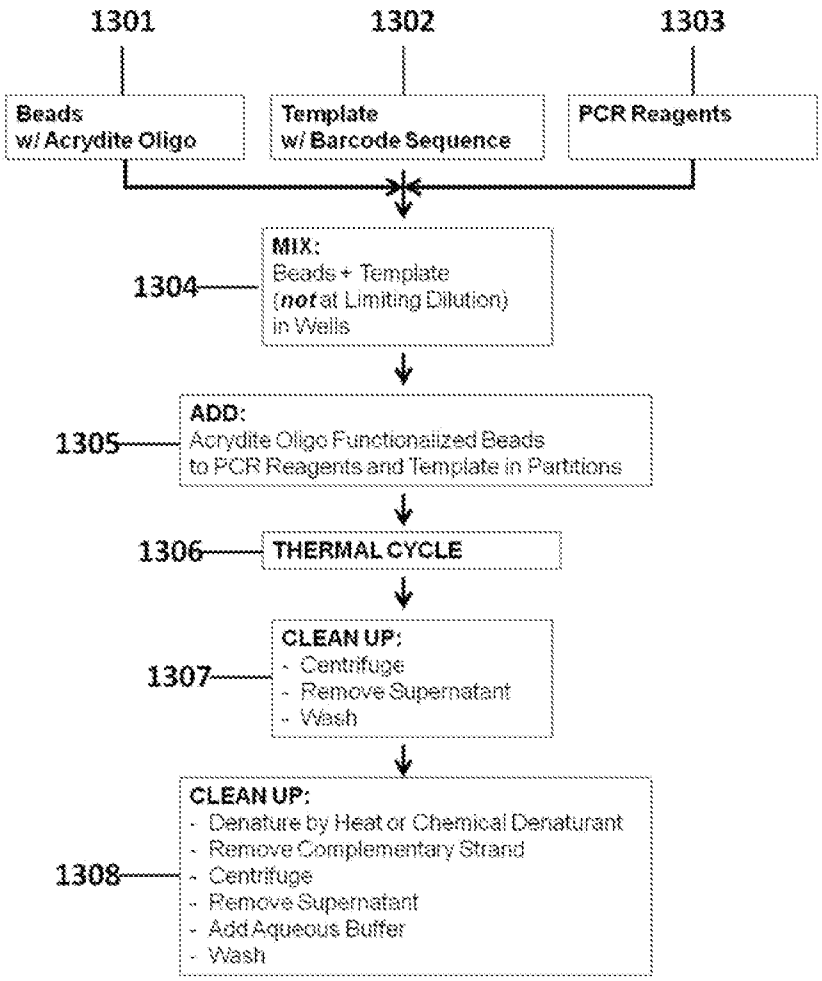

1301
1302
1303
Beads w/ Acrydite Oligo
Template w/ Barcode Sequence
PCR Reagents
1304
MIX:
Beads + Template
(not at Limiting Dilution)
in Wells
1305
ADD:
Acrydite Oligo Functionalized Beads
to PCR Reagents and Template in Partitions
1306
THERMAL CYCLE
1307
CLEAN UP:
- Centrifuge
- Remove Supernatant
- Wash
1308
CLEAN UP:
- Denature by Heat or Chemical Denaturant
- Remove Complementary Strand
- Centrifuge
- Remove Supernatant
- Add Aqueous Buffer
- Wash

Figure 13A.

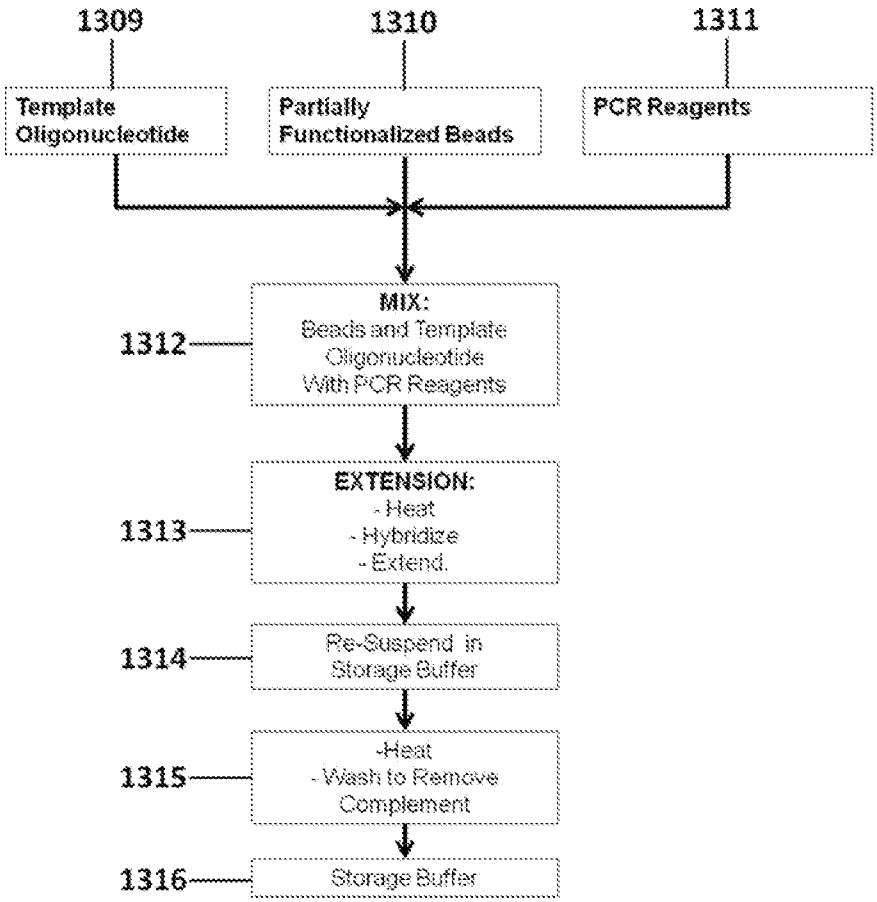

1309
1310
1311
Template Oligonucleotide
Partially Functionalized Beads
PCR Reagents
1312
MIX:
Beads and Template Oligonucleotide With PCR Reagents
1313
EXTENSION:
- Heat
- Hybridize
- Extend.
1314
Re-Suspend in Storage Buffer
1315
-Heat
- Wash to Remove Complement
1316
Storage Buffer

Figure 13B.

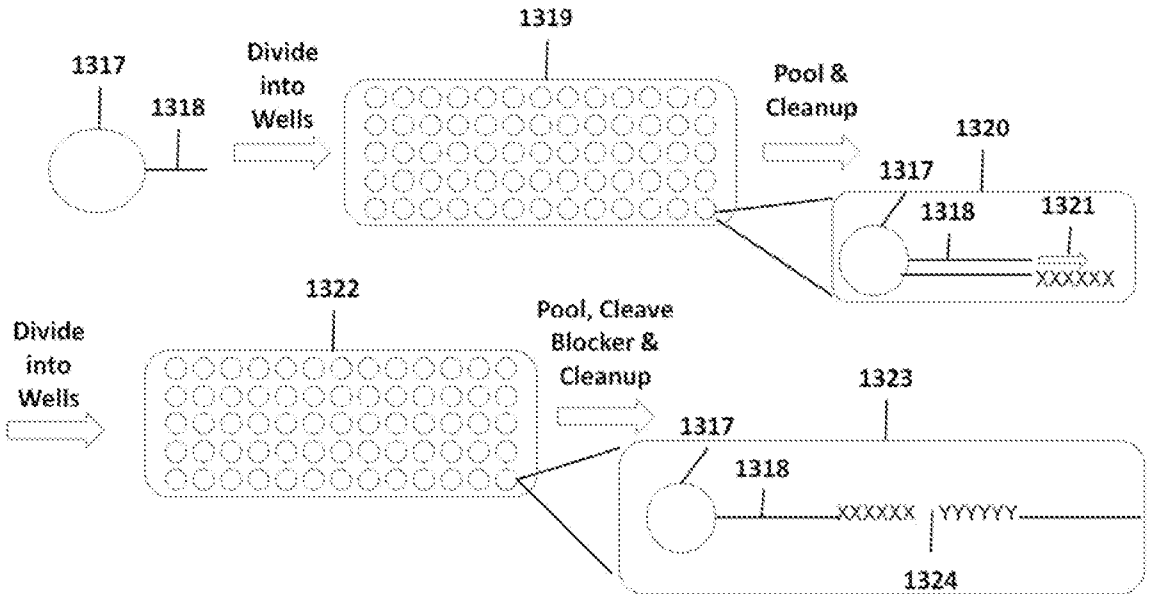

1317
1318
Divide into Wells
1319
Pool & Cleanup
1320
1317
1318
1321
XXXXXX
1322
Divide into Wells
Pool, Cleave Blocker & Cleanup
1323
1317
1318
XXXXXX
YYYYYY
1324

Figure 13C.

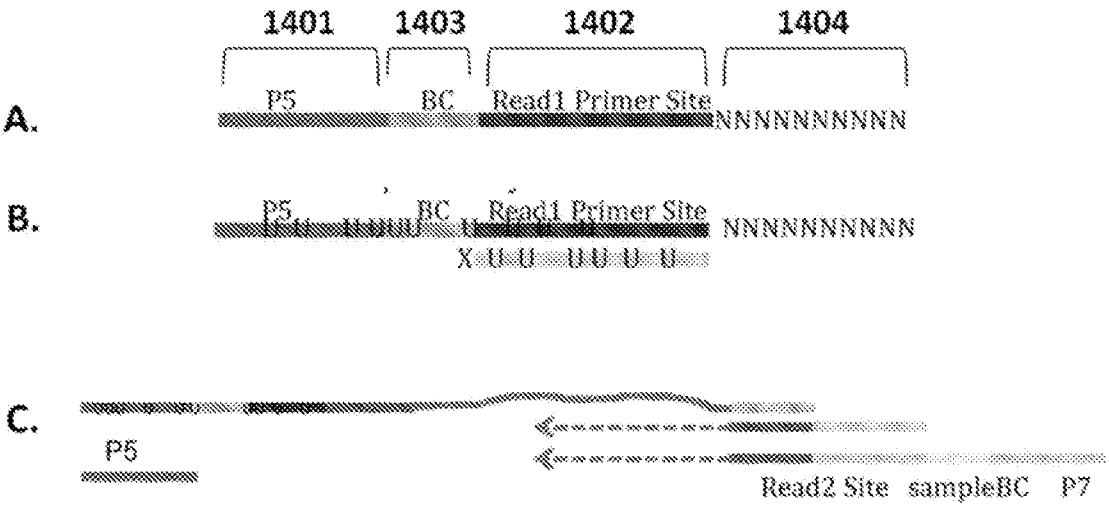

1401
1403
1402
1404
A.
P5
BC
Read1 Primer Site
NNNNNNNNNN
B.
P5
BC
Read1 Primer Site
NNNNNNNNNN
X U U U U U U
C.
P5
Read2 Site
sampleBC
P7

Figure 14.

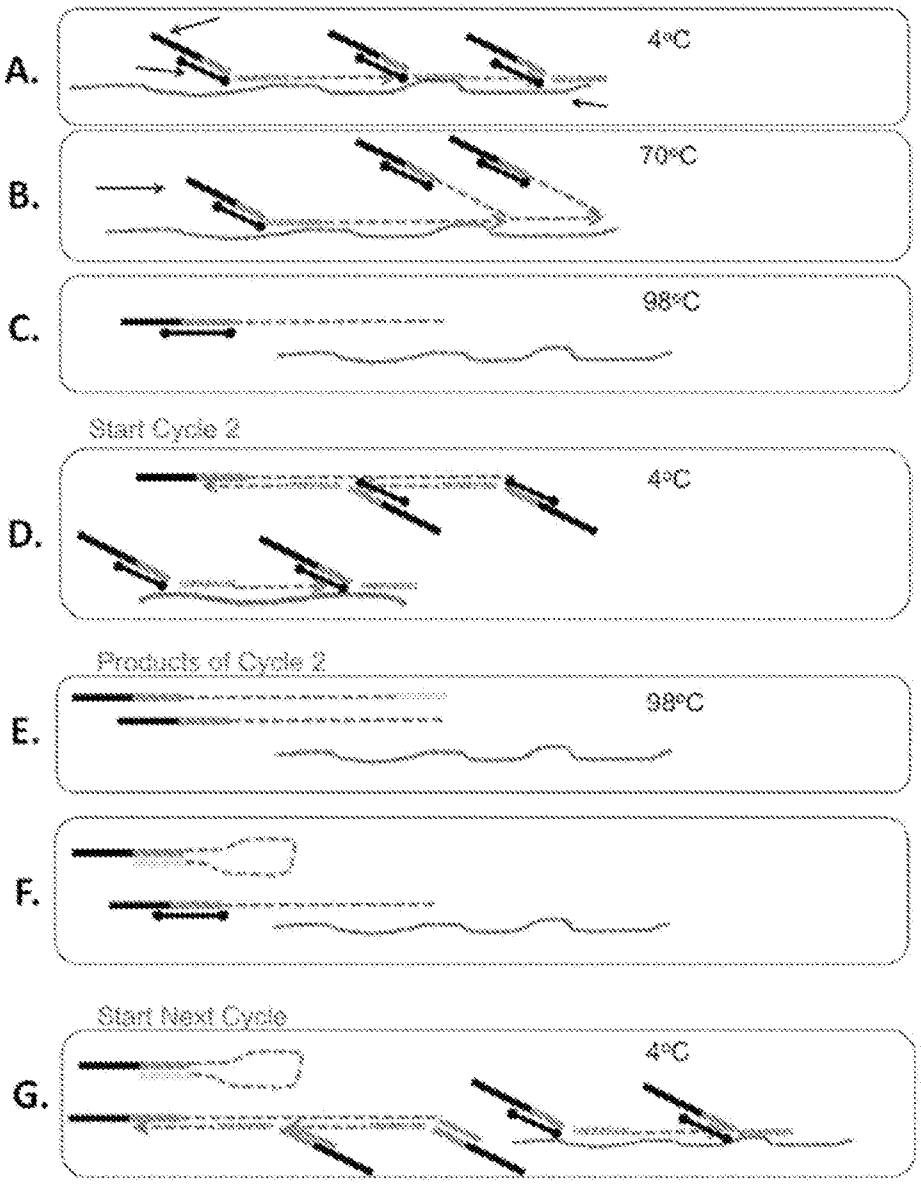
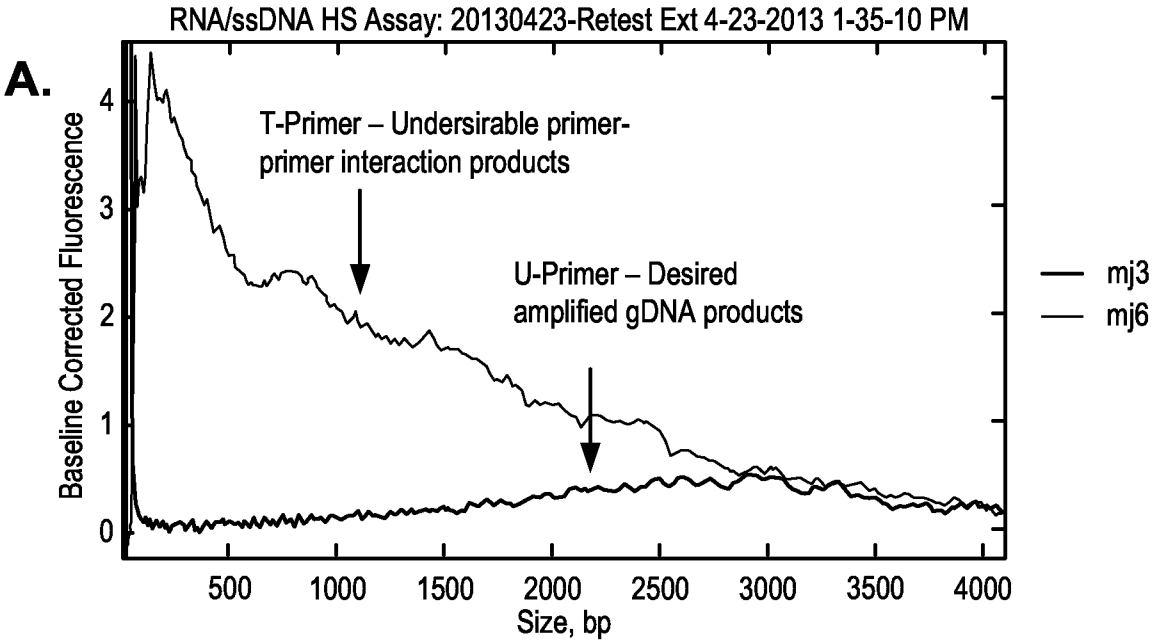

A.
4°C
B.
70°C
C.
98°C
Start Cycle 2
D.
4°C
Products of Cycle 2
E.
98°C
F.
Start Next Cycle
G.
4°C

Figure 15.


