

SPATIALLY RESOLVED SURFACE CAPTURE OF NUCLEIC ACIDS**CROSS-REFERENCE TO RELATED APPLICATIONS**

[0001] This application claims the benefit of and priority to U.S. Provisional Patent Application No. 63/398,183, filed August 15, 2022, the contents of which are incorporated by reference herein in their entirety.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0002] The contents of the electronic sequence listing (ELEM_014_001WO_SeqList_ST26.xml; Size 9,927 bytes; and Date of Creation: August 9, 2023) are herein incorporated by reference in their entireties.

TECHNICAL FIELD

[0003] The present disclosure provides compositions and methods for capturing nucleic acids from cellular samples on a support, preparing library molecules on the support, amplifying the library molecules on the support to generate nucleic acid template molecules, and analyzing the immobilized nucleic acid template molecules including detecting and/or sequencing the immobilized nucleic acid template molecules. The immobilized nucleic acid template molecules correspond to the nucleic acids from the cellular samples. The immobilized nucleic acid template molecules are spatially located on the support in an arrangement that is similar to the spatial location in the cellular sample.

BACKGROUND

[0004] Cells within a tissue of a subject have differences in cell morphology and/or function due to varied analyte levels (e.g., gene and/or protein expression) within the different cells. The specific position of a cell within a tissue (e.g., the cell's position relative to neighboring cells or the cell's position relative to the tissue microenvironment) can affect, e.g., the cell's morphology, differentiation, fate, viability, proliferation, behavior, and signaling and crosstalk with other cells in the tissue.

[0005] Spatial heterogeneity has been previously studied using techniques that only provide data for a small handful of analytes in the context of an intact tissue or a portion of a tissue, or provide a lot of analyte data for single cells, but fail to provide information regarding the position of the single cell in a parent biological sample (e.g., tissue sample).

[0006] Spatial analysis of an analyte within a biological sample may require determining the sequence of the analyte sequence or a complement thereof and the sequence of the spatial barcode or a complement thereof to identify the location of the analyte. The biological sample may be placed on a solid support to improve specificity and efficiency when being analyzed for identification or characterization of an analyte, such as DNA, RNA or other genetic material, within the sample. There exists a need for improved spatially resolved surface capture of nucleic acids on a support, and downstream assembly and sequencing techniques, which reduce the assembly, time and/or computational requirements necessary to obtain spatial sequence information. Provided herein are compositions, apparatuses and methods addressing this need.

SUMMARY

[0007] In one aspect the present disclosure provides a method for preparing spatially resolved nucleic acids, comprising: **a)** providing a support comprising (i) a low non-specific binding coating comprising at least one hydrophilic polymer, wherein the low non-specific binding coating has a water contact angle of no more than 45 degrees, and (ii) a plurality of immobilized surface capture primers covalently tethered to the low non-specific binding coating wherein each individual surface capture primer comprises a universal surface capture primer sequence (110), a universal binding site for a reverse sequencing primer (120), and an RNA capture sequence; **b)** positioning a cellular sample on the coated support under a condition suitable for the cellular sample to remain in a fixed position on the coated support, disrupting the cellular sample, releasing a plurality of RNA molecules of the cellular sample under a condition that is suitable for preserving spatial location information of the RNA molecules from the cellular sample, and collecting the plurality of RNA molecules on the plurality of immobilized surface capture primers of the support, under a condition suitable for hybridizing individual RNA molecules to individual immobilized surface capture primers to generate a plurality of capture primer-RNA duplexes, wherein each primer-RNA duplex comprises an immobilized surface capture primer hybridized to an RNA molecule; **c)** conducting a reverse transcription reaction on the coated support under a condition suitable for extending the 3' end of the immobilized surface capture primers and using the hybridized RNA as a template strand, under a condition suitable for generating a non-template polyC tail at the 3' end, thereby generating a plurality of first strand cDNA molecules with a non-template polyC tail at the 3' end, wherein the reverse transcription reaction comprises a plurality of template switching oligonucleotides, wherein each of the template switching

oligonucleotides comprises a polyG region that is capable of hybridizing to the non-template polyC tail of the first strand cDNA molecules, a universal binding site for a forward sequencing primer (130), and a universal binding site for a surface pinning primer (140); **d**) conducting a primer extension reaction from the 3' terminal end of the non-template polyC tail of the plurality of first strand cDNA molecules and using the template switching oligonucleotide as a template strand, thereby generating a plurality of immobilized full length first strand cDNAs, wherein each full length first strand cDNA comprises a universal surface capture primer sequence (110); a universal binding site for a reverse sequencing primer (120); an RNA capture sequence; a cDNA insert region; a polyC tail region; a universal binding site for a forward sequencing primer (130); and a universal binding site for a surface pinning primer (140); **e**) removing the RNA molecule while retaining the plurality of immobilized full length first strand cDNA; **f**) contacting the plurality of retained immobilized full length first strand cDNAs with a plurality of single stranded circularization oligonucleotides under a condition suitable for hybridizing individual circularization oligonucleotides to an immobilized full length first strand cDNA, thereby forming a single stranded circular molecule with a gap, wherein individual single stranded circularization oligonucleotides comprise: (i) a sequence that is complementary to the universal binding site for a reverse sequencing primer (120), (ii) a sequence that is complementary to the universal surface capture primer sequence (110), (iii) a linker region, (iv) a sequence that is complementary to the universal binding site for a surface pinning primer (140), (v) and a sequence that is complementary to the universal binding site for a forward sequencing primer (130); **g**) conducting a polymerase-catalyzed extension reaction to fill in the gap, using the polyC tail region, the cDNA insert region and the RNA capture sequence of the immobilized full length first strand cDNA as a template strand, wherein the polymerase-catalyzed extension reaction forms a single stranded circularized molecule with a nick, and wherein the nick is closed by conducting an enzymatic ligation reaction to generate a single stranded covalently closed circular molecule which is hybridized to the immobilized full length first strand cDNA; **h**) conducting a rolling circle amplification reaction using the terminal 3' end of the immobilized full length first strand cDNA as an initiation site and the covalently closed circular molecule as a template strand, thereby generating a plurality of immobilized nucleic acid concatemers that are spatially resolved on the support; **i**) sequencing the plurality of individual immobilized nucleic acid concatemers, wherein the sequencing comprises at least a portion of the cDNA insert region of the individual nucleic acid concatemers, which correspond to the individual RNA molecules eluted from the

cellular sample; and **j**) determining the location of individual nucleic acid concatemers on the coated support, which correspond to the spatial location of the individual RNA molecules eluted from the cellular sample.

[0008] In some embodiments, the cellular sample of step **b**) comprises a single cell, a plurality of cells, a tissue, an organ, an organism, or a sectioned cellular sample. In some embodiments, the cellular sample of step **b**) comprises a fresh sample, a frozen sample, a fresh frozen sample, or a formalin-fixed paraffin-embedded sample.

[0009] In some embodiments, the plurality of template switching oligonucleotides of step **c**) comprises chimeric DNA and/or RNA oligonucleotides.

[0010] In some embodiments, the linker region of the single stranded circularization oligonucleotides of step **f**) comprises at least one sample index sequence for multiplexing, at least one unique molecular index (UMI) sequence for molecular tagging and/or at least one universal binding site for a compaction oligonucleotide.

[0011] In some embodiments, the rolling circle amplification reaction of step **h**) is conducted in the presence of a plurality of compaction oligonucleotides, wherein the compaction oligonucleotides comprise single stranded oligonucleotides each having a first region at a first distal end that hybridizes to one portion of a concatemer molecule and a second region at the second distal end that hybridizes to a second portion of the same concatemer molecule, wherein the first portion of the concatemer molecule and the second portion of the concatemer molecule are brought in proximity, thereby causing compaction of the concatemer molecule to form a DNA nanoball.

[0012] In some embodiments, the sequencing of step **i**) comprises: **a.** contacting the plurality of concatemer molecules with a plurality of sequencing polymerases and a plurality of nucleic acid sequencing primers, wherein the contacting is conducted under a condition suitable to form a plurality of complexed sequencing polymerases, wherein each complexed sequencing polymerase comprises a sequencing polymerase bound to a nucleic acid duplex, and wherein the nucleic acid duplex comprises a portion of a concatemer molecule hybridized to the nucleic acid sequencing primer; **b.** contacting the plurality of complexed sequencing polymerases with a plurality of detectably labeled nucleotides comprising a blocking moiety at the 2' or 3' sugar position, wherein the contacting is conducted under a condition suitable for binding at least one nucleotide to at least one of the complexed sequencing polymerases, and wherein the condition is suitable for promoting polymerase-catalyzed nucleotide incorporation; **c.** incorporating a nucleotide into the 3' end of a sequencing primer of at least one complexed sequencing polymerase; **d.** detecting the

incorporated nucleotide and identifying the nucleo-base of the incorporated nucleotide; **e.** removing the blocking moiety from the incorporated nucleotide; and **f.** repeating steps b. to e. at least once.

[0013] In some embodiments, the sequencing of step **i)** comprises: **a.** contacting the plurality of concatemer molecules with a plurality of sequencing polymerases and a plurality of nucleic acid sequencing primers, wherein the contacting is conducted under a condition suitable to form a plurality of complexed sequencing polymerases, wherein each complexed sequencing polymerase comprises a sequencing polymerase bound to a nucleic acid duplex, and wherein the nucleic acid duplex comprises a portion of a concatemer molecule hybridized to the nucleic acid sequencing primer; **b.** contacting the plurality of complexed sequencing polymerases with a plurality of nucleotides each comprising a detectable label attached to a phosphate moiety of the phosphate chain, wherein the contacting is conducted under a condition suitable for binding at least one nucleotide to at least one of the complexed sequencing polymerases, and wherein the condition is suitable for promoting polymerase-catalyzed nucleotide incorporation; **c.** incorporating a nucleotide into the 3' end of a sequencing primer of at least one complexed sequencing polymerase; **d.** detecting the incorporated nucleotide and identifying the nucleo-base of the incorporated nucleotide; and **e.** repeating steps b. to d. at least once.

[0014] In some embodiments, the sequencing of step **i)** comprises: **a.** contacting the plurality of concatemer molecules with a plurality of a first sequencing polymerase and a plurality of nucleic acid sequencing primers, wherein the contacting is conducted under a condition suitable to bind the plurality of first polymerases to the plurality of nucleic acid template molecules and the plurality of nucleic acid primers, thereby forming a plurality of first complexed polymerases, wherein each first complexed polymerase comprises a first polymerase bound to a nucleic acid duplex, wherein the nucleic acid duplex comprises a nucleic acid template molecule hybridized to a nucleic acid primer; **b.** contacting the plurality of first complexed polymerases with a plurality of detectably labeled multivalent molecules to form a plurality of multivalent-binding complexes, wherein individual detectably labeled multivalent molecules in the plurality comprise a core attached to multiple nucleotide arms, and wherein each nucleotide arm is attached to a nucleotide unit, wherein the contacting is conducted under a condition suitable for binding complementary nucleotide units of the multivalent molecules to at least two of the plurality of first complexed polymerases thereby forming a plurality of multivalent-binding complexes, and wherein the condition is suitable for inhibiting incorporation of the complementary nucleotide units into the primers of the

plurality of multivalent-binding complexes; **c.** detecting the plurality of multivalent-binding complexes; and **d.** identifying the nucleobase of the complementary nucleotide units in the plurality of multivalent-binding complexes, thereby determining the sequence of the nucleic acid template molecules.

[0015] In some embodiments, the sequencing of step **i)** comprises: **a.** dissociating the plurality of multivalent-binding complexes, by removing the plurality of first sequencing polymerases and their bound multivalent molecules, and retaining the plurality of nucleic acid duplexes; **b.** contacting the plurality of the retained nucleic acid duplexes of step **e)** with a plurality of second sequencing polymerases under a condition suitable for binding the plurality of second polymerases to the plurality of the retained nucleic acid duplexes, thereby forming a plurality of second complexed polymerases each comprising a second polymerase bound to a nucleic acid duplex; and **c.** contacting the plurality of second complexed polymerases with a plurality of nucleotides, wherein the contacting is conducted under a condition suitable for binding complementary nucleotides from the plurality of nucleotides to at least two of the second complexed polymerases, thereby forming a plurality of nucleotide-binding complexes, and wherein the condition is suitable for promoting nucleotide incorporation of the bound complementary nucleotides into the primers of the nucleotide-binding complexes.

[0016] In some embodiments, the method further comprises: **d.** detecting the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases. In some embodiments, the method further comprises: **d.** detecting the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases; and **e.** identifying the nucleobases of the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases.

[0017] In some embodiments, the contacting the plurality of first complexed polymerases with the plurality of multivalent molecules of step **b.** is conducted in the presence of a non-catalytic divalent cation that inhibits polymerase-catalyzed nucleotide incorporation, wherein the non-catalytic divalent cation comprises strontium, barium or calcium.

[0018] In some embodiments, the individual multivalent molecules in the plurality of multivalent molecules comprise: **(a)** a core; and **(b)** a plurality of nucleotide arms which comprise: **(i)** a core attachment moiety, **(ii)** a spacer, **(iii)** a linker, and **(iv)** a nucleotide unit, wherein the core is attached to the plurality of nucleotide arms via their core attachment

moiety, wherein the spacer is attached to the linker, and wherein the linker is attached to the nucleotide unit.

[0019] In some embodiments, the linker comprises an aliphatic chain having 2-6 subunits or an oligo ethylene glycol chain having 2-6 subunits.

[0020] In some embodiments, the plurality of nucleotide arms attached to a given core have the same type of nucleotide units, and wherein the types of nucleotide units comprise dATP, dGTP, dCTP, dTTP or dUTP.

[0021] In some embodiments, the plurality of multivalent molecules comprise one type of a multivalent molecule wherein each multivalent molecule in the plurality has the same type of nucleotide unit selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP. In some embodiments, the plurality of multivalent molecules comprise a mixture of any combination of two or more types of multivalent molecules each type having nucleotide units selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP.

[0022] In some embodiments, the method further comprises forming a plurality of binding complexes, comprising the steps of: **a)** binding a first nucleic acid sequencing primer, a first sequencing polymerase, and a first multivalent molecule to a first portion of a concatemer molecule, thereby forming a first binding complex, wherein a first nucleotide unit of the first multivalent molecule binds to the first polymerase; and **b)** binding a second nucleic acid sequencing primer, a second sequencing polymerase, and the first multivalent molecule to a second portion of the same concatemer molecule, thereby forming a second binding complex, wherein a second nucleotide unit of the first multivalent molecule binds to the second polymerase, wherein the first and the second binding complexes which include the same multivalent molecule forms an avidity complex.

[0023] In some embodiments, the method further comprises forming an avidity complex comprising the steps of: **a)** contacting the plurality of the first sequencing polymerase and the plurality of nucleic acid sequencing primers with different portions of a concatemer nucleic acid template molecule to form at least first complexed polymerase and a second complexed polymerase on the same concatemer molecule; and **b)** contacting a plurality of detectably labeled multivalent molecules to the at least first complexed polymerase and the second complexed polymerase on the same concatemer molecule, under conditions suitable to bind a single multivalent molecule from the plurality to the first complexed polymerase and the second complexed polymerase, wherein at least a first nucleotide unit of the single multivalent molecule is bound to the first complexed polymerase, which comprises a first primer hybridized to a first portion of the concatemer molecule thereby forming a first

binding complex, and wherein at least a second nucleotide unit of the single multivalent molecule is bound to the second complexed polymerase which comprises a second primer hybridized to a second portion of the concatemer molecule, thereby forming a second binding complex, wherein the contacting is conducted under a condition suitable to inhibit polymerase-catalyzed incorporation of the bound first and second nucleotide units in the first and second binding complexes, and wherein the first and second binding complexes which are bound to the same multivalent molecule forms an avidity complex; **c)** detecting the first and second binding complexes on the same concatemer molecule; and **d)** identifying the first nucleotide unit in the first binding complex, thereby determining the sequence of the first portion of the concatemer molecule, and identifying the second nucleotide unit in the second binding complex thereby determining the sequence of the second portion of the concatemer molecule.

[0024] In some embodiments, the contacting the plurality of second complexed polymerases with the plurality of nucleotides of step **c.** is conducted in the presence of a catalytic divalent cation that promotes polymerase-catalyzed nucleotide incorporation, wherein the catalytic divalent cation comprises magnesium or manganese.

[0025] In some embodiments, the individual nucleotides in the plurality of nucleotides in step **c.** comprise an aromatic base, a five-carbon sugar, and 1-10 phosphate groups. In some embodiments, the plurality of nucleotides of step **c.** comprises one type of nucleotide selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP, or comprise a mixture of any combination of two or more types of nucleotides selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP.

[0026] In some embodiments, at least one of the nucleotides in the plurality of nucleotides in step **c.** is labeled with a fluorophore. In some embodiments, the plurality of nucleotides in step **c.** lack a fluorophore label.

[0027] In some embodiments, at least one of the nucleotides in the plurality of nucleotides of step **c.** comprises a removable chain terminating moiety attached to the 3' carbon position of the sugar group, wherein the removable chain terminating moiety comprises an alkyl group, alkenyl group, alkynyl group, allyl group, aryl group, benzyl group, azide group, azido group, O-azidomethyl group, amine group, amide group, keto group, isocyanate group, phosphate group, thio group, disulfide group, carbonate group, urea group, or silyl group, and wherein the removable chain terminating moiety is cleavable with a chemical compound to generate an extendible 3'OH moiety on the sugar group.

DESCRIPTION OF THE DRAWINGS

[0028] The features of the present disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings of which:

[0029] **FIG. 1** is a schematic of an exemplary low binding support comprising a glass substrate and alternating layers of hydrophilic coatings which are covalently or non-covalently adhered to the glass, and which further comprises chemically-reactive functional groups that serve as attachment sites for oligonucleotide primers (*e.g.*, capture oligonucleotides). In an alternative embodiment, the support can be made of any material such as glass, plastic or a polymer material.

[0030] **FIG. 2** is a schematic of various exemplary configurations of multivalent molecules. **Left (Class I):** schematics of multivalent molecules having a “starburst” or “helter-skelter” configuration. **Center (Class II):** a schematic of a multivalent molecule having a dendrimer configuration. **Right (Class III):** a schematic of multiple multivalent molecules formed by reacting streptavidin with 4-arm or 8-arm PEG-NHS with biotin and dNTPs. Nucleotide units are designated ‘N’, biotin is designated ‘B’, and streptavidin is designated ‘SA’.

[0031] **FIG. 3** is a schematic of an exemplary multivalent molecule comprising a generic core attached to a plurality of nucleotide-arms.

[0032] **FIG. 4** is a schematic of an exemplary multivalent molecule comprising a dendrimer core attached to a plurality of nucleotide-arms.

[0033] **FIG. 5** shows a schematic of an exemplary multivalent molecule comprising a core attached to a plurality of nucleotide-arms, where the nucleotide arms comprise biotin, spacer, linker and a nucleotide unit.

[0034] **FIG. 6** is a schematic of an exemplary nucleotide-arm comprising a core attachment moiety, spacer, linker and nucleotide unit.