A.
RNA/ssDNA HS Assay: 20130423-Retest Ext 4-23-2013 1-35-10 PM
T-Primer – Undersirable primer-primer interaction products
U-Primer – Desired amplified gDNA products
Baseline Corrected Fluorescence
Size, bp
mj3
mj6

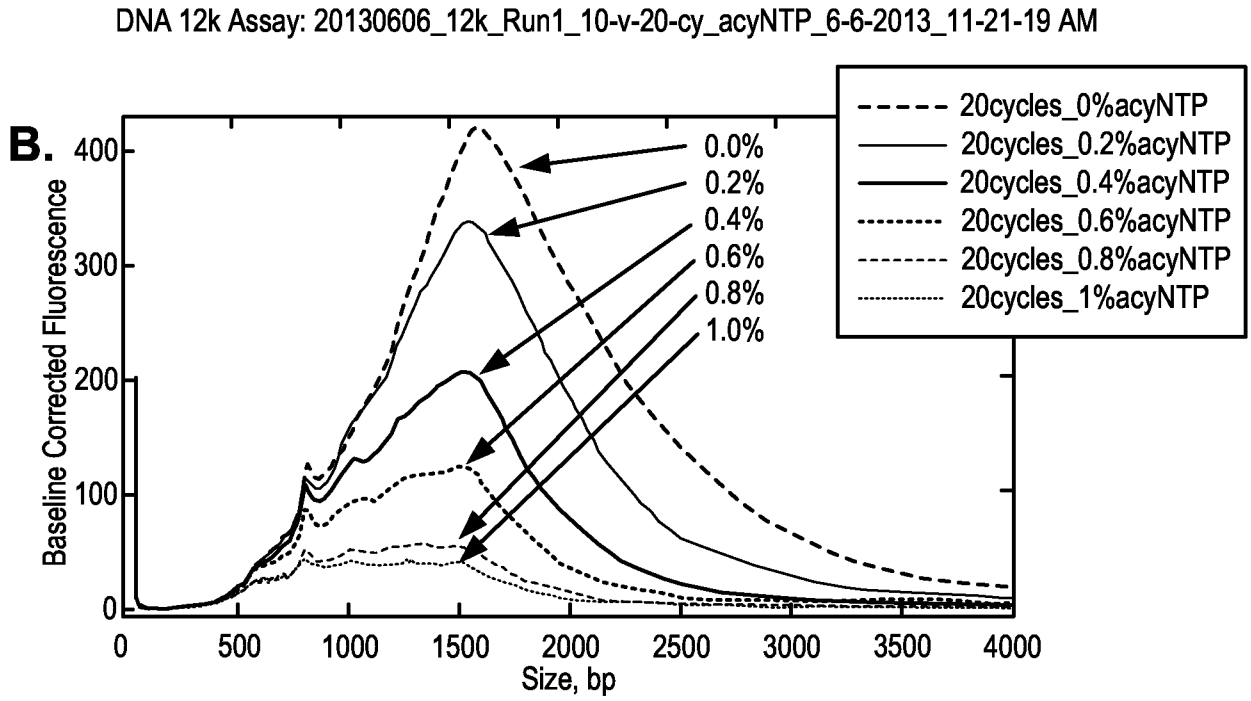

B.
DNA 12k Assay: 20130606_12k_Run1_10-v-20-cy_acyNTP_6-6-2013_11-21-19 AM
0.0%
0.2%
0.4%
0.6%
0.8%
1.0%
20cycles_0%acyNTP
20cycles_0.2%acyNTP
20cycles_0.4%acyNTP
20cycles_0.6%acyNTP
20cycles_0.8%acyNTP
20cycles_1%acyNTP
Baseline Corrected Fluorescence
Size, bp

Figure 16.

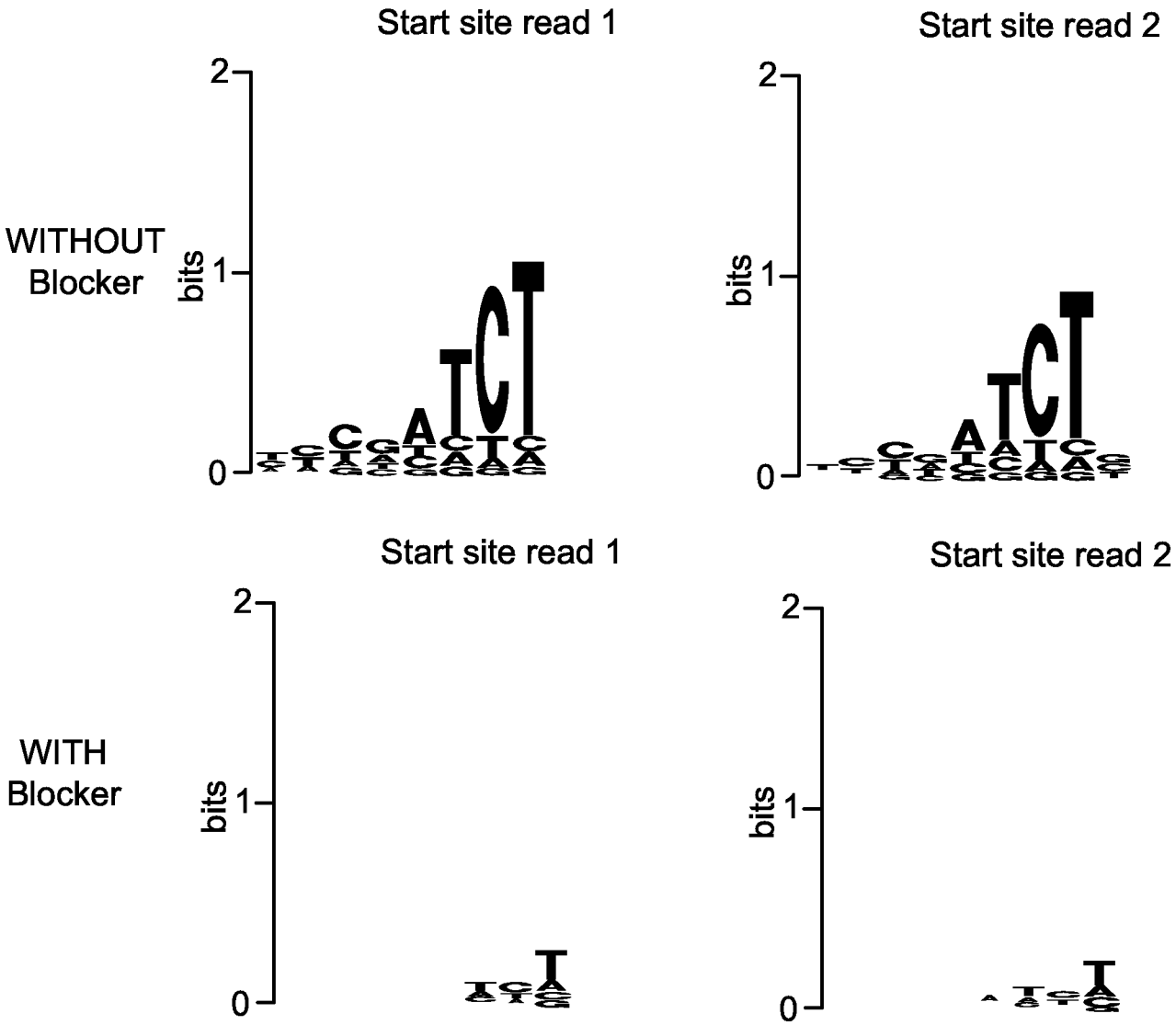

Start site read 1
Start site read 2
WITHOUT Blocker
WITH Blocker
bits
2
1
0

Figure 17.

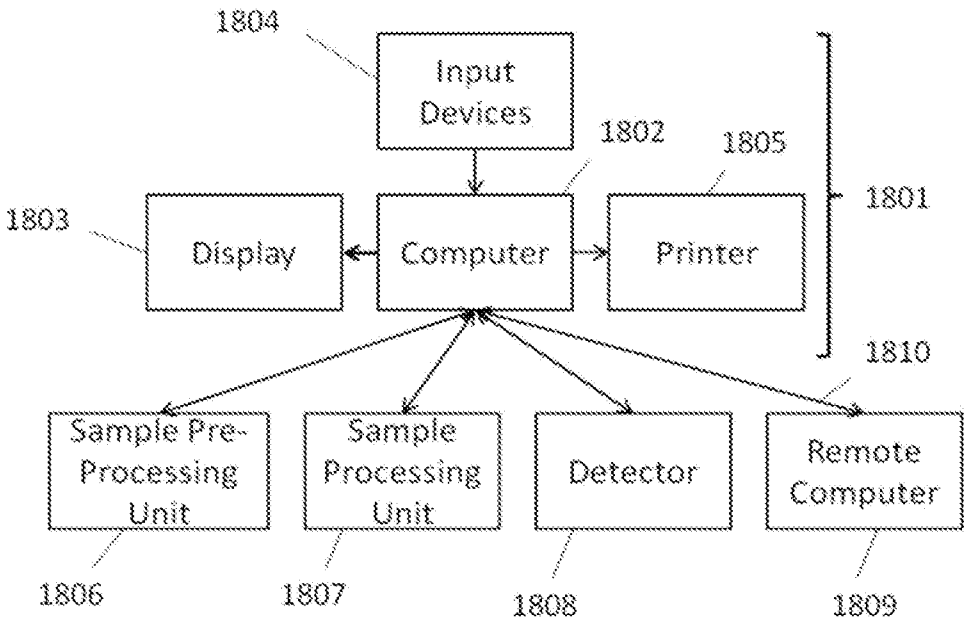

1804
Input Devices
1802
1805
1801
1803
Display
Computer
Printer
1810
Sample Pre-Processing Unit
Sample Processing Unit
Detector
Remote Computer
1806
1807
1808
1809

Figure 18.

Function	Name	Sequence 5'-3'	Mods
Primer for actual library production	P5_U	/5Acryd/iThioMC6-D/ AAU GAU ACG GCG ACC ACC GAG A	["5' Acrydite", "iThioMC6-D", 'Int deoxyUridine', 'HPLC Purification']
Reverse primer for sorting	S2-Bio-R1-rc	/5'3-Bio/ GCG TCG TGT AGG GAA AGA GTG T	["5' Dual Biotin", 'HPLC Purification']
Limiting dilution template	R1-Barcode-P5_Template	GCG TCG TGT AGG GAA AGA GTG TNN NNN NNN NNN NGT GTA GAT CTC GGT GGT CGC CGT ATC ATT	Hand mixed, with native nucleotides
Template for adding 8-mer random primer	R1-Partial-8Nmer-rc	NNN NNN NNA CAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GT	Hand mixed, with native nucleotides, no 3'blocker needed

Figure 19.

Gel Bead Aliquot	Final vol (uL)	Concentration of mag beads mg/mL	Incubation time (hr)	Selective Deposition	Recovery (Y/N)
1	100	100	24	Yes	No
2	100	80	24	Yes	No
3	100	60	24	Yes	Yes
4	100	40	24	Yes	Yes
5	100	20	24	No	No
6	100	10	24	No	No

Figure 20.

Metric	Description	Value	Target Value
Q40 Error	Error rate across all q40 positions	0.002	<=0.02
Unmapped fraction	Fraction of reads with both ends unmapped	0.996	<=0.04
Median insert size	Median insert size	399	[200:400]
IQR insert size	Interquartile range of insert size	413	<=250
Zero coverage fraction	Fraction of genome with zero coverage	0.9242	<=0.01

Figure 21.

Metric	Description	Value	Target Value
Q40 Error	Error rate across all q40 positions	0.001	<=0.02
Unmapped fraction	Fraction of reads with both ends unmapped	0.03	<=0.04
Median insert size	Median insert size	310	[200:400]
IQR insert size	Interquartile range of insert size	209	<=250
Zero coverage fraction	Fraction of genome with zero coverage	0.0093	<=0.01

Figure 22.

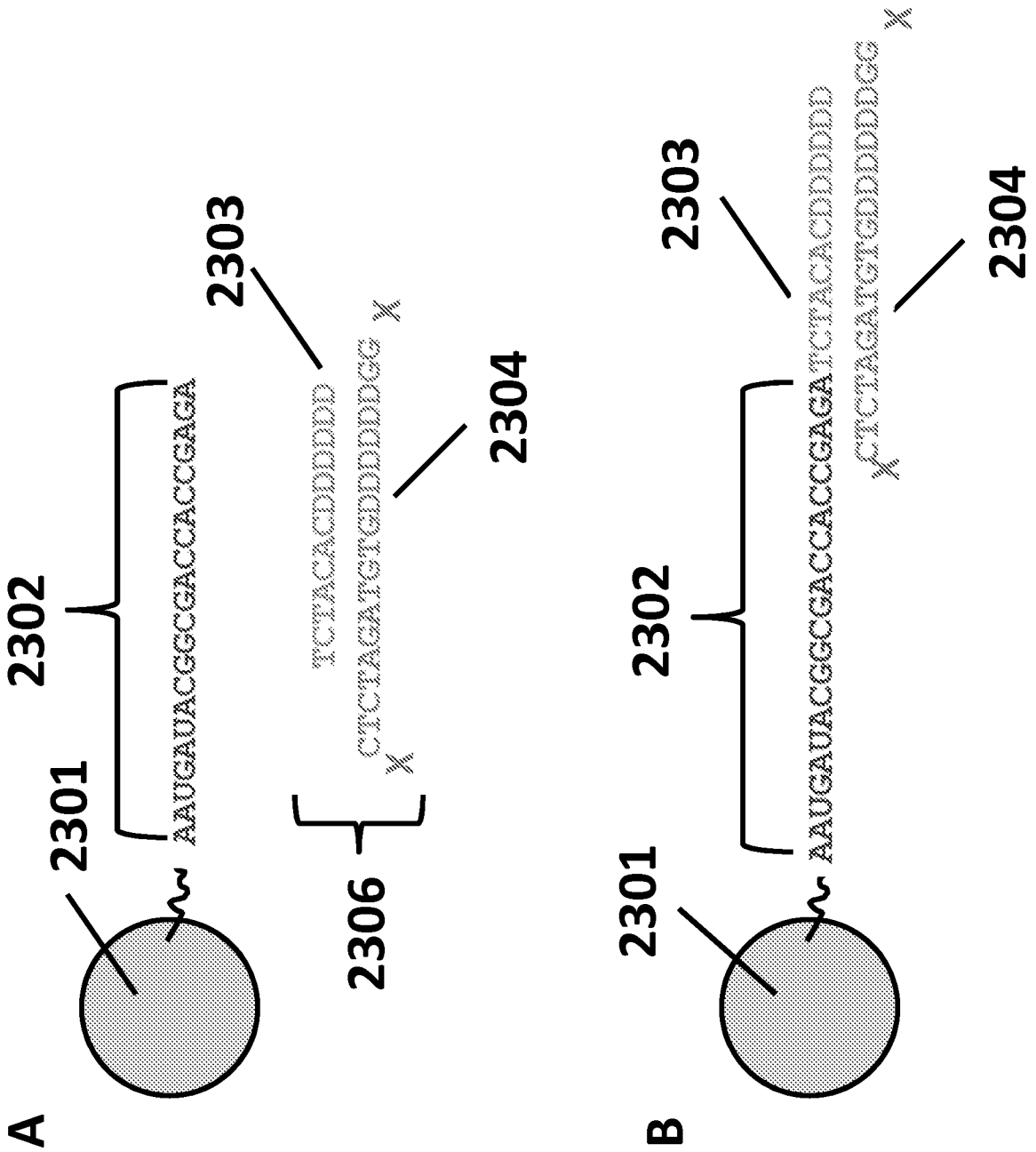

A
2301
2302
2303
2304
2306
AAUGAUACGGCGACCACCGAGA
TCTACACDDDDD
CTCTAGATGTGDDDDDGG
X
B
2301
2302
2303
2304
AAUGAUACGGCGACCACCGAGATCTACACDDDDD
CTCTAGATGTGDDDDDGG
X