[0035] **FIG. 7** shows the chemical structure of an exemplary spacer (**TOP**), and the chemical structures of various exemplary linkers, including an 11-atom Linker, 16-atom Linker, 23-atom Linker and an N3 Linker (**BOTTOM**).

[0036] **FIG. 8** shows the chemical structures of various exemplary linkers, including Linkers 1-9.

[0037] **FIG. 9A-9D** show the chemical structures of various exemplary linkers joined/attached to nucleotide units.

[0038] FIG. 10 shows the chemical structure of an exemplary biotinylated nucleotide-arm. In this example, the nucleotide unit is connected to the linker via a propargyl amine attachment at the 5 position of a pyrimidine base or the 7 position of a purine base.

[0039] FIG. 11 is a schematic of a support comprising an immobilized surface capture primer.

[0040] FIG. 12 is a schematic showing a cellular sample positioned on the support and an RNA molecule eluted from the cellular sample onto the surface capture primer immobilized to the support.

[0041] FIG. 13 is a schematic showing synthesis of a first strand cDNA (dashed line) and non-template polyC tail at the 3' end of the first strand cDNA.

[0042] FIG. 14 is a schematic showing a template switching oligonucleotide which is hybridized to the non-template polyC tail at the 3' end of the first strand cDNA. rGrGrG represents an exemplary polyG region. The polyG region of the template switching oligonucleotide may comprise ribonucleotide bases.

[0043] FIG. 15 is a schematic showing extension of the first strand cDNA (upper dashed arrow) using the template switching oligonucleotide as a template strand to generate a full length first strand cDNA.

[0044] FIG. 16 is a schematic showing the RNA strand has been removed/degraded while retaining the full length first strand cDNA which is immobilized to the support.

[0045] FIG. 17 is a schematic showing hybridizing a circularization oligonucleotide to the immobilized full length first strand cDNA to form a single stranded circular molecule with a gap.

[0046] FIG. 18 is a schematic showing a polymerase-catalyzed extension reaction to fill in the gap resulting in a nick. A ligation reaction can be conducted to ligate the nick.

[0047] FIG. 19 is a schematic showing the initial stage of a rolling circle amplification reaction on the support.

[0048] FIG. 20 is a schematic of a guanine tetrad (*e.g.*, G-tetrad).

[0049] FIG. 21 is a schematic of an exemplary intramolecular G-quadruplex structure.

DETAILED DESCRIPTION

[0050] **Definitions:**

[0051] The headings provided herein are not limitations of the various aspects of the disclosure, which aspects can be understood by reference to the specification as a whole.

[0052] Unless defined otherwise, technical and scientific terms used herein have meanings that are commonly understood by those of ordinary skill in the art unless defined otherwise. Generally, terminologies pertaining to techniques of molecular biology, nucleic acid chemistry, protein chemistry, genetics, microbiology, transgenic cell production, and hybridization described herein are those well-known and commonly used in the art. Techniques and procedures described herein are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the instant specification. For example, *see* Sambrook et al., *Molecular Cloning: A Laboratory Manual* (Third ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 2000). See also Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing Associates (1992). The nomenclatures utilized in connection with, and the laboratory procedures and techniques described herein are those well-known and commonly used in the art.

[0053] Unless otherwise required by context herein, singular terms shall include pluralities and plural terms shall include the singular. Singular forms “a”, “an” and “the”, and singular use of any word, include plural referents unless expressly and unequivocally limited on one referent.

[0054] It is understood the use of the alternative term (*e.g.*, “or”) is taken to mean either one or both or any combination thereof of the alternatives.

[0055] The term “and/or” used herein is to be taken mean specific disclosure of each of the specified features or components with or without the other. For example, the term “and/or” as used in a phrase such as “A and/or B” herein is intended to include: “A and B”; “A or B”; “A” (A alone); and “B” (B alone). In a similar manner, the term “and/or” as used in a phrase such as “A, B, and/or C” is intended to encompass each of the following aspects: “A, B, and C”; “A, B, or C”; “A or C”; “A or B”; “B or C”; “A and B”; “B and C”; “A and C”; “A” (A alone); “B” (B alone); and “C” (C alone).

[0056] As used herein and in the appended claims, terms “comprising”, “including”, “having” and “containing”, and their grammatical variants, as used herein are intended to be non-limiting so that one item or multiple items in a list do not exclude other items that can be substituted or added to the listed items. It is understood that wherever aspects are described herein with the language “comprising,” otherwise analogous aspects described in terms of “consisting of” and/or “consisting essentially of” are also provided.

[0057] As used herein, the terms “about” and “approximately” refer to a value or composition that is within an acceptable error range for the particular value or composition as

determined by one of ordinary skill in the art, which will depend in part on how the value or composition is measured or determined, i.e., the limitations of the measurement system. For example, “about” or “approximately” can mean within one or more than one standard deviation per the practice in the art. Alternatively, “about” or “approximately” can mean a range of up to 10% (i.e., $\pm 10\%$) or more depending on the limitations of the measurement system. For example, about 5 mg can include any number between 4.5 mg and 5.5 mg. Furthermore, particularly with respect to biological systems or processes, the terms can mean up to an order of magnitude or up to 5-fold of a value. When particular values or compositions are provided in the instant disclosure, unless otherwise stated, the meaning of “about” or “approximately” should be assumed to be within an acceptable error range for that particular value or composition. Also, where ranges and/or subranges of values are provided, the ranges and/or subranges can include the endpoints of the ranges and/or subranges.

[0058] The term “polymerase” and its variants, as used herein, comprises an enzyme comprising a domain that binds a nucleotide (or nucleoside) where the polymerase can form a complex having a template nucleic acid and a complementary nucleotide. The polymerase can have one or more activities including, but not limited to, base analog detection activities, DNA polymerization activity, reverse transcriptase activity, DNA binding, strand displacement activity, and nucleotide binding and recognition. A polymerase can be any enzyme that can catalyze polymerization of nucleotides (including analogs thereof) into a nucleic acid strand. Typically, but not necessarily, such nucleotide polymerization can occur in a template-dependent fashion. Typically, a polymerase comprises one or more active sites at which nucleotide binding and/or catalysis of nucleotide polymerization can occur. In some embodiments, a polymerase includes other enzymatic activities, such as for example, 3' to 5' exonuclease activity or 5' to 3' exonuclease activity. In some embodiments, a polymerase has strand displacing activity. A polymerase can include without limitation naturally occurring polymerases and any subunits and truncations thereof, mutant polymerases, variant polymerases, recombinant, fusion or otherwise engineered polymerases, chemically modified polymerases, synthetic molecules or assemblies, and any analogs, derivatives, or fragments thereof that retain the ability to catalyze nucleotide polymerization (*e.g.*, catalytically active fragment). The polymerase includes catalytically inactive polymerases, catalytically active polymerases, reverse transcriptases, and other enzymes comprising a nucleotide binding domain. In some embodiments, a polymerase can be isolated from a cell, or generated using recombinant DNA technology or chemical synthesis methods. In some embodiments, a polymerase can be expressed in prokaryote, eukaryote, viral, or phage organisms. In some

embodiments, a polymerase can be post-translationally modified proteins or fragments thereof. A polymerase can be derived from a prokaryote, eukaryote, virus or phage. A polymerase comprises DNA-directed DNA polymerase and RNA-directed DNA polymerase.

[0059] As used herein, the term “strand displacing” refers to the ability of a polymerase to locally separate strands of double-stranded nucleic acids and synthesize a new strand in a template-based manner. Strand displacing polymerases displace a complementary strand from a template strand and catalyze new strand synthesis. Strand displacing polymerases include mesophilic and thermophilic polymerases. Strand displacing polymerases include wild type enzymes, and variants including exonuclease minus mutants, mutant versions, chimeric enzymes and truncated enzymes. Examples of strand displacing polymerases include, without limitation, phi29 DNA polymerase, large fragment of Bst DNA polymerase, large fragment of Bsu DNA polymerase (exo-), Bca DNA polymerase (exo-), Klenow fragment of E. coli DNA polymerase, T5 polymerase, M-MuLV reverse transcriptase, HIV viral reverse transcriptase, Deep Vent DNA polymerase and KOD DNA polymerase. The phi29 DNA polymerase can be wild type phi29 DNA polymerase (*e.g.*, MagniPhi from Expedeon), or variant EquiPhi29 DNA polymerase (*e.g.*, from Thermo Fisher Scientific), or chimeric QualiPhi DNA polymerase (*e.g.*, from 4basebio).

[0060] The terms “nucleic acid”, “polynucleotide” and “oligonucleotide” and other related terms used herein are used interchangeably and refer to polymers of nucleotides and are not limited to any particular length. Nucleic acids include recombinant and chemically-synthesized forms. Nucleic acids can be isolated. Nucleic acids include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs (*e.g.*, peptide nucleic acids (PNA) and non-naturally occurring nucleotide analogs), and chimeric forms containing DNA and RNA. Nucleic acids can be single-stranded or double-stranded. Nucleic acids comprise polymers of nucleotides, where the nucleotides may include natural or non-natural bases, and/or sugars. Nucleic acids comprise naturally-occurring internucleosidic linkages, for example phosphodiester linkages. Nucleic acids can lack a phosphate group. Nucleic acids can comprise non-natural internucleoside linkages, including phosphorothioate, phosphorothiolate, and/or peptide nucleic acid (PNA) linkages. In some embodiments, nucleic acids comprise one type of polynucleotides or a mixture of two or more different types of polynucleotides.