Fig. 23

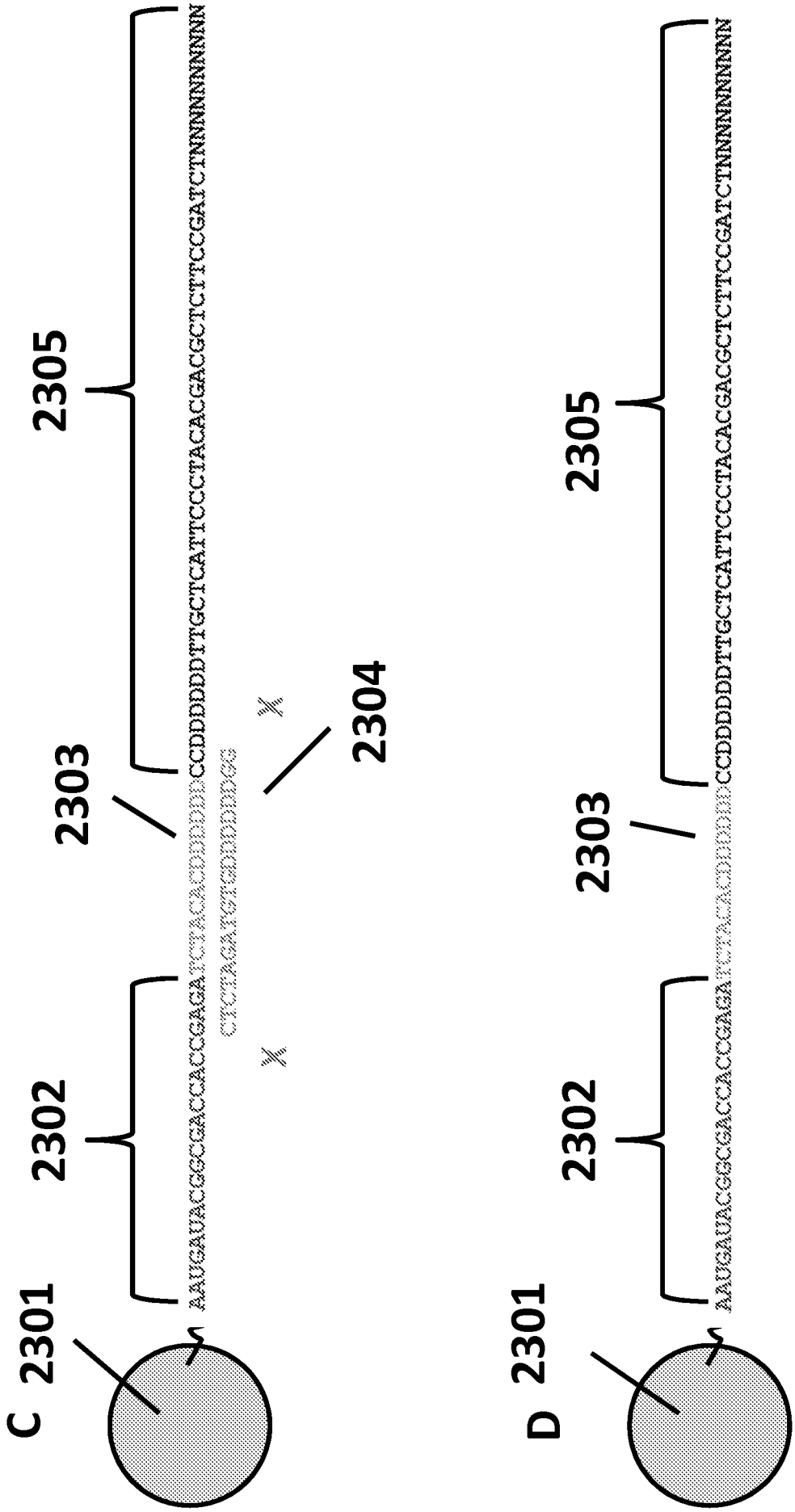

C
2301
2302
2303
2305
2304
AAUGAUACGGCGACCACCGAGATCTACACDDDDDDCCDDDDDTTGCTCATTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNN
CTCTAGATGTGDDDDDG
D
2301
2302
2303
2305
AAUGAUACGGCGACCACCGAGATCTACACDDDDDDCCDDDDDTTGCTCATTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNN

Fig. 23 (cont.)

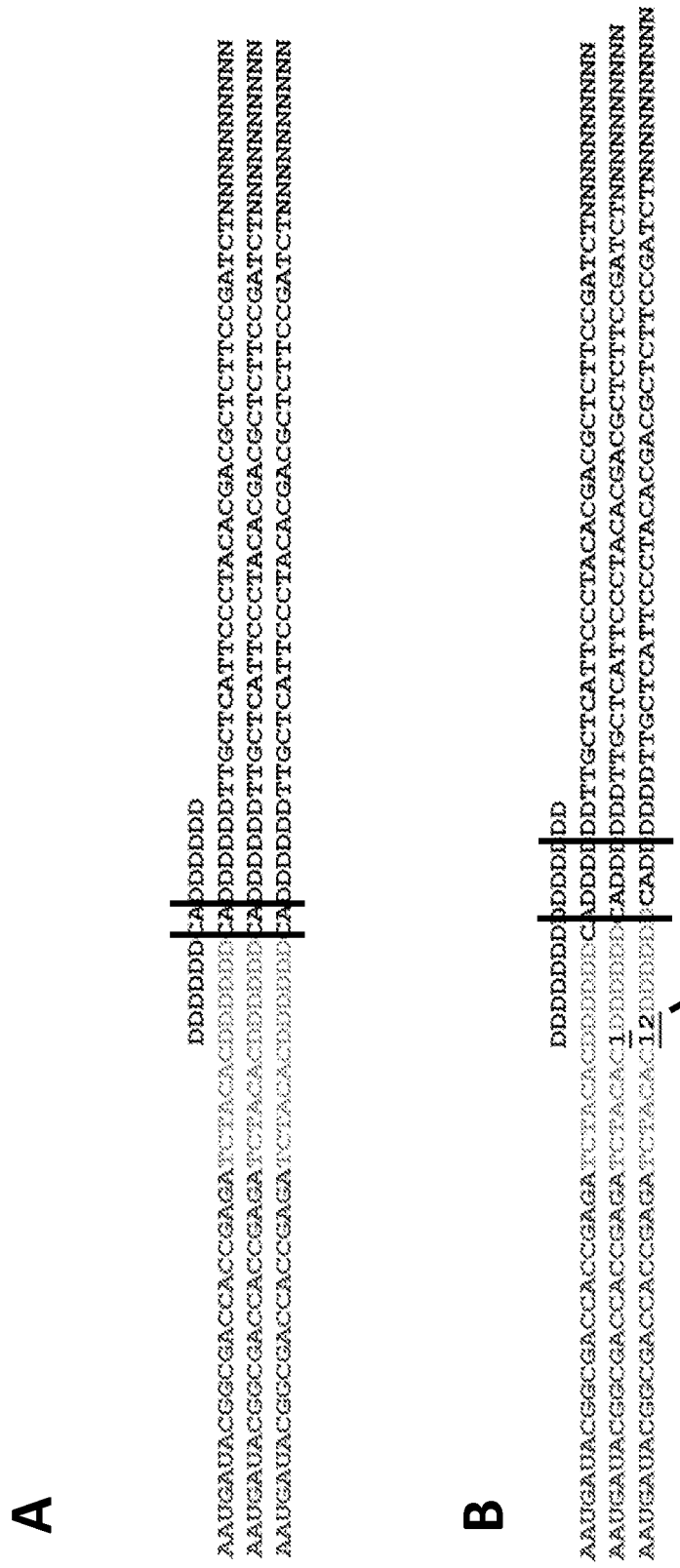


Fig. 24

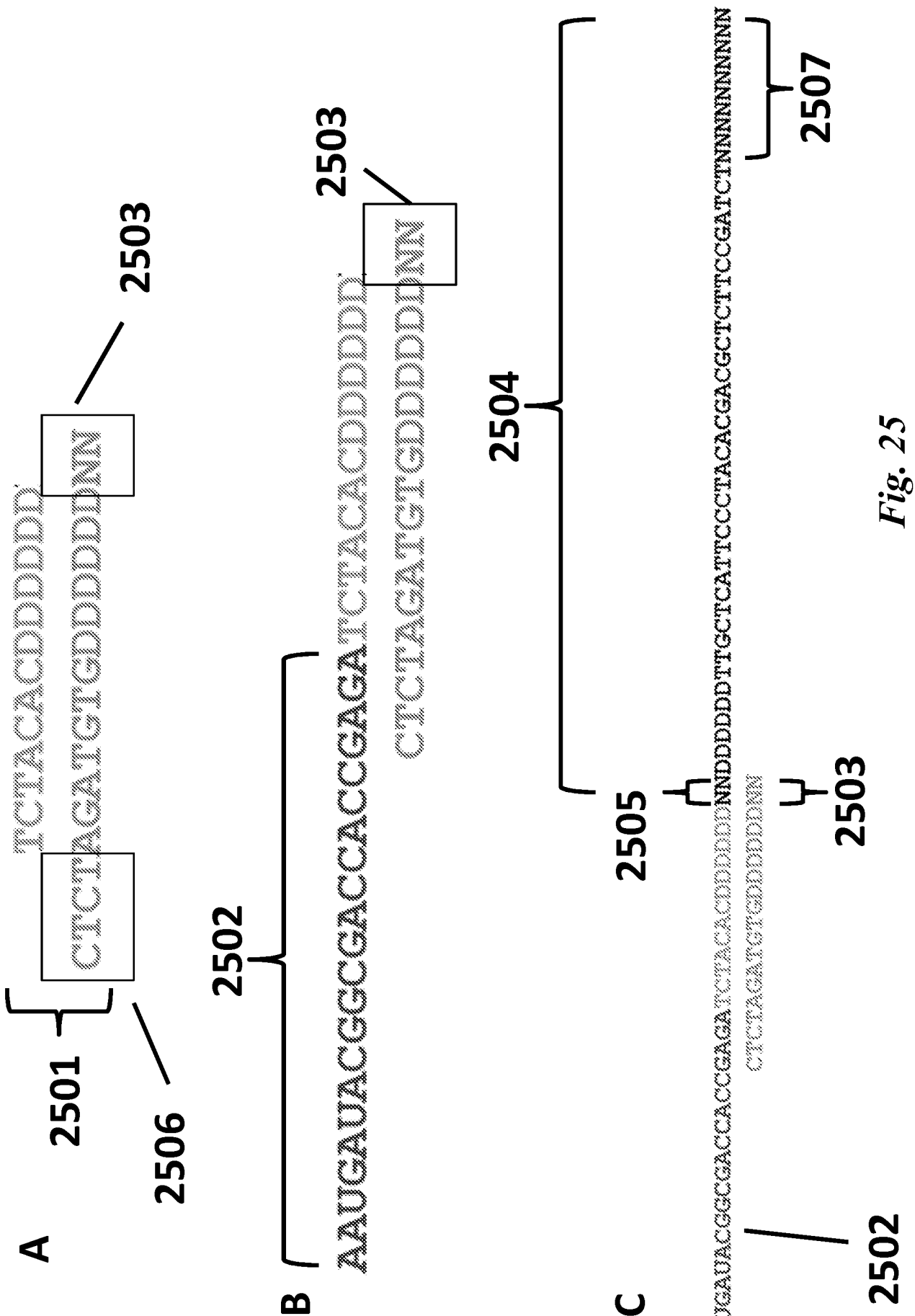

A
2501
2506
2503
TCTACACDDDDD
CTCTAGATGTGDDDDDDMN
B
2502
2503
AAUGAUACGGCGACCACCGAGATCTACACDDDDD
CTCTAGATGTGDDDDDDMN
C
2504
2505
2503
2507
2502
AAUGAUACGGCGACCACCGAGATCTACACDDDDDDNNDDDDDTTGCTCATTCCCTACACGACGCTCTTCCGATCTNNNNNNNNN
CTCTAGATGTGDDDDDDMN

Fig. 25

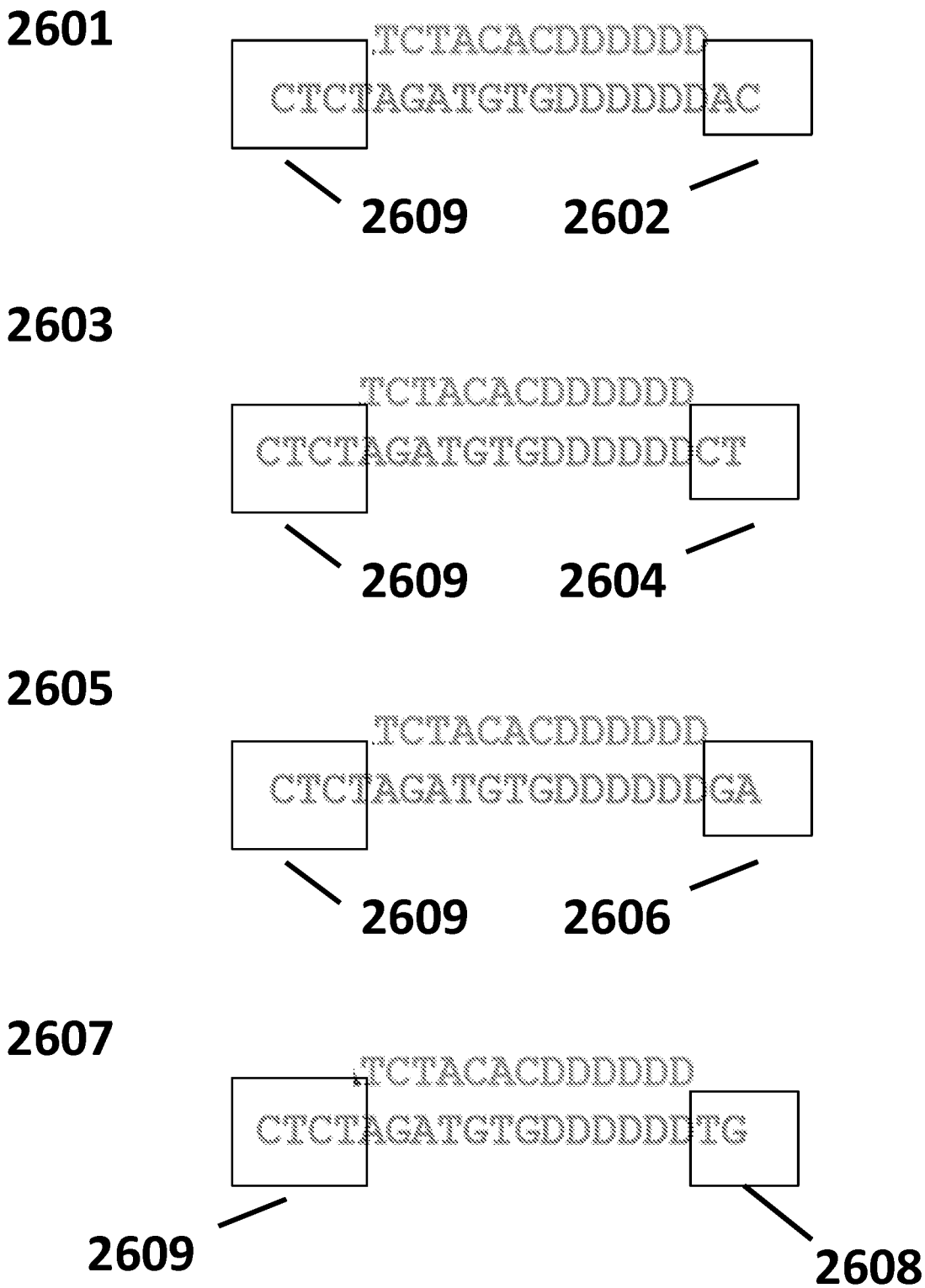

2601
TCTACACDDDDDD
CTCTAGATGTGDDDDDDDAC
2609
2602
2603
TCTACACDDDDDD
CTCTAGATGTGDDDDDDDCT
2609
2604
2605
TCTACACDDDDDD
CTCTAGATGTGDDDDDDDGA
2609
2606
2607
TCTACACDDDDDD
CTCTAGATGTGDDDDDDDTG
2609
2608

Fig. 26

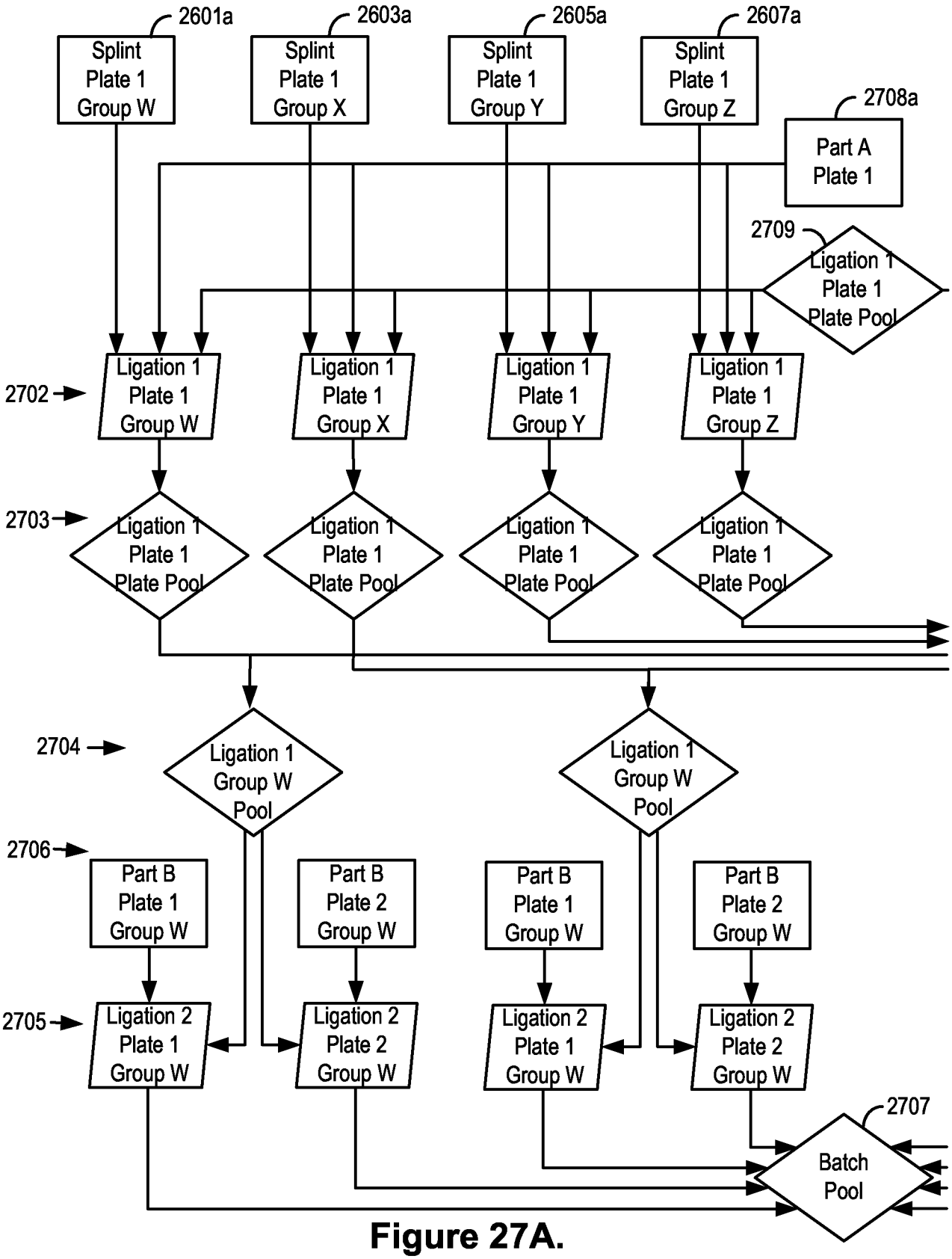


Figure 27A.

Figure 27A	Figure 27B

Figure 27.

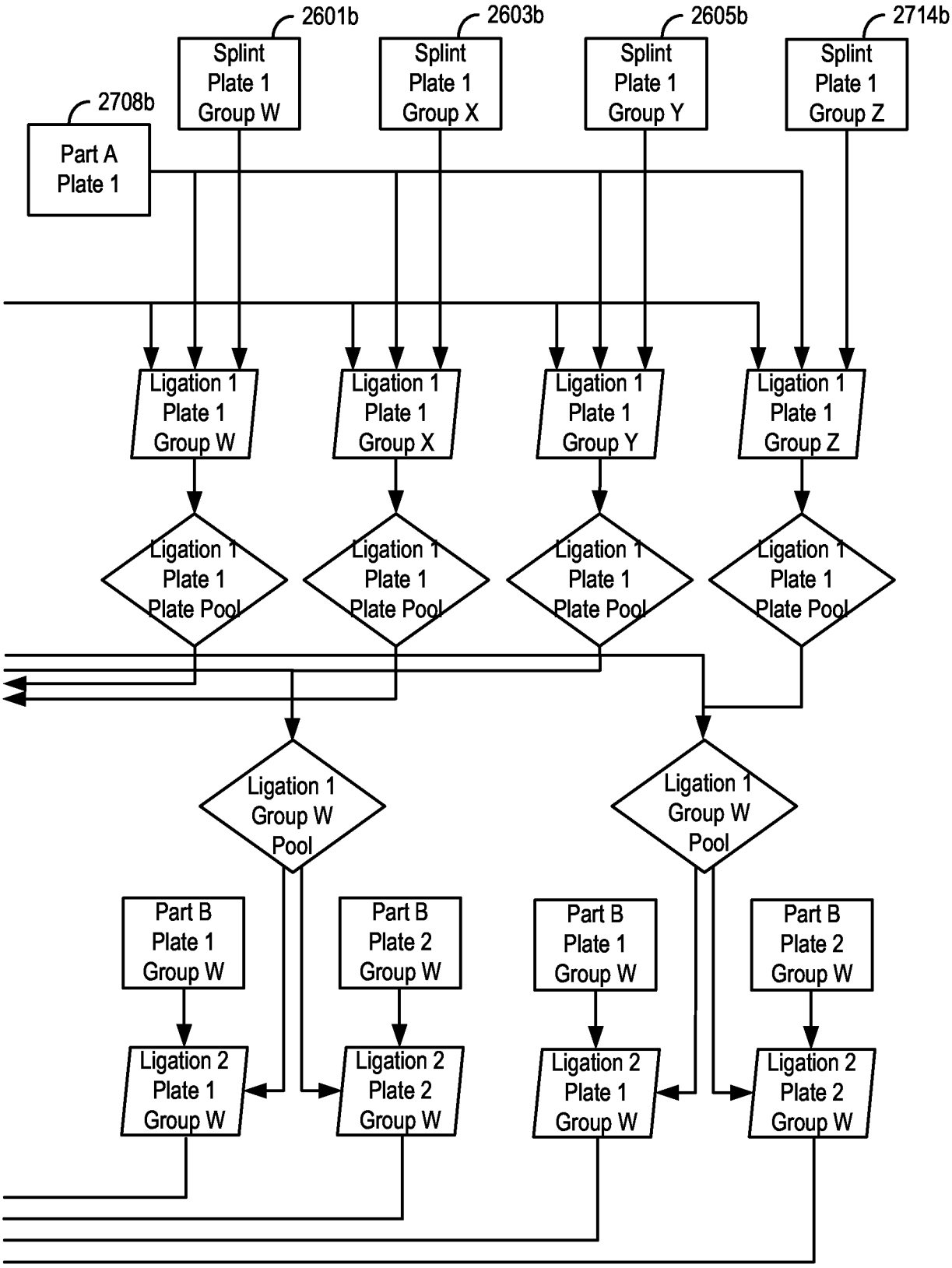

2601b
2603b
2605b
2714b
Splint Plate 1 Group W
Splint Plate 1 Group X
Splint Plate 1 Group Y
Splint Plate 1 Group Z
2708b
Part A Plate 1
Ligation 1 Plate 1 Group W
Ligation 1 Plate 1 Group X
Ligation 1 Plate 1 Group Y
Ligation 1 Plate 1 Group Z
Ligation 1 Plate 1 Plate Pool
Ligation 1 Plate 1 Plate Pool
Ligation 1 Plate 1 Plate Pool
Ligation 1 Plate 1 Plate Pool
Ligation 1 Group W Pool
Ligation 1 Group W Pool
Part B Plate 1 Group W
Part B Plate 2 Group W
Part B Plate 1 Group W
Part B Plate 2 Group W
Ligation 2 Plate 1 Group W
Ligation 2 Plate 2 Group W
Ligation 2 Plate 1 Group W
Ligation 2 Plate 2 Group W

Figure 27B.

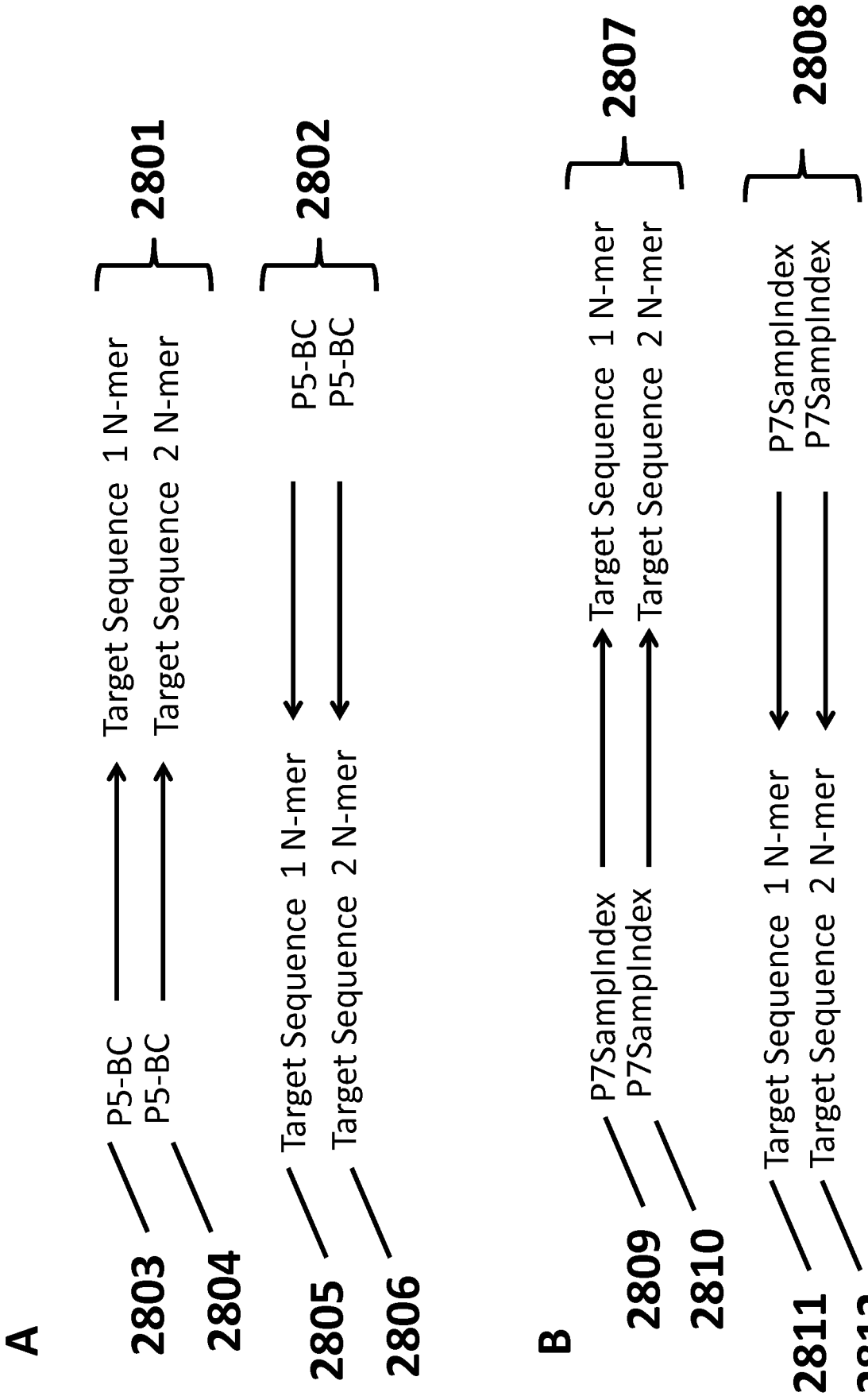
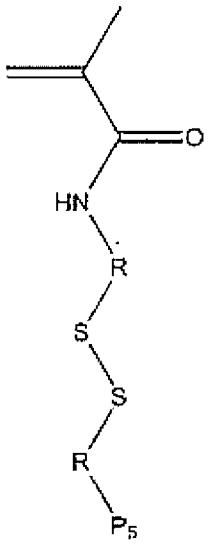

A
2803
2804
2805
2806
P5-BC
P5-BC
Target Sequence 1 N-mer
Target Sequence 2 N-mer
2801
2802
B
2809
2810
2811
2812
P7SampIndex
P7SampIndex
Target Sequence 1 N-mer
Target Sequence 2 N-mer
2807
2808

Fig. 28

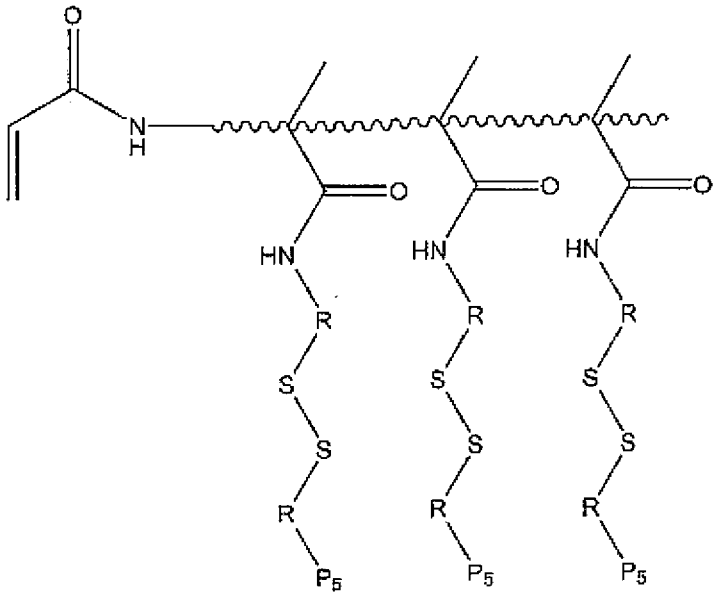
A

P_5 monomer



B

Multi-P_5 monomer



C

Multi-P_5 crosslinker

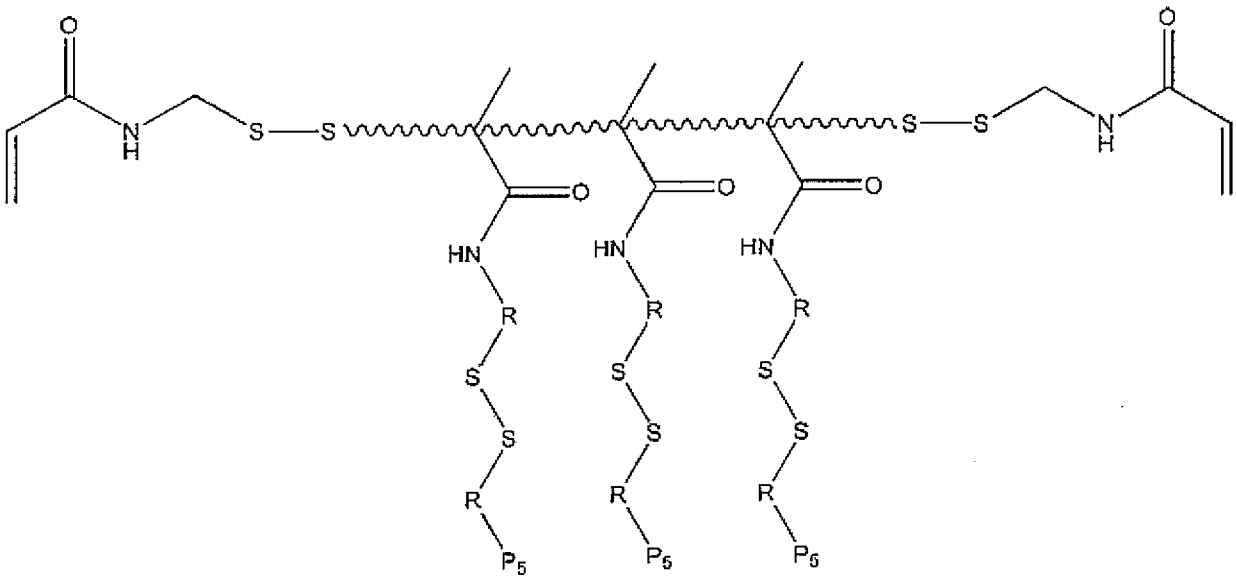


Fig. 29

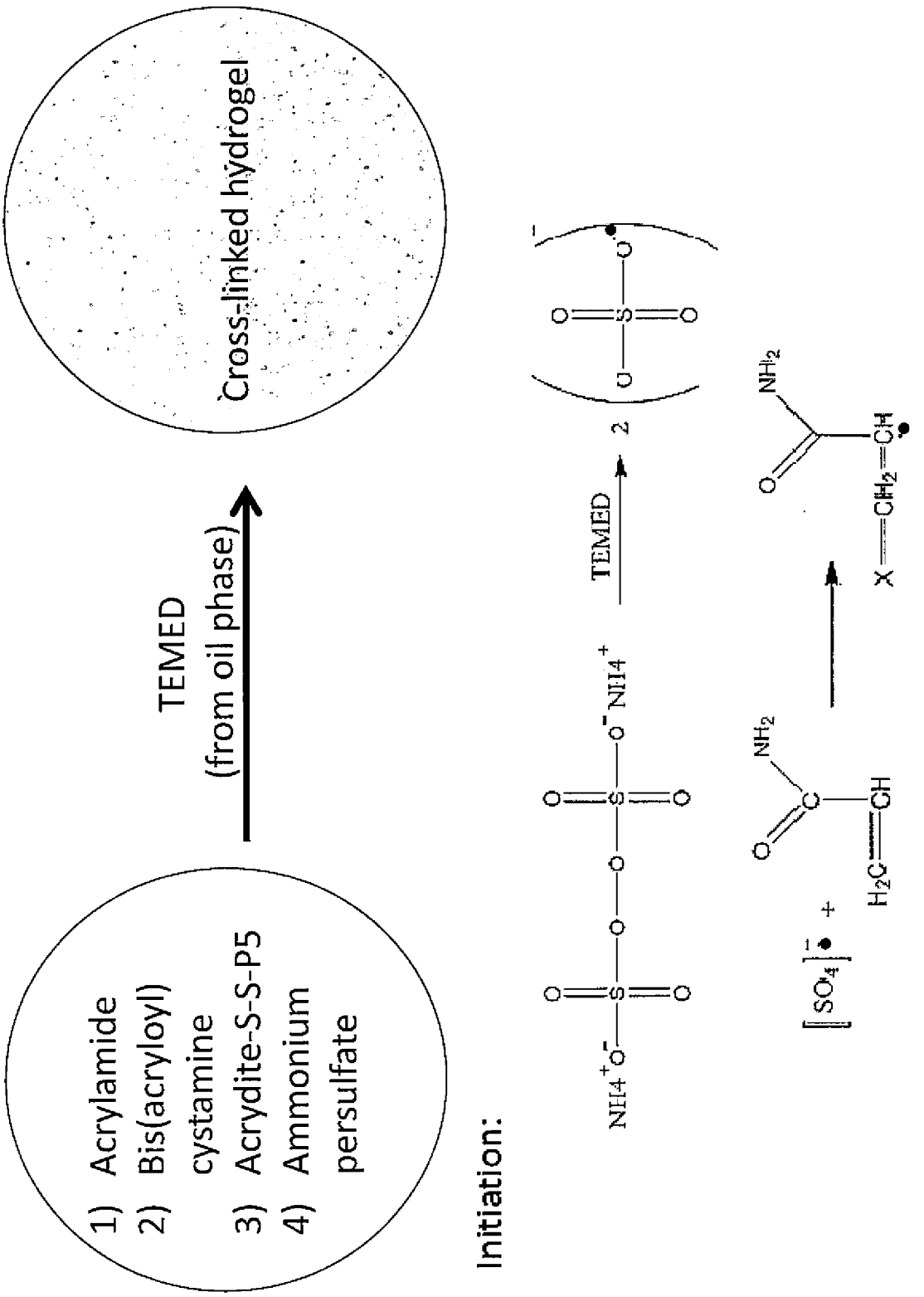

1) Acrylamide
2) Bis(acryloyl) cystamine
3) Acrydite-S-S-P5
4) Ammonium persulfate
TEMED
(from oil phase)
Cross-linked hydrogel
Initiation:
TEMED
2