[0061] The term “operably linked” and “operably joined” or related terms as used herein refers to juxtaposition of components. The juxtapositioned components can be linked together covalently. For example, two nucleic acid components can be enzymatically ligated together

where the linkage that joins together the two components comprises phosphodiester linkage. A first and second nucleic acid component can be linked together, where the first nucleic acid component can confer a function on a second nucleic acid component. For example, linkage between a primer binding sequence and a sequence of interest forms a nucleic acid library molecule having a portion that can bind to a primer. In another example, a transgene (*e.g.*, a nucleic acid encoding a polypeptide or a nucleic acid sequence of interest) can be ligated to a vector where the linkage permits expression or functioning of the transgene sequence contained in the vector. In some embodiments, a transgene is operably linked to a host cell regulatory sequence (*e.g.*, a promoter sequence) that affects expression of the transgene. In some embodiments, the vector comprises at least one host cell regulatory sequence, including a promoter sequence, enhancer, transcription and/or translation initiation sequence, transcription and/or translation termination sequence, polypeptide secretion signal sequences, and the like. In some embodiments, the host cell regulatory sequence controls expression of the level, timing and/or location of the transgene.

[0062] The terms “linked”, “joined”, “attached”, “appended” and variants thereof comprise any type of fusion, bond, adherence or association between any combination of compounds or molecules that is of sufficient stability to withstand use in the particular procedure. The procedure can include but is not limited to: nucleotide binding; nucleotide incorporation; de-blocking (*e.g.*, removal of chain-terminating moiety); washing; removing; flowing; detecting; imaging and/or identifying. Such linkage can comprise, for example, covalent, ionic, hydrogen, dipole-dipole, hydrophilic, hydrophobic, or affinity bonding, bonds or associations involving van der Waals forces, mechanical bonding, and the like. In some embodiments, such linkage occurs intramolecularly, for example linking together the ends of a single-stranded or double-stranded linear nucleic acid molecule to form a circular molecule. In some embodiments, such linkage can occur between a combination of different molecules, or between a molecule and a non-molecule, including but not limited to: linkage between a nucleic acid molecule and a solid surface; linkage between a protein and a detectable reporter moiety; linkage between a nucleotide and detectable reporter moiety; and the like. Some examples of linkages can be found, for example, in Hermanson, G., “Bioconjugate Techniques”, Second Edition (2008); Aslam, M., Dent, A., “Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences”, London: Macmillan (1998); Aslam, M., Dent, A., “Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences”, London: Macmillan (1998).

[0063] The term “primer” and related terms as used herein refers to an oligonucleotide that is capable of hybridizing with a DNA and/or RNA polynucleotide template to form a duplex molecule. Primers can comprise natural nucleotides and/or nucleotide analogs. Primers can be recombinant nucleic acid molecules. Primers may have any length, but typically range from 4-50 nucleotides. A typical primer comprises a 5' end and 3' end. The 3' end of the primer can include a 3' OH moiety which serves as a nucleotide polymerization initiation site in a polymerase-catalyzed primer extension reaction. Alternatively, the 3' end of the primer can lack a 3' OH moiety, or can include a terminal 3' blocking group that inhibits nucleotide polymerization in a polymerase-catalyzed reaction. Any one nucleotide, or more than one nucleotide, along the length of the primer can be labeled with a detectable reporter moiety. A primer can be in solution (*e.g.*, a soluble primer) or can be immobilized to a support (*e.g.*, a capture primer).

[0064] The term “template nucleic acid”, “template polynucleotide”, “target nucleic acid” “target polynucleotide”, “template strand” and other variations refer to a nucleic acid strand that serves as the basis nucleic acid molecule for any of the reiterative sequencing methods described herein. The template nucleic acid can be single-stranded or double-stranded, or the template nucleic acid can have single-stranded or double-stranded portions. The template nucleic acid can be obtained from a naturally-occurring source, recombinant form, or chemically synthesized to include any type of nucleic acid analog. The template nucleic acid can be linear, circular, or other forms. The template nucleic acids can include an insert portion having an insert sequence. The template nucleic acids can also include at least one adaptor sequence. The insert portion can be isolated in any form, including chromosomal, genomic, organellar (*e.g.*, mitochondrial, chloroplast or ribosomal), recombinant molecules, cloned, amplified, cDNA, RNA such as precursor mRNA or mRNA, oligonucleotides, whole genomic DNA, obtained from fresh frozen paraffin embedded tissue, needle biopsies, circulating tumor cells, cell free circulating DNA, or any type of nucleic acid library. The insert portion can be isolated from any source including from organisms such as prokaryotes, eukaryotes (*e.g.*, humans, plants and animals), fungus, viruses cells, tissues, normal or diseased cells or tissues, body fluids including blood, urine, serum, lymph, tumor, saliva, anal and vaginal secretions, amniotic samples, perspiration, semen, environmental samples, culture samples, or synthesized nucleic acid molecules prepared using recombinant molecular biology or chemical synthesis methods. The insert portion can be isolated from any organ, including head, neck, brain, breast, ovary, cervix, colon, rectum, endometrium, gallbladder, intestines, bladder, prostate, testicles, liver, lung, kidney, esophagus, pancreas, thyroid,

pituitary, thymus, skin, heart, larynx, or other organs. The template nucleic acid can be subjected to nucleic acid analysis, including sequencing and composition analysis.

[0065] The term “adaptor” and related terms refers to oligonucleotides that can be operably linked to a target polynucleotide, where the adaptor confers a function to the co-joined adaptor-target molecule. Adaptors comprise DNA, RNA, chimeric DNA/RNA, or analogs thereof. Adaptors can include at least one ribonucleoside residue. Adaptors can be single-stranded, double-stranded, or have single-stranded and/or double-stranded portions. Adaptors can be configured to be linear, stem-looped, hairpin, or Y-shaped forms. Adaptors can be any length, including 4-100 nucleotides or longer. Adaptors can have blunt ends, overhang ends, or a combination of both. Overhang ends include 5' overhang and 3' overhang ends. The 5' end of a single-stranded adaptor, or one strand of a double-stranded adaptor, can have a 5' phosphate group or lack a 5' phosphate group. Adaptors can include a 5' tail that does not hybridize to a target polynucleotide (*e.g.*, tailed adaptor), or adaptors can be non-tailed. An adaptor can include a sequence that is complementary to at least a portion of a primer, such as an amplification primer, a sequencing primer, or a capture primer (*e.g.*, soluble or immobilized capture primers). Adaptors can include a random sequence or degenerate sequence. Adaptors can include at least one inosine residue. Adaptors can include at least one phosphorothioate, phosphorothiolate and/or phosphoramidate linkage. Adaptors can include a barcode sequence which can be used to distinguish polynucleotides (*e.g.*, insert sequences) from different sample sources in a multiplex assay. Adaptors can include a unique identification sequence (*e.g.*, unique molecular index, UMI; or a unique molecular tag) that can be used to uniquely identify a nucleic acid molecule to which the adaptor is appended. In some embodiments, a unique identification sequence can be used to increase error correction and accuracy, reduce the rate of false-positive variant calls and/or increase sensitivity of variant detection. Adaptors can include at least one restriction enzyme recognition sequence, including any one or any combination of two or more selected from a group consisting of type I, type II, type III, type IV, type Hs or type IIB.

[0066] In some embodiments, any of the amplification primer sequences, sequencing primer sequences, capture primer sequences, target capture sequences, circularization anchor sequences, sample barcode sequences, spatial barcode sequences, or anchor region sequences can be about 3-50 nucleotides in length, or about 5-40 nucleotides in length, or about 5-25 nucleotides in length.

[0067] The term “universal sequence” and related terms refers to a sequence in a nucleic acid molecule that is common among two or more polynucleotide molecules. For example,

an adaptor having a universal sequence can be operably joined to a plurality of polynucleotides so that the population of co-joined molecules carry the same universal adaptor sequence. Examples of universal adaptor sequences include an amplification primer sequence, a sequencing primer sequence or a capture primer sequence (*e.g.*, soluble or immobilized capture primers).

[0068] When used in reference to nucleic acid molecules, the terms “hybridize” or “hybridizing” or “hybridization” or other related terms refers to hydrogen bonding between two different nucleic acids to form a duplex nucleic acid. Hybridization also includes hydrogen bonding between two different regions of a single nucleic acid molecule to form a self-hybridizing molecule having a duplex region. Hybridization can comprise Watson-Crick or Hoogsteen binding to form a duplex double-stranded nucleic acid, or a double-stranded region within a nucleic acid molecule. The double-stranded nucleic acid, or the two different regions of a single nucleic acid, may be wholly complementary, or partially complementary. Complementary nucleic acid strands need not hybridize with each other across their entire length. The complementary base pairing can be the standard A-T or C-G base pairing, or can be other forms of base-pairing interactions. Duplex nucleic acids can include mismatched base-paired nucleotides.

[0069] When used in reference to nucleic acids, the terms “extend”, “extending”, “extension” and other variants, refers to incorporation of one or more nucleotides into a nucleic acid molecule. Nucleotide incorporation comprises polymerization of one or more nucleotides into the terminal 3' OH end of a nucleic acid strand, resulting in extension of the nucleic acid strand. Nucleotide incorporation can be conducted with natural nucleotides and/or nucleotide analogs. Typically, but not necessarily, nucleotide incorporation occurs in a template-dependent fashion. Any suitable method of extending a nucleic acid molecule may be used, including primer extension catalyzed by a DNA polymerase or RNA polymerase.

[0070] The term “nucleotides” and related terms refers to a molecule comprising an aromatic base, a five-carbon sugar (*e.g.*, ribose or deoxyribose), and at least one phosphate group. Canonical or non-canonical nucleotides are consistent with use of the term. The phosphate in some embodiments comprises a monophosphate, diphosphate, or triphosphate, or corresponding phosphate analog. The term “nucleoside” refers to a molecule comprising an aromatic base and a sugar. Nucleotides and nucleosides can be non-labeled or labeled with a detectable reporter moiety.