Fig. 30A

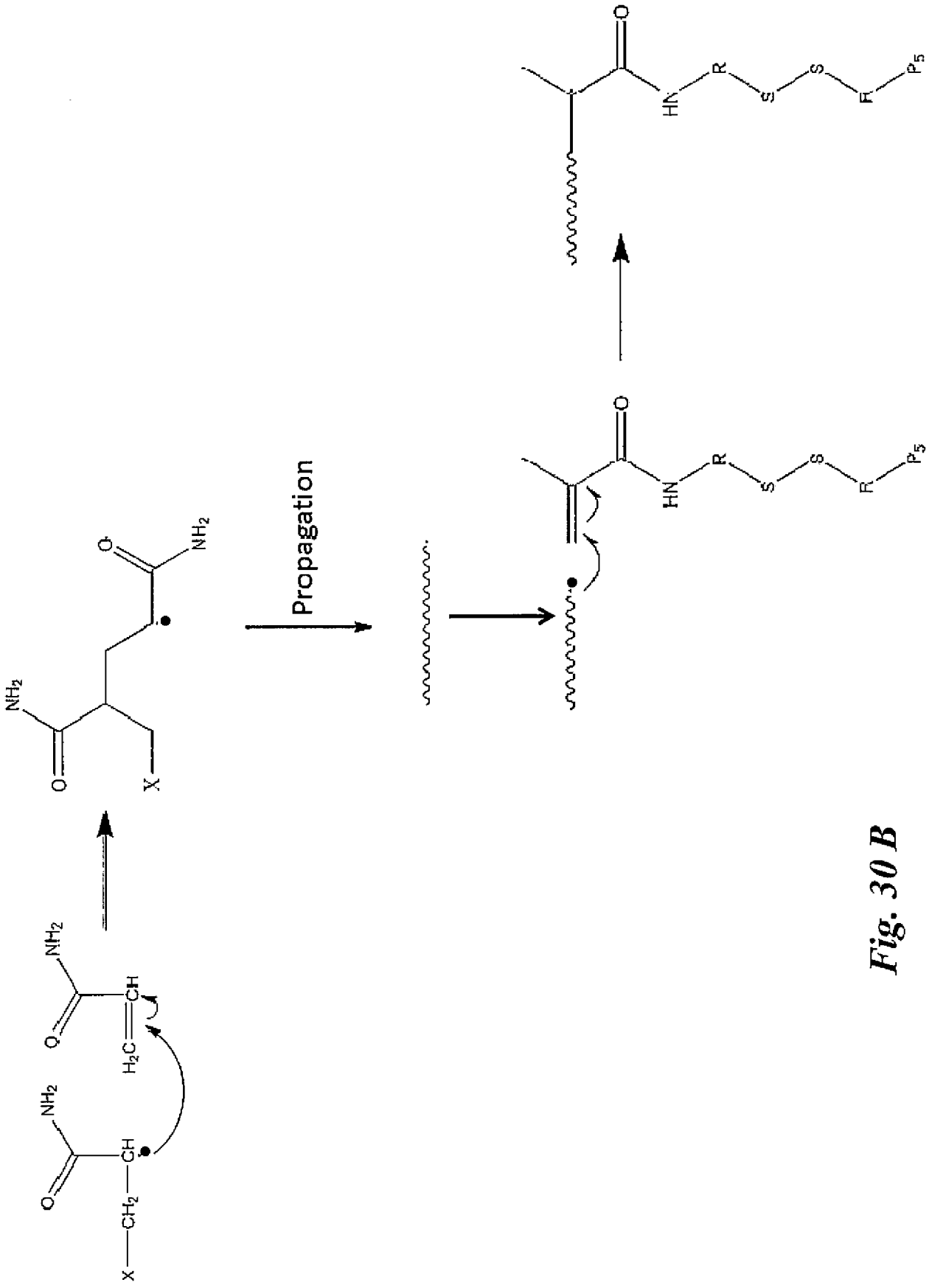


Fig. 30 B

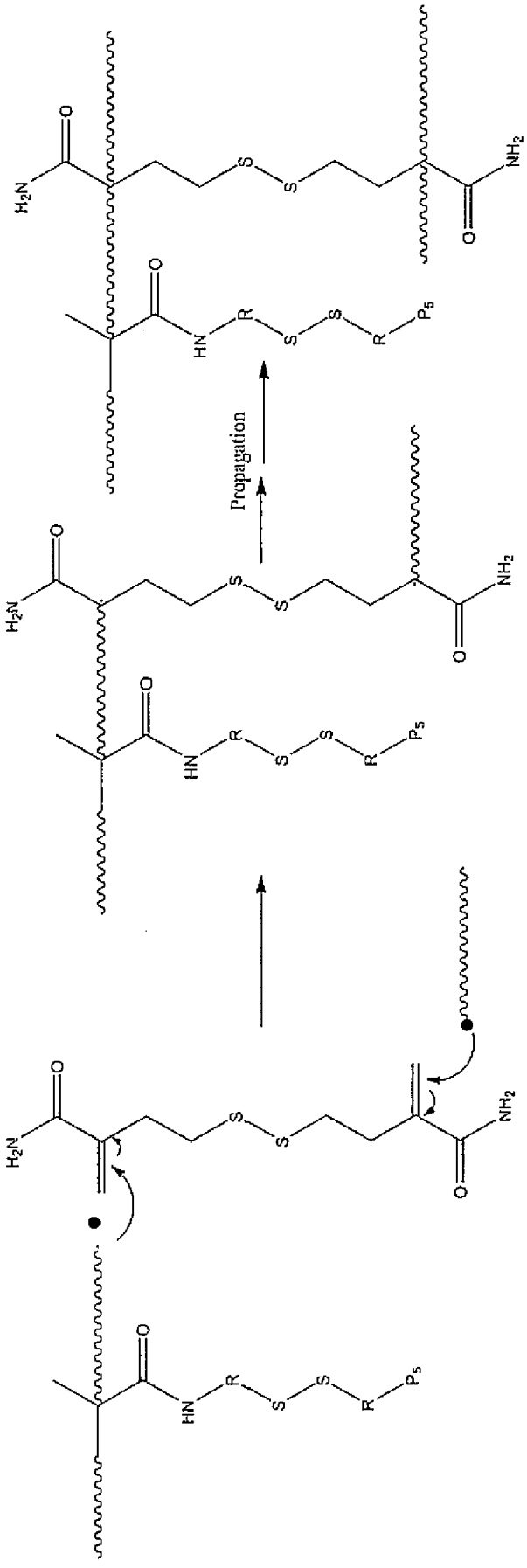


Fig. 30C

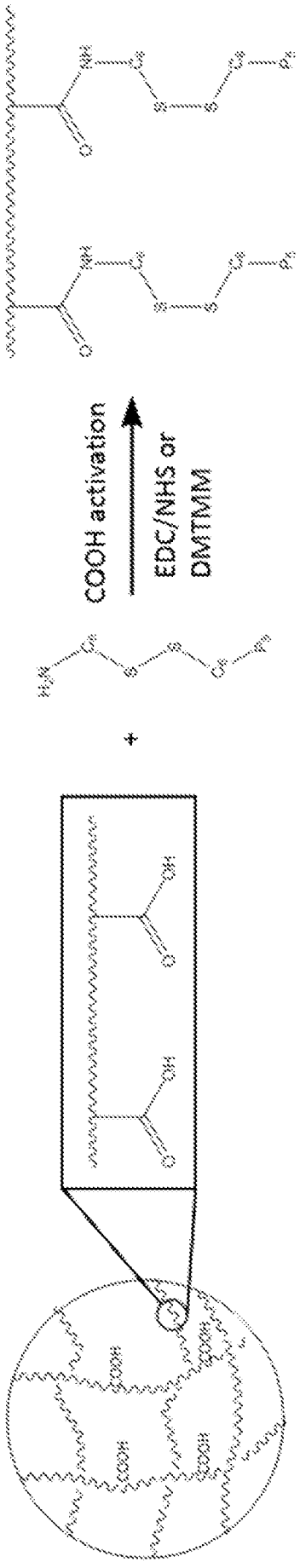
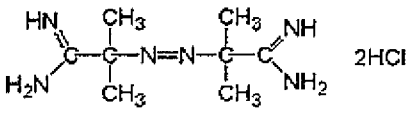
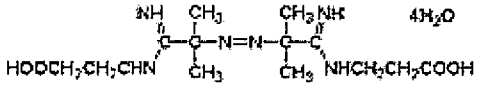


Fig. 31

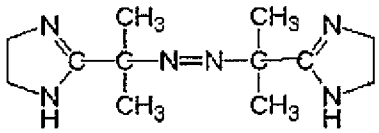
I



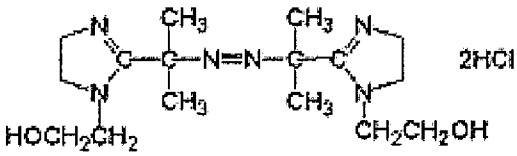
II



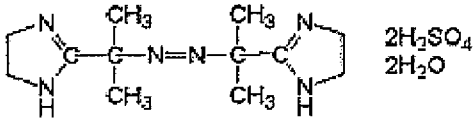
III



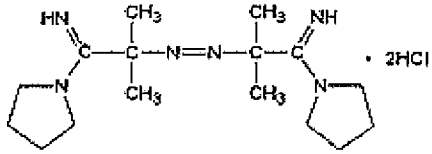
IV



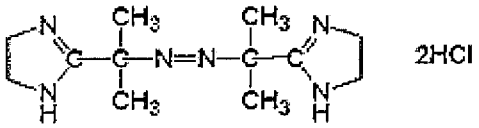
V



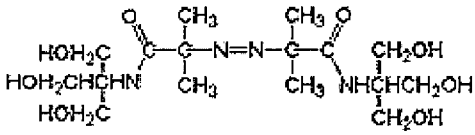
VI



VII



VIII



IX

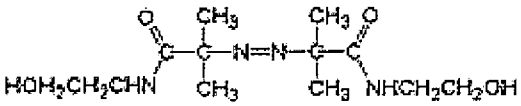


Fig. 32

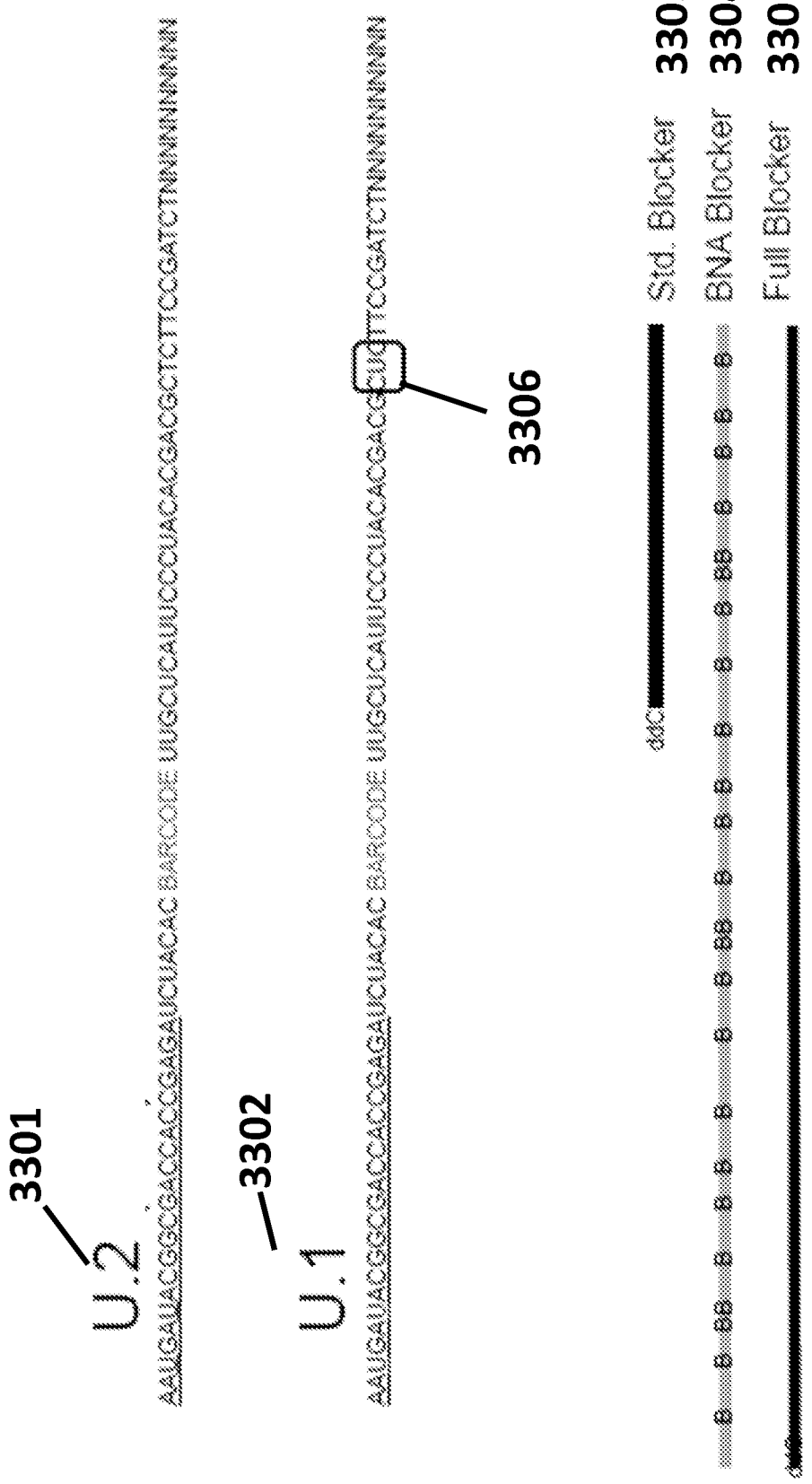

3301
U.2
AAUGAUACGGCGACCACCGAGAUCUACAC BARCODE UUGCUCAUUCCCUACACGACGCTCTTCCGATCTNNNNNNNNN
3302
U.1
AAUGAUACGGCGACCACCGAGAUCUACAC BARCODE UUGCUCAUUCCCUACACGACGCUUTTCCGATCTNNNNNNNNN
3306
dMC
Std. Blocker 3303
BNA Blocker 3304
Full Blocker 3305