[0071] Nucleotides (and nucleosides) typically comprise a hetero cyclic base including substituted or unsubstituted nitrogen-containing parent heteroaromatic ring which are

commonly found in nucleic acids, including naturally-occurring, substituted, modified, or engineered variants, or analogs of the same. The base of a nucleotide (or nucleoside) is capable of forming Watson-Crick and/or Hoogsteen hydrogen bonds with an appropriate complementary base. Exemplary bases include, but are not limited to, purines and pyrimidines such as: 2-aminopurine, 2,6-diaminopurine, adenine (A), ethenoadenine, N⁶-Δ²-isopentenyladenine (6iA), N⁶-Δ²-isopentenyl-2-methylthioadenine (2ms6iA), N⁶-methyladenine, guanine (G), isoguanine, N²-dimethylguanine (dmG), 7-methylguanine (7mG), 2-thiopyrimidine, 6-thioguanine (6sG), hypoxanthine and O⁶-methylguanine; 7-deaza-purines such as 7-deazaadenine (7-deaza-A) and 7-deazaguanine (7-deaza-G); pyrimidines such as cytosine (C), 5-propynylcytosine, isocytosine, thymine (T), 4-thiothymine (4sT), 5,6-dihydrothymine, O⁴-methylthymine, uracil (U), 4-thiouracil (4sU) and 5,6-dihydrouracil (dihydrouracil; D); indoles such as nitroindole and 4-methylindole; pyrroles such as nitropyrrole; nebularine; inosines; hydroxymethylcytosines; 5-methylcytosines; base (Y); as well as methylated, glycosylated, and acylated base moieties; and the like. Additional exemplary bases can be found in Fasman, 1989, in "Practical Handbook of Biochemistry and Molecular Biology", pp. 385-394, CRC Press, Boca Raton, Fla.

[0072] Nucleotides (and nucleosides) typically comprise a sugar moiety, such as carbocyclic moiety (Ferraro and Gotor 2000 Chem. Rev. 100: 4319-48), acyclic moieties (Martinez, et al., 1999 Nucleic Acids Research 27: 1271-1274; Martinez, et al., 1997 Bioorganic & Medicinal Chemistry Letters vol. 7: 3013-3016), and other sugar moieties (Joeng, et al., 1993 J. Med. Chem. 36: 2627-2638; Kim, et al., 1993 J. Med. Chem. 36: 30-7; Eschenmosser 1999 Science 284:2118-2124; and U.S. Pat. No. 5,558,991). The sugar moiety comprises: ribosyl; 2'-deoxyribosyl; 3'-deoxyribosyl; 2',3'-dideoxyribosyl; 2',3'-didehydrodideoxyribosyl; 2'-alkoxyribosyl; 2'-azidoribosyl; 2'-aminoribosyl; 2'-fluororibosyl; 2'-mercaptoriboxyl; 2'-alkylthioribosyl; 3'-alkoxyribosyl; 3'-azidoribosyl; 3'-aminoribosyl; 3'-fluororibosyl; 3'-mercaptoriboxyl; 3'-alkylthioribosyl carbocyclic; acyclic or other modified sugars.

[0073] In some embodiments, nucleotides comprise a chain of one, two or three phosphorus atoms where the chain is typically attached to the 5' carbon of the sugar moiety via an ester or phosphoramidate linkage. In some embodiments, the nucleotide is an analog having a phosphorus chain in which the phosphorus atoms are linked together with intervening O, S, NH, methylene or ethylene. In some embodiments, the phosphorus atoms in the chain include substituted side groups including O, S or BH₃. In some embodiments, the

chain includes phosphate groups substituted with analogs including phosphoramidate, phosphorothioate, phosphordithioate, and O-methylphosphoroamidite groups.

[0074] As used herein, a “nucleotide unit” or “nucleotide moiety” refers to nucleotides (*e.g.*, dATP, dTTP, dGTP, dCTP, or dUTP), or analogs thereof, comprising comprises a base, sugar and at least one phosphate group. Nucleotide units can be attached to the multivalent molecules used in the sequencing reactions described herein. In general, all nucleotide units attached to the same multivalent molecule will have the same identity (*e.g.*, all A, all T, all C, or all G), although the skilled artisan will appreciate that there may be situations in which a multivalent molecule comprising nucleotide units of differing identity will be advantageous.

[0075] The term “reporter moiety”, “reporter moieties” or related terms refers to a compound that generates, or causes to generate, a detectable signal. A reporter moiety is sometimes called a “label”. Any suitable reporter moiety may be used, including luminescent, photoluminescent, electroluminescent, bioluminescent, chemiluminescent, fluorescent, phosphorescent, chromophore, radioisotope, electrochemical, mass spectrometry, Raman, hapten, affinity tag, atom, or an enzyme. A reporter moiety generates a detectable signal resulting from a chemical or physical change (*e.g.*, heat, light, electrical, pH, salt concentration, enzymatic activity, or proximity events). A proximity event includes two reporter moieties approaching each other, or associating with each other, or binding each other. It is well known to one skilled in the art to select reporter moieties so that each absorbs excitation radiation and/or emits fluorescence at a wavelength distinguishable from the other reporter moieties to permit monitoring the presence of different reporter moieties in the same reaction or in different reactions. Two or more different reporter moieties can be selected having spectrally distinct emission profiles, or having minimal overlapping spectral emission profiles. Reporter moieties can be linked (*e.g.*, operably linked) to nucleotides, nucleosides, nucleic acids, enzymes (*e.g.*, polymerases or reverse transcriptases), or support (*e.g.*, surfaces).

[0076] A reporter moiety (or label) comprises a fluorescent label or a fluorophore. Exemplary fluorescent moieties which may serve as fluorescent labels or fluorophores include, but are not limited to fluorescein and fluorescein derivatives such as carboxyfluorescein, tetrachlorofluorescein, hexachlorofluorescein, carboxynaphthofluorescein, fluorescein isothiocyanate, NHS-fluorescein, iodoacetamidofluorescein, fluorescein maleimide, SAMSA-fluorescein, fluorescein thiosemicarbazide, carbohydrazinomethylthioacetyl-amino fluorescein, rhodamine and rhodamine derivatives such as TRITC, TMR, lissamine rhodamine, Texas Red, rhodamine B, rhodamine 6G,

rhodamine 10, NHS-rhodamine, TMR-iodoacetamide, lissamine rhodamine B sulfonyl chloride, lissamine rhodamine B sulfonyl hydrazine, Texas Red sulfonyl chloride, Texas Red hydrazide, coumarin and coumarin derivatives such as AMCA, AMCA-NHS, AMCA-sulfo-NHS, AMCA-HPDP, DCIA, AMCE-hydrazide, BODIPY and derivatives such as BODIPY FL C3-SE, BODIPY 530/550 C3, BODIPY 530/550 C3-SE, BODIPY 530/550 C3 hydrazide, BODIPY 493/503 C3 hydrazide, BODIPY FL C3 hydrazide, BODIPY FL IA, BODIPY 530/551 IA, Br-BODIPY 493/503, Cascade Blue and derivatives such as Cascade Blue acetyl azide, Cascade Blue cadaverine, Cascade Blue ethylenediamine, Cascade Blue hydrazide, Lucifer Yellow and derivatives such as Lucifer Yellow iodoacetamide, Lucifer Yellow CH, cyanine and derivatives such as indolium based cyanine dyes, benzo-indolium based cyanine dyes, pyridium based cyanine dyes, thiozolium based cyanine dyes, quinolinium based cyanine dyes, imidazolium based cyanine dyes, Cy 3, Cy5, lanthanide chelates and derivatives such as BCPDA, TBP, TMT, BHHCT, BCOT, Europium chelates, Terbium chelates, Alexa Fluor dyes, DyLight dyes, Atto dyes, LightCycler Red dyes, CAL Flour dyes, JOE and derivatives thereof, Oregon Green dyes, WellRED dyes, IRD dyes, phycoerythrin and phycobilin dyes, Malachite green, stilbene, DEG dyes, NR dyes, near-infrared dyes and others known in the art such as those described in Haugland, *Molecular Probes Handbook*, (Eugene, Oreg.) 6th Edition; Lakowicz, *Principles of Fluorescence Spectroscopy*, 2nd Ed., Plenum Press New York (1999), or Hermanson, *Bioconjugate Techniques*, 2nd Edition, or derivatives thereof, or any combination thereof. Cyanine dyes may exist in either sulfonated or non-sulfonated forms, and consist of two indolenin, benzo-indolium, pyridium, thiozolium, and/or quinolinium groups separated by a polymethine bridge between two nitrogen atoms. Commercially available cyanine fluorophores include, for example, Cy3, (which may comprise 1-[6-(2,5-dioxopyrrolidin-1-yloxy)-6-oxohexyl]-2-(3-{1-[6-(2,5-dioxopyrrolidin-1-yloxy)-6-oxohexyl]-3,3-dimethyl-1,3-dihydro-2H-indol-2-ylidene}prop-1-en-1-yl)-3,3-dimethyl-3H-indolium or 1-[6-(2,5-dioxopyrrolidin-1-yloxy)-6-oxohexyl]-2-(3-{1-[6-(2,5-dioxopyrrolidin-1-yloxy)-6-oxohexyl]-3,3-dimethyl-5-sulfo-1,3-dihydro-2H-indol-2-ylidene}prop-1-en-1-yl)-3,3-dimethyl-3H-indolium-5-sulfonate), Cy5 (which may comprise 1-(6-((2,5-dioxopyrrolidin-1-yl)oxy)-6-oxohexyl)-2-((1E,3E)-5-((E)-1-(6-((2,5-dioxopyrrolidin-1-yl)oxy)-6-oxohexyl)-3,3-dimethyl-5-indolin-2-ylidene)penta-1,3-dien-1-yl)-3,3-dimethyl-3H-indol-1-ium or 1-(6-((2,5-dioxopyrrolidin-1-yl)oxy)-6-oxohexyl)-2-((1E,3E)-5-((E)-1-(6-((2,5-dioxopyrrolidin-1-yl)oxy)-6-oxohexyl)-3,3-dimethyl-5-sulfoindolin-2-ylidene)penta-1,3-dien-1-yl)-3,3-dimethyl-3H-indol-1-ium-5-sulfonate), and Cy7 (which may comprise 1-(5-carboxypentyl)-2-[(1E,3E,5E,7Z)-7-(1-ethyl-1,3-dihydro-2H-

indol-2-ylidene)hepta-1,3,5-trien-1-yl]-3H-indolium or 1-(5-carboxypentyl)-2-[(1E,3E,5E,7Z)-7-(1-ethyl-5-sulfo-1,3-dihydro-2H-indol-2-ylidene)hepta-1,3,5-trien-1-yl]-3H-indolium-5-sulfonate), where “Cy” stands for 'cyanine', and the first digit identifies the number of carbon atoms between two indolenine groups. Cy2 which is an oxazole derivative rather than indolenin, and the benzo-derivatized Cy3.5, Cy5.5 and Cy7.5 are exceptions to this rule.