Fig. 33A

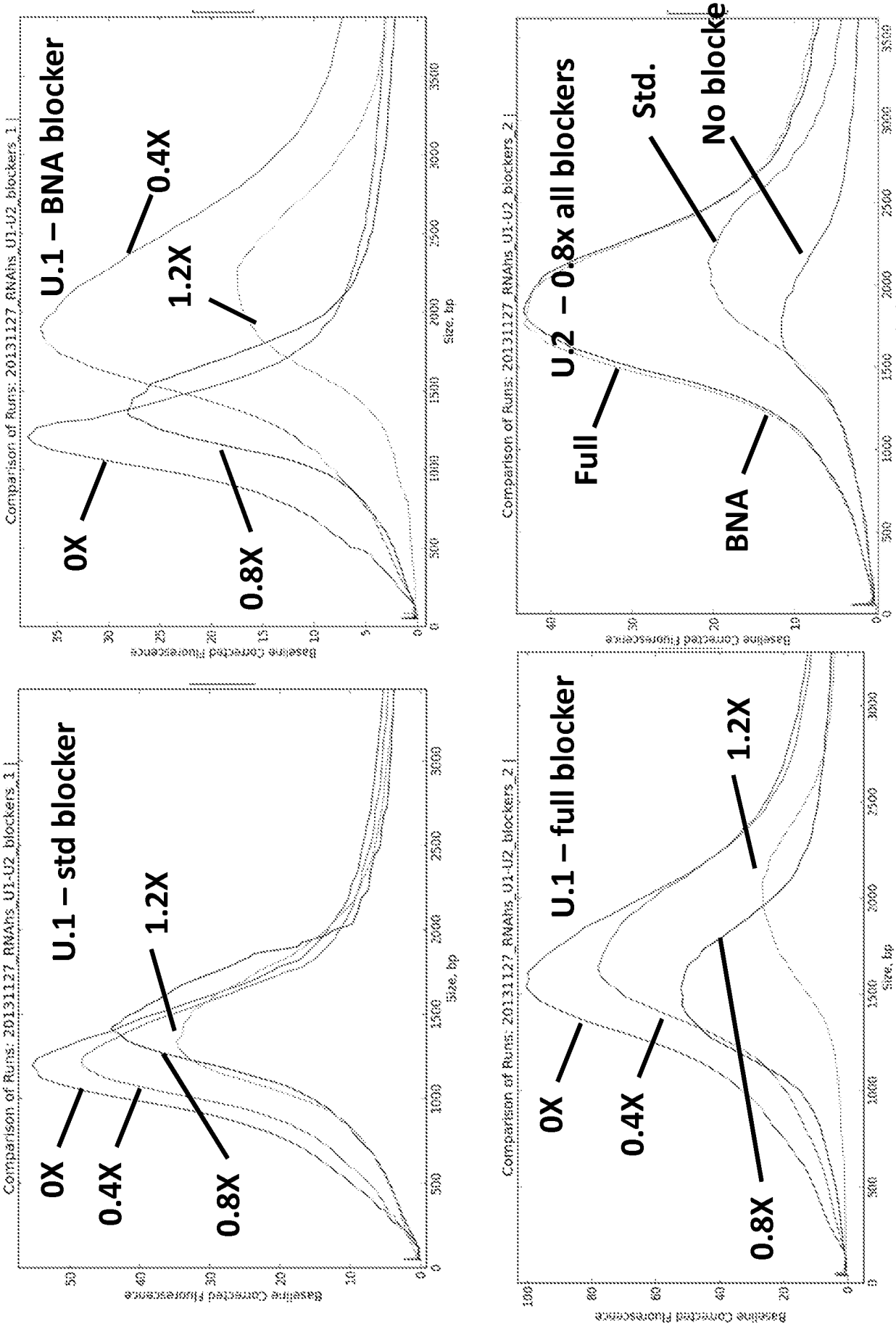

Comparison of Runs: 20131127_RNAIns_U1-U2_blockers_1
U.1 – std blocker
0X
0.4X
1.2X
0.8X
Baseline Corrected Fluorescence
Size, bp
Comparison of Runs: 20131127_RNAIns_U1-U2_blockers_1
U.1 – BNA blocker
0X
0.4X
1.2X
0.8X
Baseline Corrected Fluorescence
Size, bp
Comparison of Runs: 20131127_RNAIns_U1-U2_blockers_2
U.1 – full blocker
0X
0.4X
1.2X
0.8X
Baseline Corrected Fluorescence
Size, bp
Comparison of Runs: 20131127_RNAIns_U1-U2_blockers_2
U.2 – 0.8x all blockers
Full
Std.
BNA
No blocker
Baseline Corrected Fluorescence
Size, bp

Fig. 33B

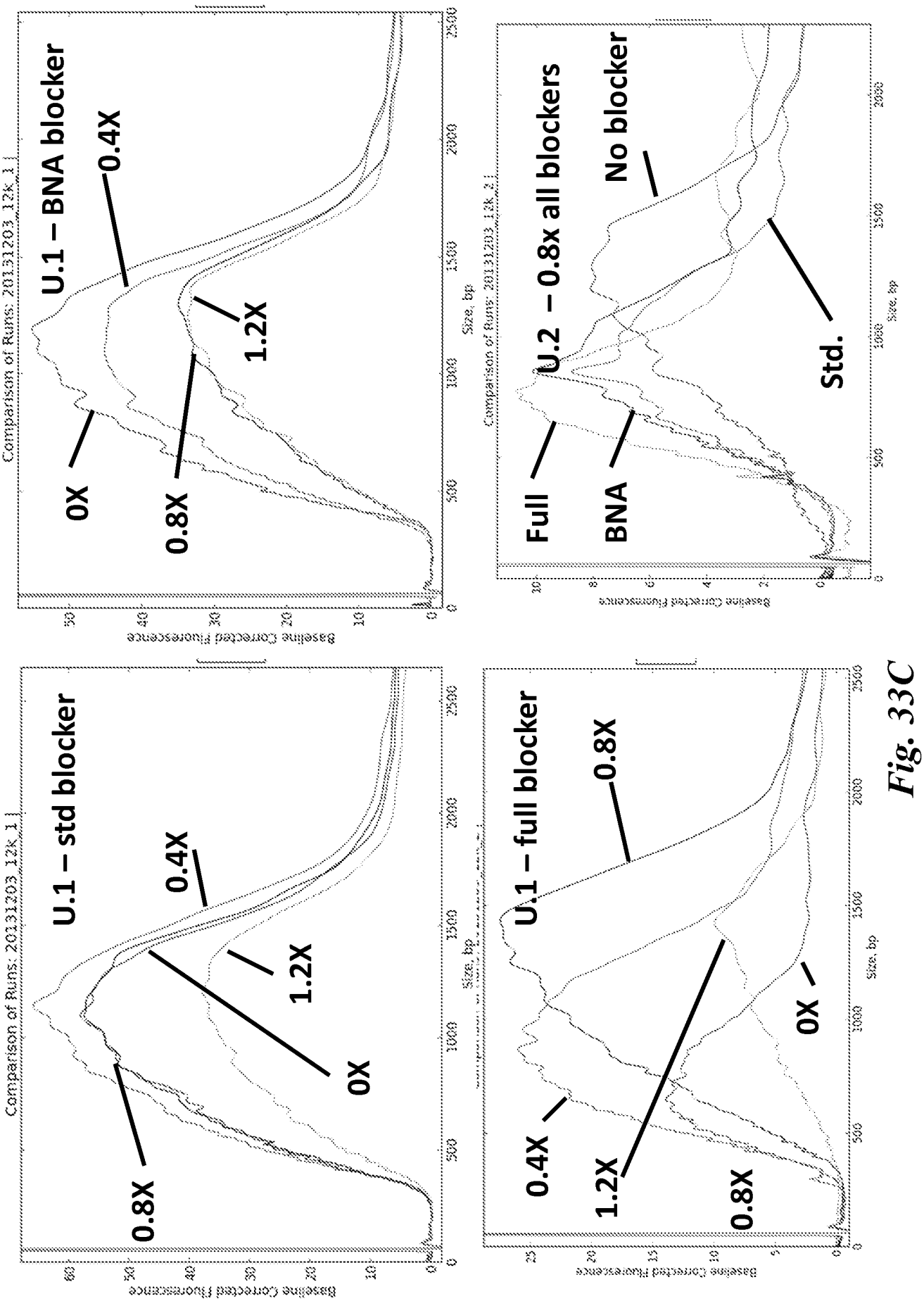
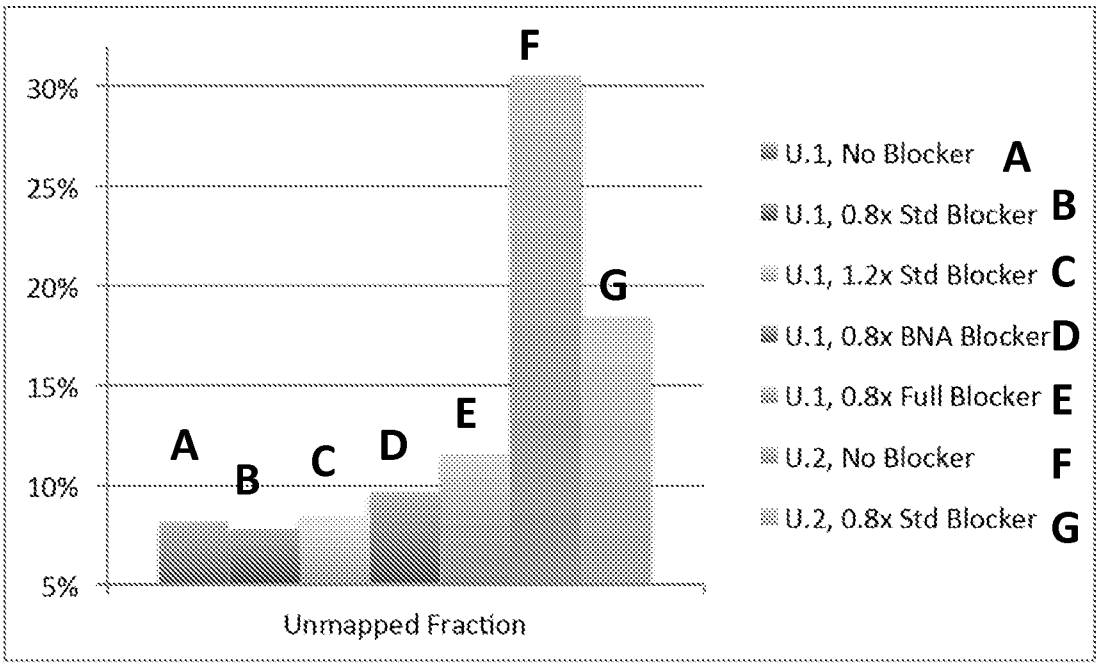

U.1 – std blocker
0.8X
0.4X
1.2X
0X
U.1 – BNA blocker
0X
0.4X
0.8X
1.2X
U.1 – full blocker
0.4X
1.2X
0.8X
0X
U.2 – 0.8x all blockers
Full
BNA
No blocker
Std.
Comparison of Runs: 20131203_12k_1
Baseline Corrected Fluorescence
Size, bp

Fig. 33C

I


30%
25%
20%
15%
10%
5%
A
B
C
D
E
F
G
Unmapped Fraction
U.1, No Blocker A
U.1, 0.8x Std Blocker B
U.1, 1.2x Std Blocker C
U.1, 0.8x BNA Blocker D
U.1, 0.8x Full Blocker E
U.2, No Blocker F
U.2, 0.8x Std Blocker G

II

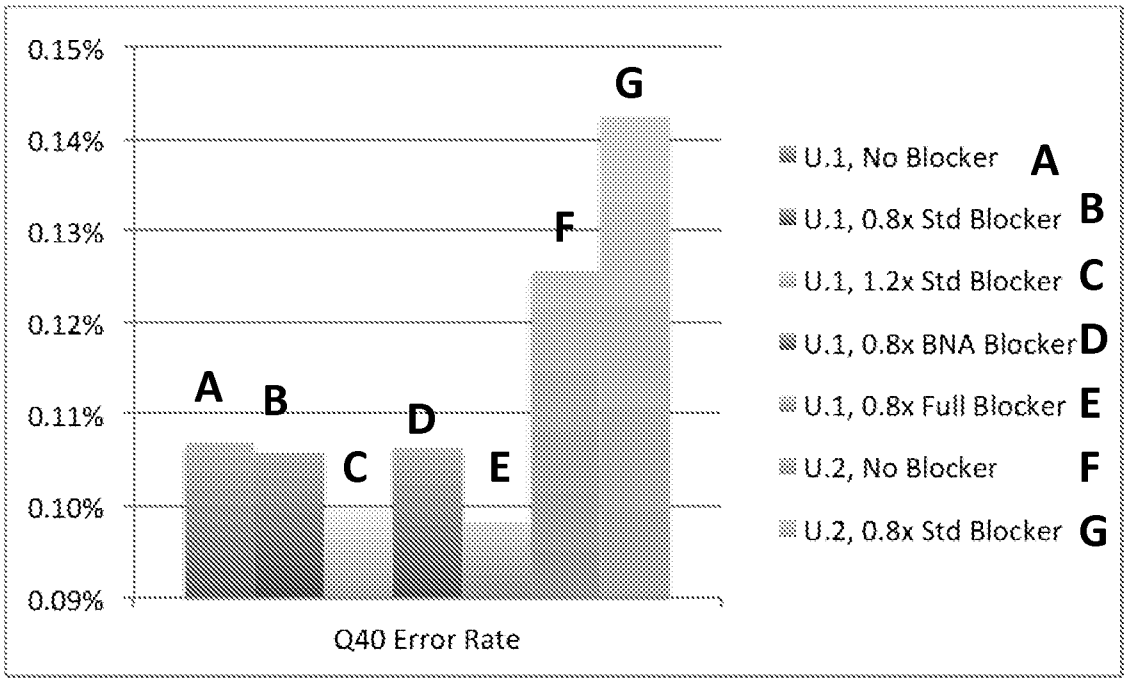
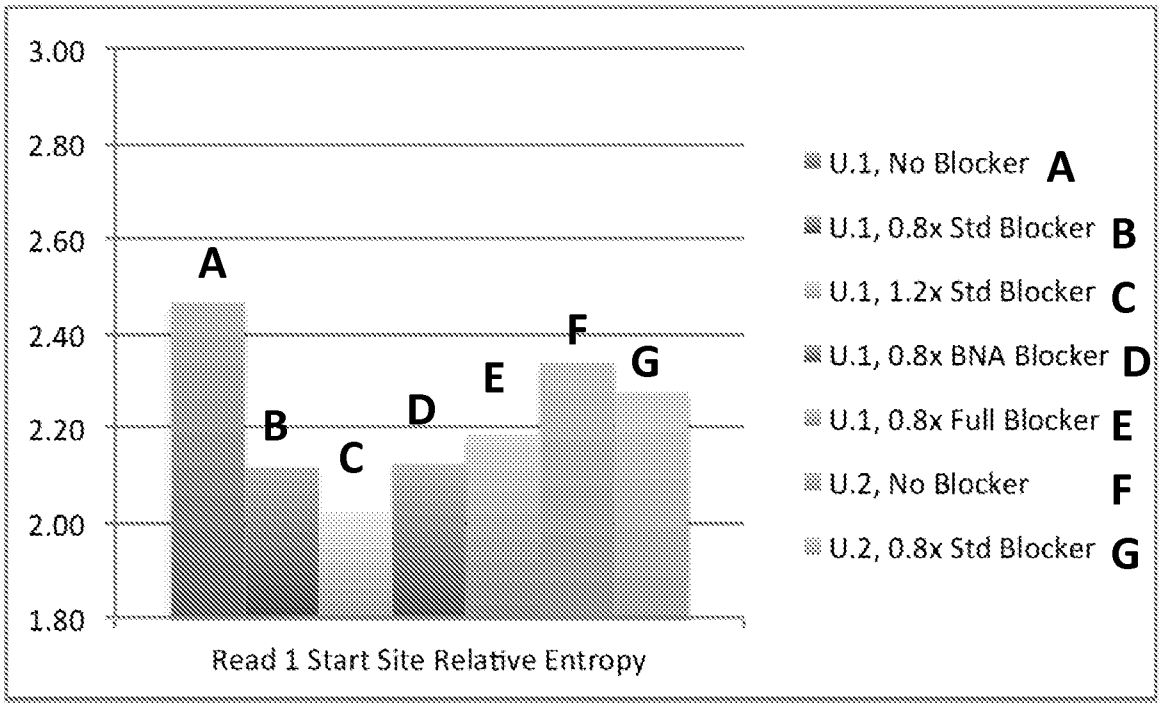

0.15%
0.14%
0.13%
0.12%
0.11%
0.10%
0.09%
A
B
C
D
E
F
G
Q40 Error Rate
U.1, No Blocker A
U.1, 0.8x Std Blocker B
U.1, 1.2x Std Blocker C
U.1, 0.8x BNA Blocker D
U.1, 0.8x Full Blocker E
U.2, No Blocker F
U.2, 0.8x Std Blocker G

Fig. 33D

III



IV

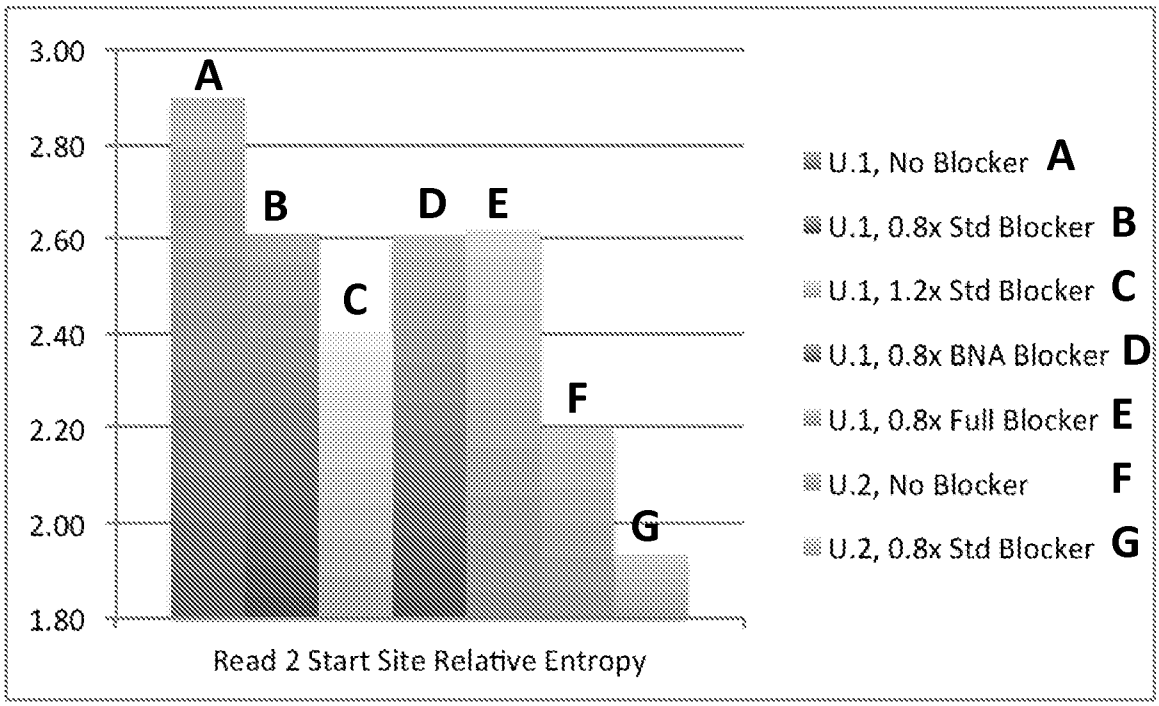


Fig. 33E

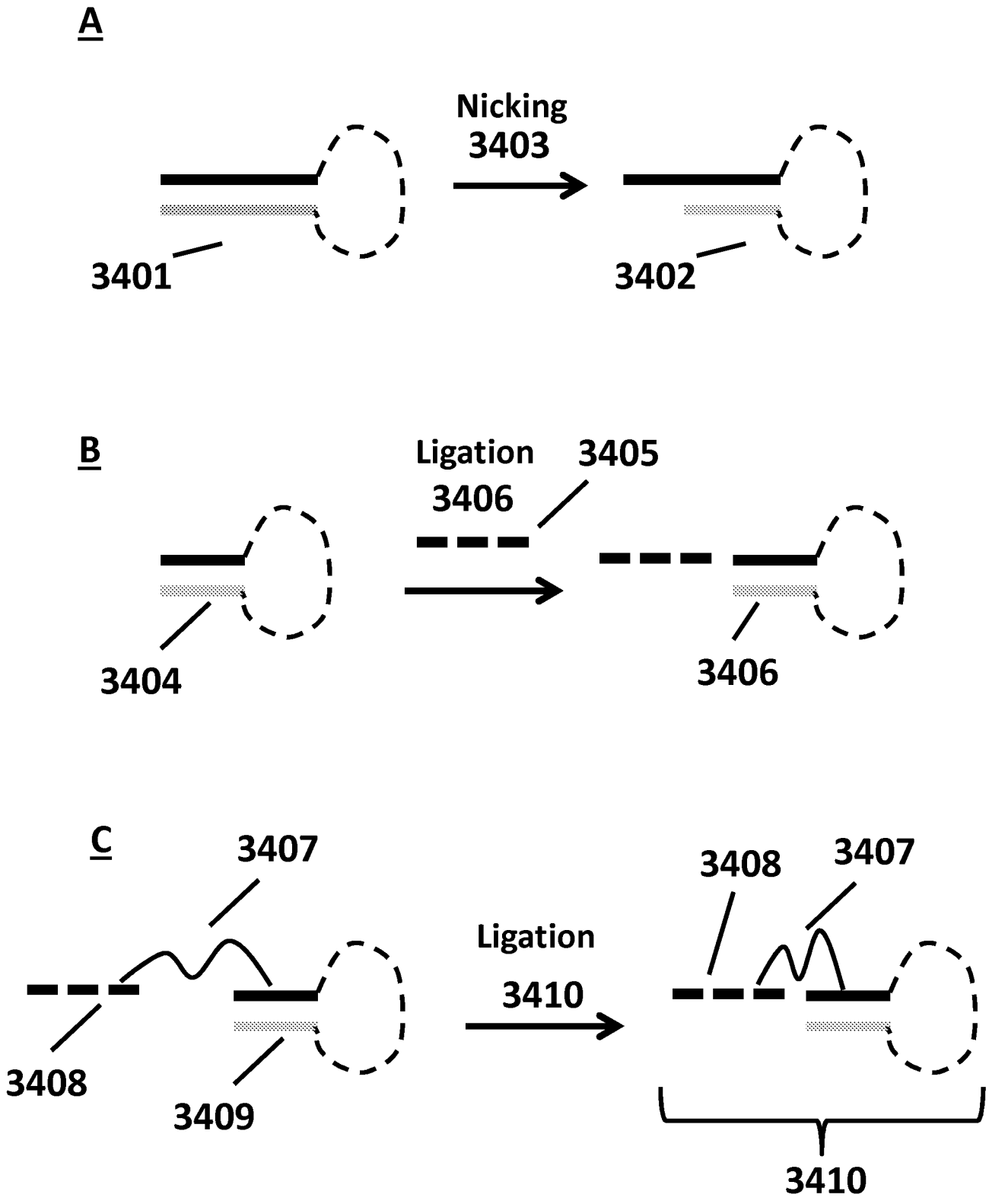

A
Nicking
3403
3401
3402
B
Ligation
3405
3406
3404
3406
C
3407
3408
3409
Ligation
3410
3408
3407
3410

Fig. 34

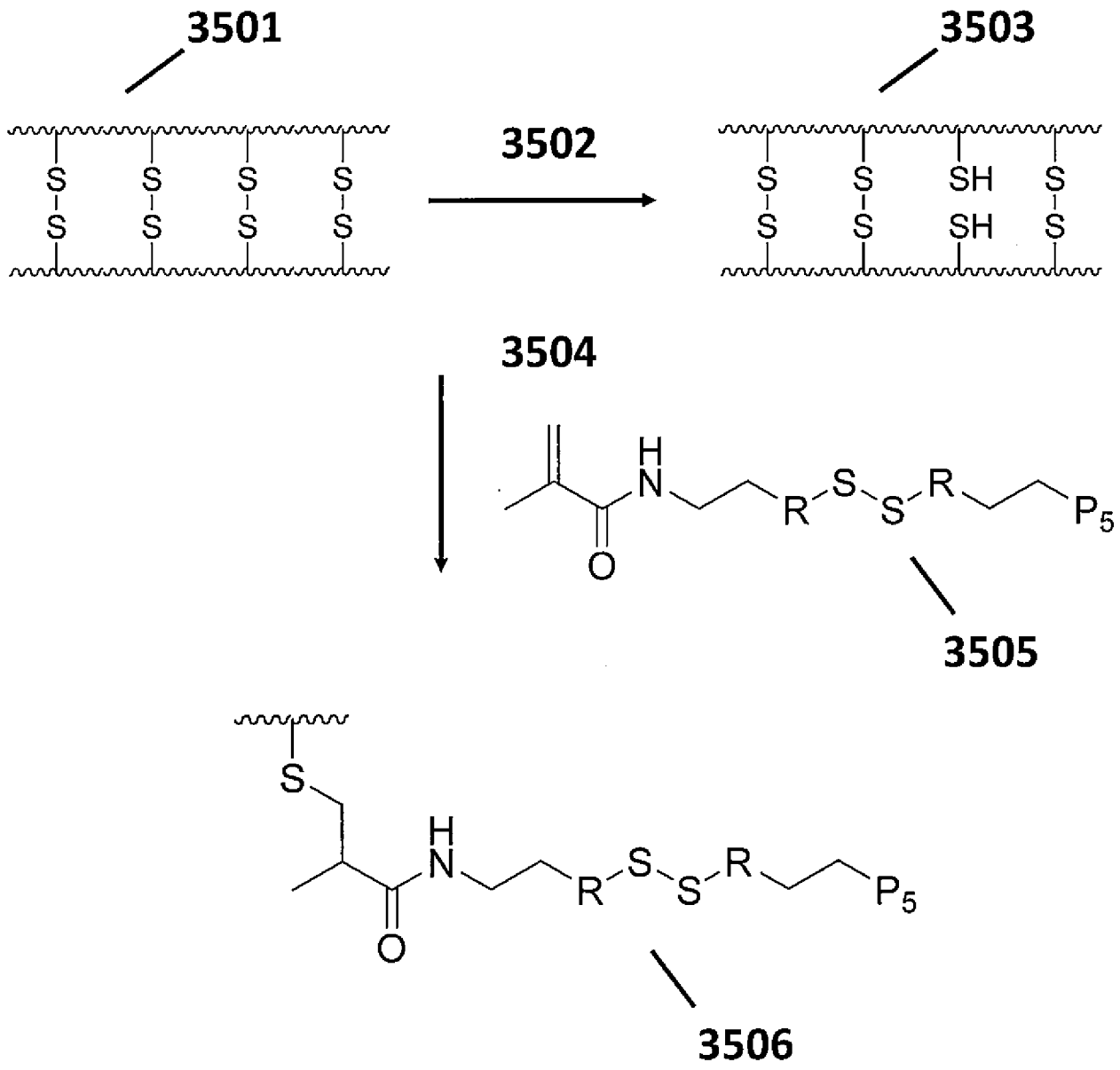


Fig. 35A

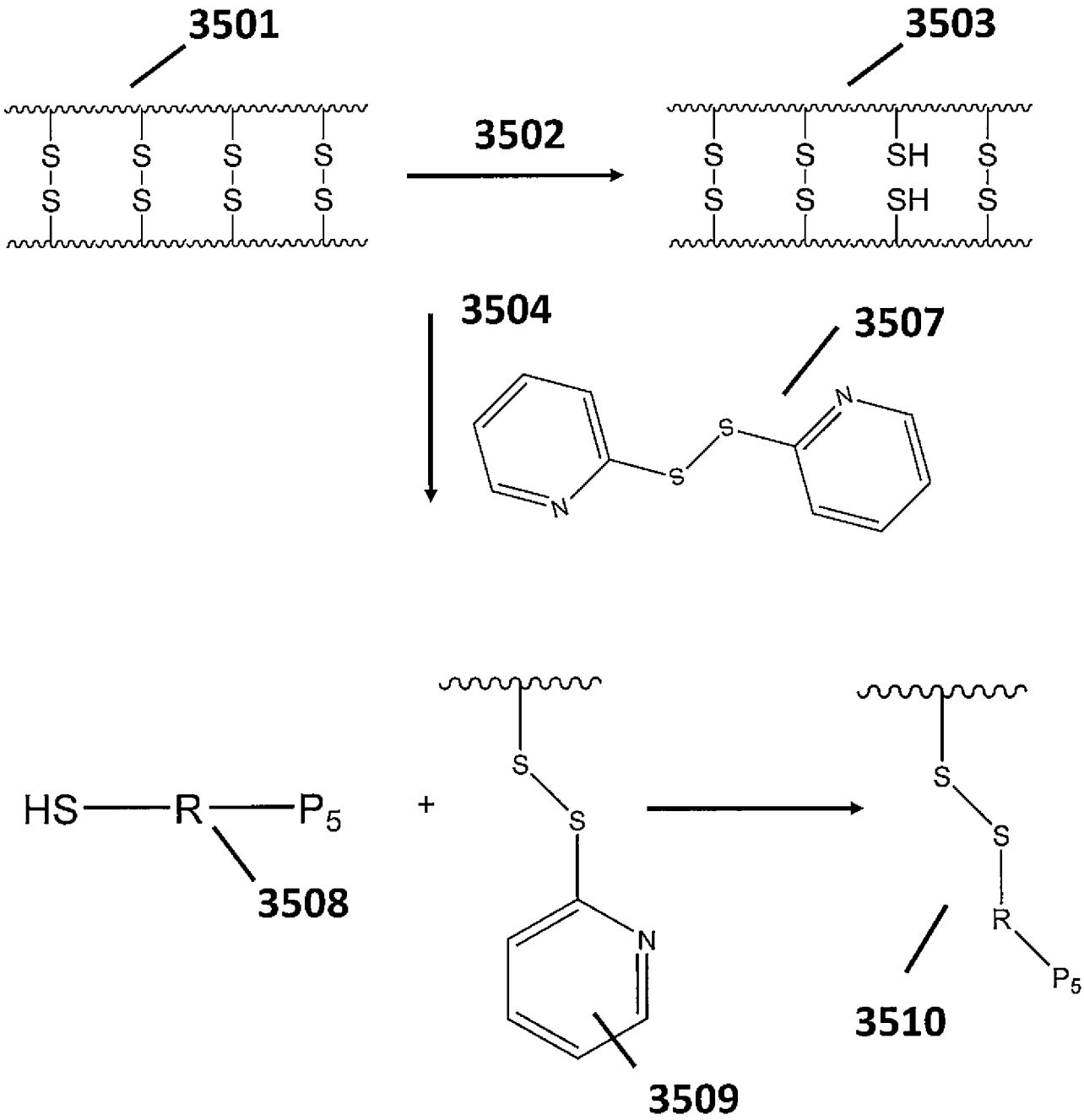
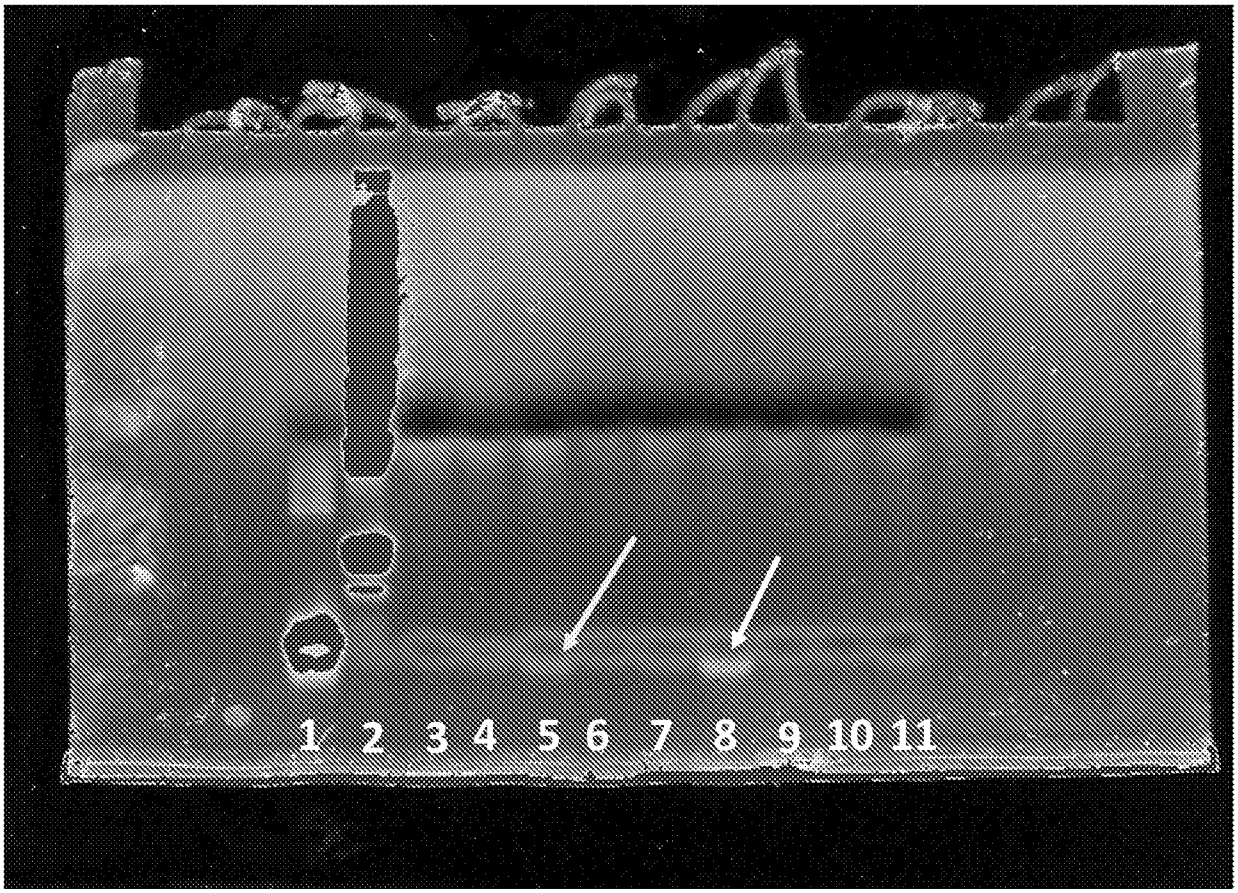


Fig. 35 B



1: 50 pmoles P5-acrydite
2: 25 bp ladder
3-5: TCEP:GB = 2.5 billion
6-8: TCEP:GB = 10.0 billion
9-11: TCEP:GB = 0
3,6,9: P5:GB = 50 million
4,7,10: P5:GB = 500 million
5,8,11: P5:GB = 5 billion