[0077] In some embodiments, the reporter moiety can be a FRET pair, *e.g.*, such that multiple classifications can be performed under a single excitation and imaging step. As used herein, FRET may comprise excitation exchange (Forster) transfers, or electron-exchange (Dexter) transfers.

[0078] When used in reference to nucleic acids, the terms “amplify”, “amplifying”, “amplification”, and other related terms include producing multiple copies of an original polynucleotide template molecule, where the copies comprise a sequence that is complementary to the template sequence, or the copies comprise a sequence that is the same as the template sequence. In some embodiments, the copies comprise a sequence that is substantially identical to a template sequence, or is substantially identical to a sequence that is complementary to the template sequence.

[0079] The term “support” as used herein refers to a substrate that is designed for deposition of biological molecules or biological samples for assays and/or analyses. Examples of biological molecules to be deposited onto a support include nucleic acids (*e.g.*, DNA, RNA), polypeptides, saccharides, lipids, a single cell or multiple cells. Examples of biological samples include, but are not limited to saliva, phlegm, mucus, blood, plasma, serum, urine, stool, sweat, tears and fluids from tissues or organs.

[0080] In some embodiments, the support is solid, semi-solid, or a combination of both. In some embodiments, the support is porous, semi-porous, non-porous, or any combination of porosity. In some embodiments, the support can be substantially planar, concave, convex, or any combination thereof. In some embodiments, the support can be cylindrical, for example comprising a capillary or interior surface of a capillary.

[0081] In some embodiments, the surface of the support can be substantially smooth. In some embodiments, the support can be regularly or irregularly textured, including bumps, etched, pores, three-dimensional scaffolds, or any combination thereof.

[0082] In some embodiments, the support comprises a bead having any shape, including spherical, hemi-spherical, cylindrical, barrel-shaped, toroidal, disc-shaped, rod-like, conical, triangular, cubical, polygonal, tubular or wire-like.

[0083] The support can be fabricated from any material, including but not limited to glass, fused-silica, silicon, a polymer (*e.g.*, polystyrene (PS), macroporous polystyrene (MPPS), polymethylmethacrylate (PMMA), polycarbonate (PC), polypropylene (PP), polyethylene (PE), high density polyethylene (HDPE), cyclic olefin polymers (COP), cyclic olefin copolymers (COC), polyethylene terephthalate (PET)), or any combination thereof. Various compositions of both glass and plastic substrates are contemplated.

[0084] In some aspects, the support can have a plurality (*e.g.*, two or more) of nucleic acid templates immobilized thereon. In some embodiments, the plurality of immobilized nucleic acid templates has the same sequence. In some embodiments, the plurality of immobilized nucleic acid templates has different sequences. In some embodiments, individual nucleic acid template molecules in the plurality of nucleic acid templates are immobilized to a different site on the support. In some embodiments, two or more individual nucleic acid template molecules in the plurality of nucleic acid templates are immobilized to a site on the support.

[0085] The term “array” refers to a support comprising a plurality of sites located at pre-determined locations on the support to form an array of sites. The sites can be discrete and separated by interstitial regions. In some embodiments, the pre-determined sites on the support can be arranged in one dimension in a row or a column, or arranged in two dimensions in rows and columns. In some embodiments, the plurality of pre-determined sites is arranged on the support in an organized fashion. In some embodiments, the plurality of pre-determined sites is arranged in any organized pattern, including rectilinear, hexagonal patterns, grid patterns, patterns having reflective symmetry, patterns having rotational symmetry, or the like. The pitch between different pairs of sites can be that same or can vary. In some embodiments, the support comprises at least 10^2 sites, at least 10^3 sites, at least 10^4 sites, at least 10^5 sites, at least 10^6 sites, at least 10^7 sites, at least 10^8 sites, at least 10^9 sites, at least 10^{10} sites, at least 10^{11} sites, at least 10^{12} sites, at least 10^{13} sites, at least 10^{14} sites, at least 10^{15} sites, or more, where the sites are located at pre-determined locations on the support. In some embodiments, a plurality of pre-determined sites on the support (*e.g.*, 10^2 – 10^{15} sites or more, *e.g.*, 10^2 sites, 10^3 sites, 10^4 sites, 10^5 sites, 10^6 sites, 10^7 sites, 10^8 sites, 10^9 sites, 10^{10} sites, 10^{11} sites, 10^{12} sites, 10^{13} sites, 10^{14} sites, 10^{15} sites, or more) are immobilized with nucleic acid templates to form a nucleic acid template array. In some embodiments, the nucleic acid templates that are immobilized at a plurality of pre-determined sites by hybridization to immobilized surface capture primers, or the nucleic acid templates are covalently attached to the surface capture primer. In some embodiments, the nucleic acid

templates that are immobilized at a plurality of pre-determined sites, for example immobilized at $10^2 - 10^{15}$ sites or more (*e.g.*, $10^2 - 10^{15}$ sites or more, *e.g.*, 10^2 sites, 10^3 sites, 10^4 sites, 10^5 sites, 10^6 sites, 10^7 sites, 10^8 sites, 10^9 sites, 10^{10} sites, 10^{11} sites, 10^{12} sites, 10^{13} sites, 10^{14} sites, 10^{15} sites, or more). In some embodiments, the immobilized nucleic acid templates are clonally-amplified to generate immobilized nucleic acid clusters at the plurality of pre-determined sites. In some embodiments, individual immobilized nucleic acid clusters comprise linear clusters, or comprise single-stranded or double-stranded concatemers.

[0086] In some embodiments, a support comprising a plurality of sites located at random locations on the support is referred to herein as a support having randomly located sites thereon. In such embodiments, the location of the randomly located sites on the support are not pre-determined. Consequently, the plurality of randomly-located sites is arranged on the support in a disordered and/or unpredictable fashion. In some embodiments, the support comprises at least 10^2 sites, at least 10^3 sites, at least 10^4 sites, at least 10^5 sites, at least 10^6 sites, at least 10^7 sites, at least 10^8 sites, at least 10^9 sites, at least 10^{10} sites, at least 10^{11} sites, at least 10^{12} sites, at least 10^{13} sites, at least 10^{14} sites, at least 10^{15} sites, or more, where the sites are randomly located on the support. In some embodiments, a plurality of randomly located sites on the support (*e.g.*, $10^2 - 10^{15}$ sites or more, *e.g.*, 10^2 sites, 10^3 sites, 10^4 sites, 10^5 sites, 10^6 sites, 10^7 sites, 10^8 sites, 10^9 sites, 10^{10} sites, 10^{11} sites, 10^{12} sites, 10^{13} sites, 10^{14} sites, 10^{15} sites, or more) are immobilized with nucleic acid templates to form a support immobilized with nucleic acid templates. In some embodiments, the nucleic acid templates that are immobilized at a plurality of randomly located sites by hybridization to immobilized surface capture primers, or the nucleic acid templates are covalently attached to the surface capture primer. In some embodiments, the nucleic acid templates that are immobilized at a plurality of randomly located sites, for example immobilized at $10^2 - 10^{15}$ sites or more (*e.g.*, $10^2 - 10^{15}$ sites or more, *e.g.*, 10^2 sites, 10^3 sites, 10^4 sites, 10^5 sites, 10^6 sites, 10^7 sites, 10^8 sites, 10^9 sites, 10^{10} sites, 10^{11} sites, 10^{12} sites, 10^{13} sites, 10^{14} sites, 10^{15} sites, or more). In some embodiments, the immobilized nucleic acid templates are clonally-amplified to generate immobilized nucleic acid clusters at the plurality of randomly located sites. In some embodiments, individual immobilized nucleic acid clusters comprise linear clusters, or comprise single-stranded or double-stranded concatemers.

[0087] In some embodiments, the plurality of immobilized surface capture primers on the support are in fluid communication with each other to permit flowing a solution of reagents (*e.g.*, nucleic acid template molecules, soluble primers, enzymes, nucleotides, divalent cations, buffers, and the like) onto the support so that the plurality of immobilized surface

capture primers on the support can be essentially simultaneously reacted with the reagents in a massively parallel manner. In some embodiments, the fluid communication of the plurality of immobilized surface capture primers can be used to conduct nucleic acid amplification reactions (*e.g.*, RCA, MDA, PCR, and bridge amplification) essentially simultaneously on the plurality of immobilized surface capture primers.

[0088] In some embodiments, the plurality of immobilized nucleic acid clusters on the support are in fluid communication with each other to permit flowing a solution of reagents (*e.g.*, enzymes, nucleotides, divalent cations, and the like) onto the support so that the plurality of immobilized nucleic acid clusters on the support can be essentially simultaneously reacted with the reagents in a massively parallel manner. In some embodiments, the fluid communication of the plurality of immobilized nucleic acid clusters can be used to conduct nucleotide binding assays and/or conduct nucleotide polymerization reactions (*e.g.*, primer extension or sequencing) essentially simultaneously on the plurality of immobilized nucleic acid clusters, and optionally to conduct detection and imaging for massively parallel sequencing.

[0089] When used in reference to immobilized enzymes, the term “immobilized” and related terms refer to enzymes (*e.g.*, polymerases) that are attached to a support through covalent bond or non-covalent interaction, or attached to a coating on the support, or buried within a matrix formed by a coating on the support.

[0090] When used in reference to immobilized nucleic acids, the term “immobilized” and related terms refer to nucleic acid molecules that are attached to a support through covalent bond or non-covalent interaction, or attached to a coating on the support, or buried within a matrix formed by a coating on the support, where the nucleic acid molecules include surface capture primers, nucleic acid template molecules and extension products of capture primers. Extension products of capture primers can include nucleic acid concatemers (*e.g.*, nucleic acid clusters).

[0091] In some embodiments, one or more nucleic acid templates are immobilized on the support, for example immobilized at the sites on the support. In some embodiments, the one or more nucleic acid templates are clonally-amplified. In some embodiments, the one or more nucleic acid templates are clonally-amplified off the support (*e.g.*, in-solution) and then deposited onto the support and immobilized on the support. In some embodiments, the clonal amplification reaction of the one or more nucleic acid templates is conducted on the support resulting in immobilization on the support. In some embodiments, the one or more nucleic acid templates are clonally-amplified (*e.g.*, in solution or on the support) using a nucleic acid

amplification reaction, including any one or any combination of: polymerase chain reaction (PCR), multiple displacement amplification (MDA), transcription-mediated amplification (TMA), nucleic acid sequence-based amplification (NASBA), strand displacement amplification (SDA), real-time SDA, bridge amplification, isothermal bridge amplification, rolling circle amplification (RCA), circle-to-circle amplification, helicase-dependent amplification, recombinase-dependent amplification, and/or single-stranded binding (SSB) protein-dependent amplification.

[0092] The term “persistence time” and related terms refers to the length of time that a binding complex, which is formed between the target nucleic acid, a polymerase, a conjugated or unconjugated nucleotide, remains stable without any binding component dissociates from the binding complex. The persistence time is indicative of the stability of the binding complex and strength of the binding interactions. Persistence time can be measured by observing the onset and/or duration of a binding complex, such as by observing a signal from a labeled component of the binding complex. For example, a labeled nucleotide or a labeled reagent comprising one or more nucleotides may be present in a binding complex, thus allowing the signal from the label to be detected during the persistence time of the binding complex. One exemplary label is a fluorescent label.

INTRODUCTION

[0093] The present disclosure provides compositions, apparatuses and methods for capturing nucleic acids from cellular samples on a support, preparing library molecules on the support, amplifying the library molecules on the support to generate nucleic acid template molecules, and analyzing the immobilized nucleic acid template molecules including detecting and/or sequencing the immobilized nucleic acid template molecules. The immobilized nucleic acid template molecules correspond to the nucleic acids from the cellular samples. The immobilized nucleic acid template molecules are spatially located on the support in an arrangement that is similar to the spatial location in the cellular sample.

Methods for Preparing Spatially Resolved Nucleic Acids

[0094] The present disclosure provides methods for preparing spatially resolved nucleic acids, comprising: **step (a)** providing a support which is passivated with at least one layer of a low non-specific binding coating. In some embodiments, the low non-specific binding coating comprises a high contrast-to-noise (CNR) coating that provides a low-binding/low scatter base layer. In some embodiments, the low non-specific binding coating comprises at

least one hydrophilic polymer layer. In some embodiments, the low non-specific binding coating has a water contact angle of no more than 45 degrees. In some embodiments, the low non-specific binding coating can form a continuous layer on the support, or the low non-specific binding coating can be placed on the support in an organized pattern across the support, for example as spots, grids and/or lines. In some embodiments, the low non-specific binding coating can provide a surface having a high contrast-to-noise (CNR) ratio. Additional descriptions of the low non-specific binding coating is provided below. In some embodiments, the support is passivated with a cell-binding coating which comprises polylysine, polyarginine, other polycations, tissue-specific antibodies, detergents, amphipathic peptides, tissue-specific or generic receptor ligands especially cell surface receptor ligands, and/or lectins. In some embodiments, the cell-binding coating can be placed on the support in an organized pattern across the support, for example as spots, grids and/or lines. In some embodiments, the cell-binding coating can be present at low density which is optimized to provide cell or tissue binding without interfering with the low binding/high CNR properties of the low non-specific binding coating(s).

[0095] In some embodiments in step (a), the support is solid, semi-solid, or a combination of both. In some embodiments, the support is porous, semi-porous, non-porous, or any combination of porosity. In some embodiments, the support can be substantially planar, concave, convex, or any combination thereof.

[0096] In some embodiments in step (a), the low non-specific binding coating comprises a plurality of surface capture primers that selectively hybridize (capture) nucleic acids from a cellular sample such as RNA and/or DNA. The surface capture primers can hybridize to any portion of the RNA or DNA from the cellular sample. In some embodiments, the surface capture primers comprise an RNA capture sequence which comprises a polyT sequence, a target-specific sequence and/or a random sequence. In some embodiments, the surface capture primers comprise a universal adaptor sequence, such as a universal surface capture primer sequence, a universal surface pinning primer sequence, a universal binding site for a reverse sequencing primer, a universal binding site for a forward sequencing primer and/or a universal binding site for a compaction oligonucleotide. In some embodiments, the polyT sequence comprises 3-30 nucleotides having a thymine base. In some embodiments, individual surface capture primers comprise (*e.g.*, arranged in a 5' to 3' orientation): a universal surface capture primer sequence (110), a universal binding site for a reverse sequencing primer (120), and an RNA capture sequence (*e.g.*, *see* **FIG. 11**). In some embodiments, the RNA capture sequence comprises a poly-T sequence and/or an RNA

target-specific sequence. In some embodiments, the 5' ends of the surface capture primers are tethered to the low non-specific binding coating. In some embodiments, the support further comprises a plurality of surface pinning primers which are tethered to the low non-specific binding coating. In some embodiments, the immobilized surface pinning primers serve to pin down at least one portion of concatemer molecules to the support (*e.g.*, concatemer molecules are described below). In some embodiments, the immobilized surface pinning primers have a non-extendible 3' end and cannot be used for amplification. In some embodiments, the surface capture primers and surface pinning primers can be covalently tethered to the low non-specific binding coating.

[0097] In some embodiments, the universal surface capture primer sequence (110) comprises the sequence: 5'- AGTCGTCGCAGCCTCACCTGATC -3' (**SEQ ID NO: 1**).

[0098] In some embodiments, the universal surface capture primer sequence (110) comprises the sequence: 5'- TCGTATGCCGTCTTCTGCTTG -3' (**SEQ ID NO: 2**).

[0099] In some embodiments, universal binding site for a reverse sequencing primer (120) comprises the sequence: 5'- ATGTCGGAAGGTGTGCAGGCTACCGCTTGTCAACT -3' (**SEQ ID NO: 3**).

[00100] In some embodiments, universal binding site for a reverse sequencing primer (120) comprises the sequence: 5'- AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -3' (**SEQ ID NO: 4**).

[00101] In some embodiments, universal binding site for a reverse sequencing primer (120) comprises the sequence: 5'- CTGTCTCTTATACACATCTCCGAGCCCACGAGAC -3' (**SEQ ID NO: 5**).

[00102] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (b)**: positioning a cellular sample on the support under conditions suitable for the cell sample to remain in a fixed position on the coated support. In some embodiments, the cellular sample can be manually positioned on the support, for example using a tool such as forceps. In some embodiments, the cellular sample can be a tissue slice which is flowed or placed onto the support. In some embodiments, the support comprises a flowcell which can be a closed flowcell or a resealable flowcell. In some embodiments, the cellular sample can be permeabilized to permit nucleic acids to elute from inside the cell to outside of the cell and contact the support. In some embodiments, RNA can elute from inside the cell and onto the surface capture primers that are immobilized on the coated support. In some embodiments, the eluted RNA can hybridize to the surface capture primers to generate a plurality of capture primer-RNA duplexes each comprising a surface capture primer

hybridized to an RNA molecule (*e.g.*, *see* **FIG. 12**). In some embodiments, RNA can elute from inside the cell and onto the support by diffusion, by thermally assisted diffusion, by radiation assisted diffusion, by electrophoresis, by centrifugation, or by other methods. In some embodiments, RNA can elute from inside the cell and onto the surface capture primers in a manner that preserves the spatial location information of the RNA molecules in the cellular sample. In some embodiments, the location of the RNA molecules hybridized to surface capture primers on the coated support corresponds to the spatial location of the RNA molecules when they were previously located inside the cellular sample. In some embodiments, the cellular sample can be permeabilized prior to, or after or concurrently, with positioning a cellular sample on the support. In some embodiments, the cellular sample can be treated with a chemical fixation reagent or the cellular sample is not chemically fixed. In some embodiments, the cellular sample can be stained, de-stained or un-stained. In some embodiments, the cellular sample can be imaged, or stained and imaged at step (b). The cellular sample can be stained by employing any of the techniques described in “Histological and Histochemical Methods: Theory and Practice”, 4th Edition, Kiernan, J.A., Ed. (2008). In some embodiments, imaging may be conducted using negative stains or by fluorescent or interferometric methods. In some embodiments, in step (b), the cellular sample (*e.g.*, tissue slice) can be positioned on the support for automated or semi-automated scanning. In some embodiments, the cellular sample can be removed from the support after eluting the nucleic acids from the cellular sample and onto the immobilized surface capture primers on the support of step (b). In some embodiments, the cellular sample can be removed from the support after any of steps (c) – (i) which are described below. In some embodiments, the cellular sample can be imaged, or stained and imaged during any of steps (c) – (i).