Fig. 36

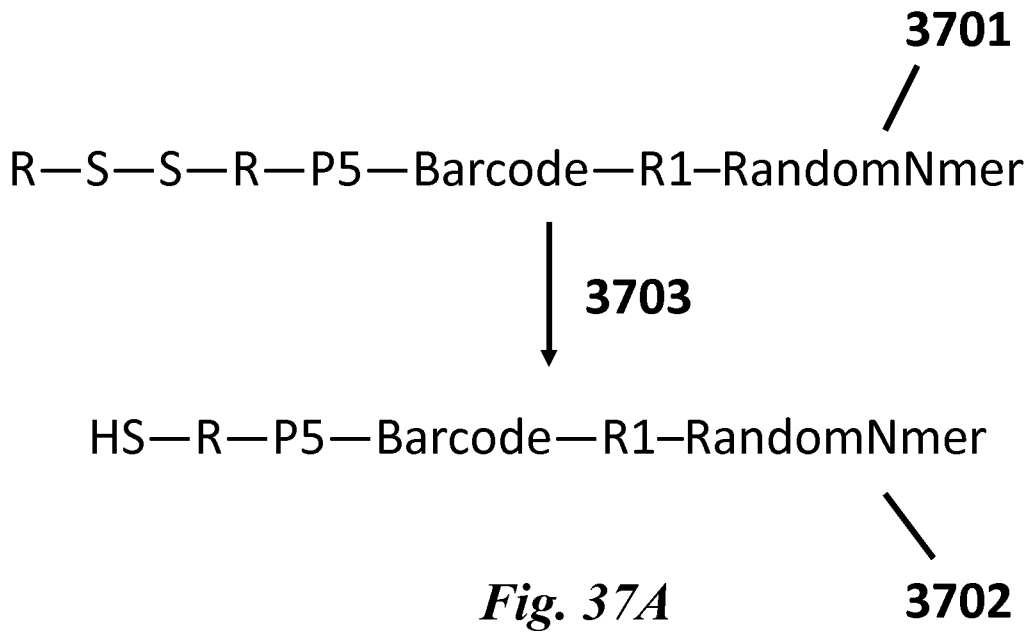
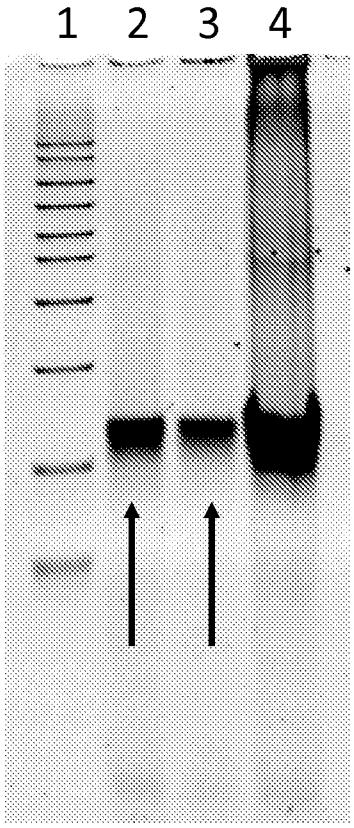
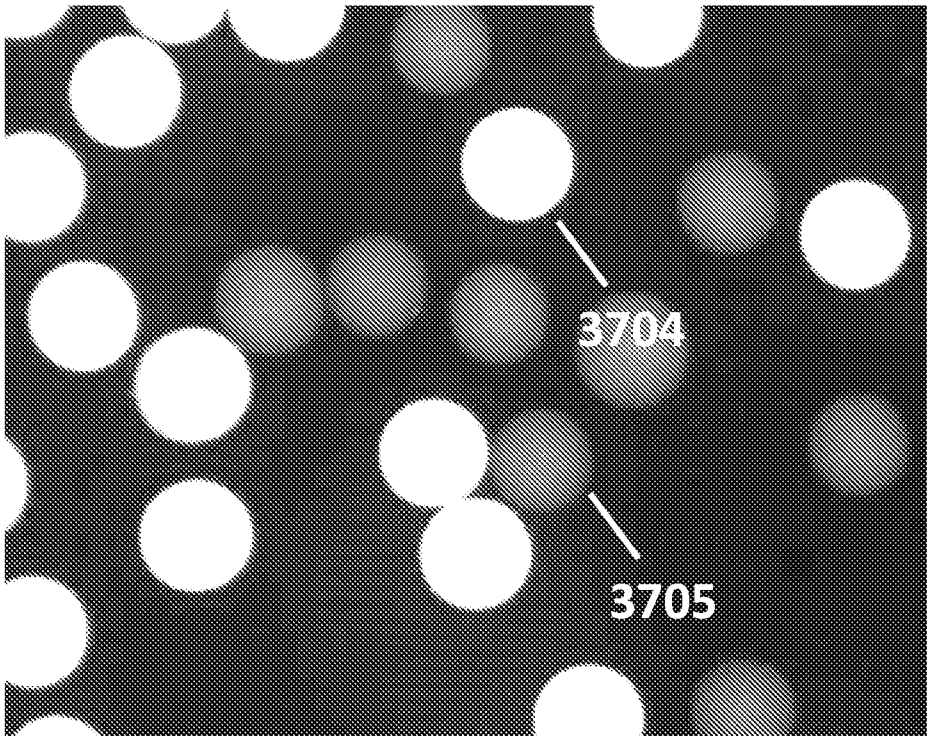


Fig. 37A



1: 50 bp ladder
2: FCBC:GB = 400 million
3: FCBC:GB = 1.6 billion
4: Acrydite Polymerization Chemistry

Fig. 37B



Bright beads: Acrydite Polymerization Chemistry
(Functionalization During Gel Bead Production)

Dim beads: Thiol-Disulfide Exchange Chemistry
(Functionalization Post-Production of Gel Beads)

Fig. 37C

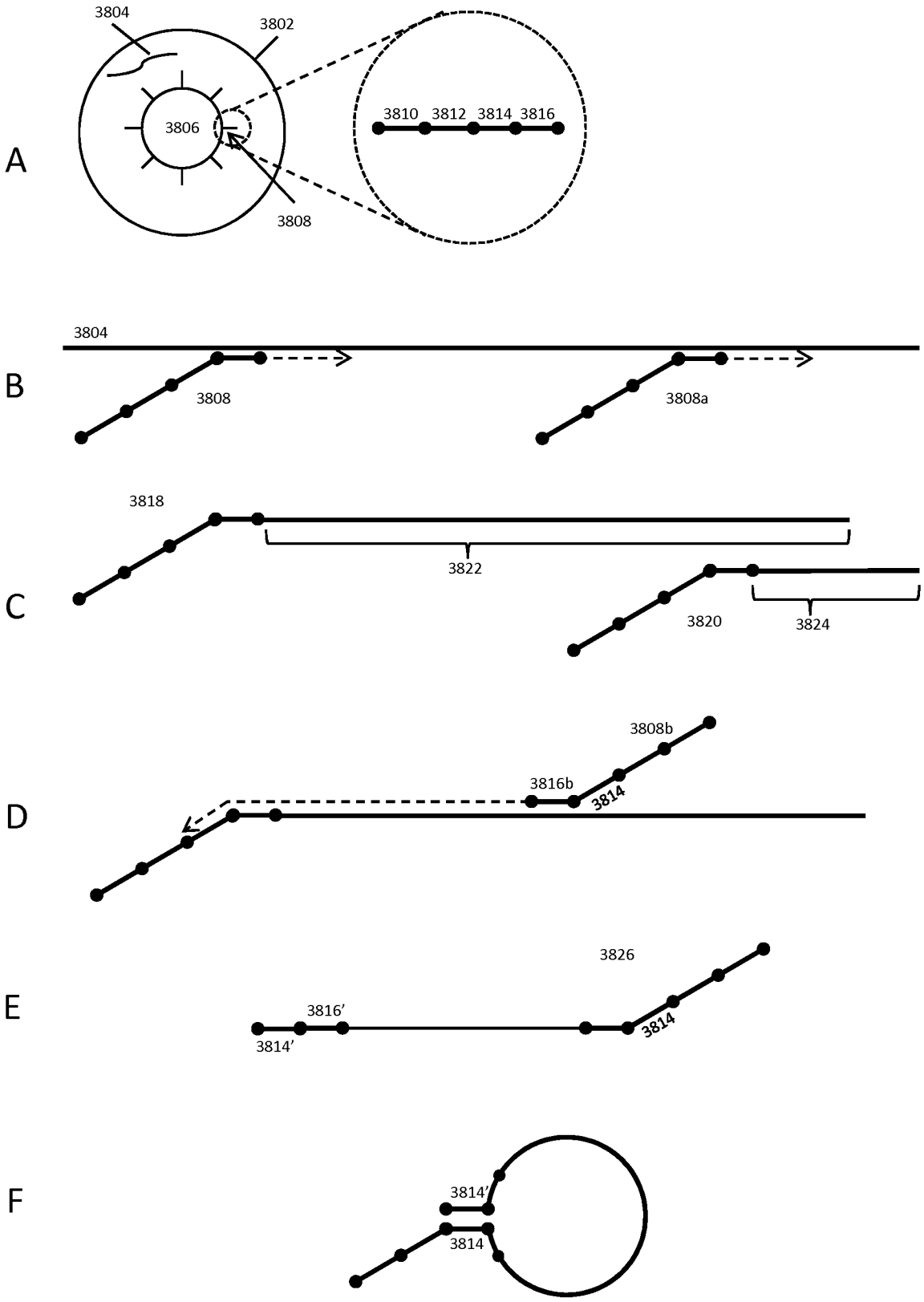

A
3804
3802
3806
3808
3810 3812 3814 3816
B
3804
3808
3808a
C
3818
3822
3820
3824
D
3808b
3816b
3814
E
3826
3816'
3814'
3814
F
3814'
3814

Fig. 38

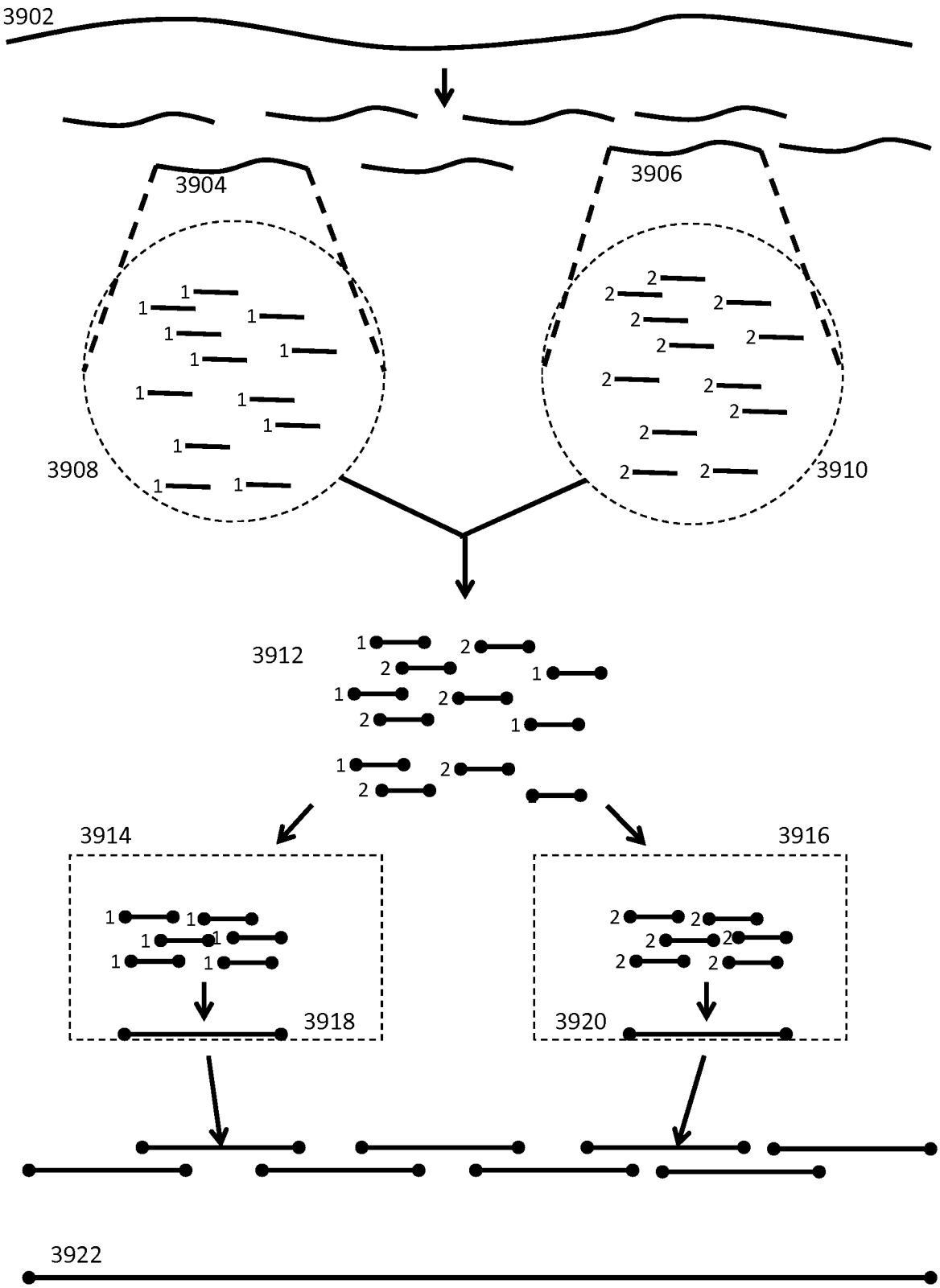


Fig. 39

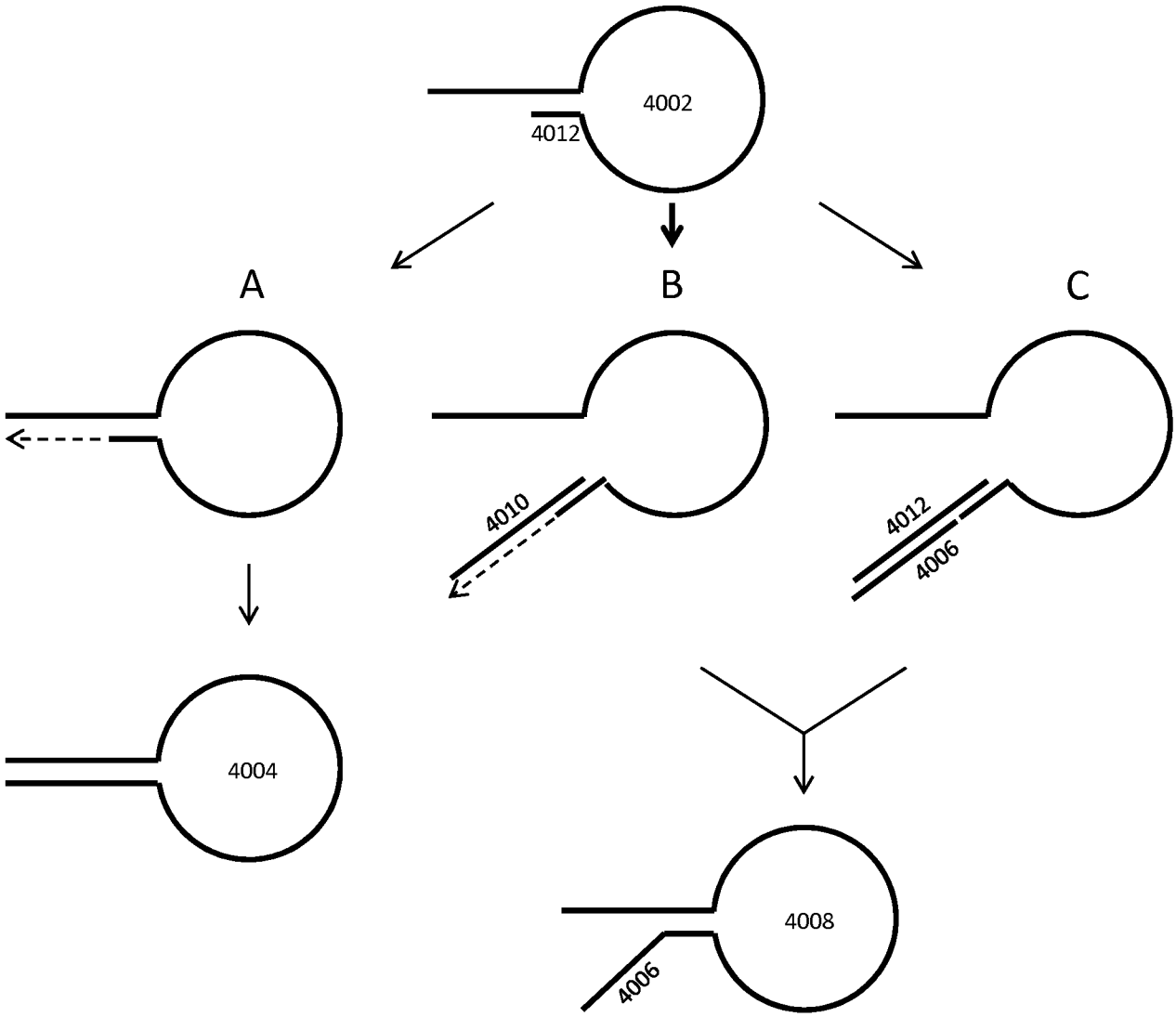


Fig. 40