[00103] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (c)**: conducting a reverse transcription reaction on the support by contacting the capture primer-RNA duplexes with Moloney Murine Leukemia Virus (MMLV) reverse transcriptase, a plurality of nucleotides (*e.g.*, dATP, dGTP, dCTP, dTTP and/or dUTP). In some embodiments, the reverse transcription reaction generates a plurality of first strand cDNA molecules by extending the 3' end of the immobilized surface capture primers and using the hybridized RNA as a template strand. In some embodiments, the reverse transcription reaction can be conducted under a condition suitable for the MMLV reverse transcription enzyme to generate a non-template polyC tail at the 3' end of the first strand cDNA (*e.g.*, *see* **FIG. 13**). Alternatively, a non-template polyC tail can be appended to the 3' end of the first strand cDNA using terminal deoxynucleotidyl transferase and dCTP.

[00104] In some embodiments, the reverse transcription reaction of step (c) can be conducted in the presence of a plurality of template switching oligonucleotides. Individual template switching oligonucleotides can hybridize to the non-template polyC tail at the 3' end of the first strand of the cDNA (*e.g.*, see **FIG. 14**). In some embodiments, the template switching oligonucleotides comprise a DNA/RNA chimeric oligonucleotide. In some embodiments, the template switching oligonucleotides comprise a polyG region which can hybridize with the polyC tail at the 3' end of the first strand of the cDNA. In some embodiments, the polyG region of the template switching oligonucleotide comprises ribonucleotide bases (*e.g.*, see **FIG. 14**). In some embodiments, the template switching oligonucleotide further comprises at least one universal adaptor sequence, for example a universal binding site for a surface pinning primer (140), and a universal binding site for a forward sequencing primer (130) (*e.g.*, see **FIG. 14**).

[00105] In some embodiments, the universal binding site for a forward sequencing primer (130) comprises the sequence: 5'- CGTGCTGGATTGGCTCACCAGACACCTTCCGACAT -3' (**SEQ ID NO: 6**).

[00106] In some embodiments, the universal binding site for a forward sequencing primer (130) comprises the sequence: 5'- ACACTCTTTCCTACACGACGCTCTTCCGATCT -3' (**SEQ ID NO: 7**).

[00107] In some embodiments, the universal binding site for a forward sequencing primer (130) comprises the sequence: 5'- TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG -3' (**SEQ ID NO: 8**).

[00108] In some embodiments, the universal binding site for a surface pinning primer (140) comprises the sequence: 5'- CATGTAATGCACGTACTTTCAGGGT -3' (**SEQ ID NO: 9**).

[00109] In some embodiments, the universal binding site for a surface pinning primer (140) comprises the sequence: 5'- AATGATACGGCGACCACCGA -3' (**SEQ ID NO: 10**).

[00110] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (d)**: extending the first strand of the cDNA using MMLV reverse transcriptase, a plurality of nucleotides (*e.g.*, dATP, dGTP, dCTP, dTTP and/or dUTP), and at least a portion of the template switching oligonucleotide (TSO) as a template strand (see the upper dashed line with an arrow in **FIG. 15** labelled '1st strand cDNA (TSO)'). In some embodiments, the extension reaction of step (d) extends from the 3' terminal end of the polyC tail of the first strand cDNA. The extension reaction of step (d) can generate a full length first strand of cDNA, which comprises (in a 5' to 3' orientation): a universal surface capture

primer sequence (110); a universal binding site for a reverse sequencing primer (120); an RNA capture sequence; a cDNA insert corresponding to the RNA; a polyC tail; a universal binding site for forward sequencing primer (130); and a universal binding site for surface pinning primer (140).

[00111] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (e)**: removing the RNA while retaining the immobilized full length first strand cDNA. In some embodiments, the RNA can be removed using RNase. In some embodiments, the retained immobilized full length first strand cDNA comprises (in a 5' to 3' orientation): a universal surface capture primer sequence (110); a universal binding site for a reverse sequencing primer (120); an RNA capture sequence; a cDNA insert corresponding to the RNA; a polyC tail; a universal binding site for forward sequencing primer (130); and a universal binding site for surface pinning primer (140) (*e.g.*, see **FIG. 16**).

[00112] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (f)**: contacting the retained immobilized full length first strand cDNA with a plurality of single stranded circularization oligonucleotides. In some embodiments, individual circularization oligonucleotides comprise: an adaptor sequence that is complementary to the universal binding site for a reverse sequencing primer (120); an adaptor sequence that is complementary to the universal surface capture primer sequence (110); a linker region; an adaptor sequence that is complementary to the universal binding site for surface pinning primer (140); and an adaptor sequence that is complementary to the universal binding site for forward sequencing primer (130) (*e.g.*, see **FIG. 17**). In some embodiments, the circularization oligonucleotide lacks an adaptor sequence that is complementary to the universal surface capture primer sequence (110). The retained immobilized full length first strand cDNA are contacted with a plurality of single stranded circularization oligonucleotides under a condition suitable for the circularization oligonucleotide to hybridize to the universal adaptor sequences of the immobilized first strand cDNA but not hybridize to the RNA capture sequence, the cDNA insert region or the polyC tail. Hybridization of the circularization oligonucleotide to the immobilized full length first strand cDNA forms a single stranded circular molecule with a gap. In some embodiments, the linker region can include any one or any combination of two or more of: at least one sample index sequence for multiplexing; at least one unique molecular index (UMI) sequence for molecular tagging; and/or a universal binding site for a compaction oligonucleotide. In some embodiments, the sample index comprises 5-20 bases which can be used to distinguish polynucleotides (*e.g.*, insert sequences) from different sample sources in a

multiplex assay. In some embodiments, the unique molecular index (UMI) comprises 3-20 bases which can be used to uniquely identify individual nucleic acid molecules to which the UMI sequence is appended (*e.g.*, molecular tagging). In some embodiments, the unique molecular index (UMI) comprises a random sequence.

[00113] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (g)**: conducting a polymerase-catalyzed extension reaction to fill in the gap, using as a template strand the polyC tail region, the first strand cDNA insert region, and the RNA capture sequence region. The polymerase-catalyzed extension reaction forms a single stranded circularized molecule with a nick. In some embodiments, step (g) further comprises conducting an enzymatic ligation reaction to generate a single stranded, covalently closed circular molecule, which is hybridized to the immobilized full length first strand cDNA (*e.g.*, see **FIG. 18**).

[00114] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (h)**: conducting a rolling circle amplification reaction using the terminal 3' end of the immobilized full length 1st strand cDNA as an initiation site to generate nucleic acid concatemers immobilized to the support. In some embodiments, the rolling circle amplification reaction employs a DNA polymerase having a strand displacing activity and a plurality of nucleotides (*e.g.*, dATP, dGTP, dCTP, dTTP and/or dUTP). The rolling circle amplification reaction can use the covalently closed circular molecule as a template strand to generate concatemers having a plurality of tandem repeat units, where each unit comprises: a linker region (or a complementary sequence thereof); a universal surface capture primer sequence (110) (or a complementary sequence thereof); a universal binding site for a reverse sequencing primer (120) (or a complementary sequence thereof); an RNA capture sequence (or a complementary sequence thereof); a first strand cDNA insert region (or a complementary sequence thereof); a polyC region (or a complementary sequence thereof); a universal binding site for forward sequencing primer (130) (or a complementary sequence thereof); and a universal binding site for surface pinning primer (140) (or a complementary sequence thereof) (*e.g.*, see **FIG. 19**).

[00115] In some embodiments, the rolling circle amplification reaction of step (h) can be conducted in the presence or absence of a plurality of compaction oligonucleotides. In some embodiments, the compaction oligonucleotides comprise single stranded oligonucleotides having a first region at one end that hybridizes to a portion of a concatemer molecule and a second region at the other end that hybridizes to another portion of the same concatemer molecule, where hybridization of the compaction oligonucleotide to a given concatemer

compacts the size and/or shape of the concatemer. In some embodiments, the 5' and 3' regions of the compaction oligonucleotide can hybridize to different portions of the same concatemer to pull together distal portions of the concatemer causing compaction of the concatemer to form a DNA nanoball. Additional descriptions of compaction oligonucleotides can be found below.

[00116] In some embodiments, the DNA polymerase having a strand displacing activity of step (h) can be selected from a group consisting of phi29 DNA polymerase, large fragment of Bst DNA polymerase, large fragment of Bsu DNA polymerase, and Bca (exo-) DNA polymerase, Klenow fragment of E. coli DNA polymerase, T5 polymerase, M-MuLV reverse transcriptase, HIV viral reverse transcriptase, or Deep Vent DNA polymerase. In some embodiments, the phi29 DNA polymerase can be wild type phi29 DNA polymerase (*e.g.*, MagniPhi from Expedeon), or variant EquiPhi29 DNA polymerase (*e.g.*, from Thermo Fisher Scientific), and chimeric QualiPhi DNA polymerase (*e.g.*, from 4basebio).

[00117] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (i)**: sequencing the plurality of immobilized concatemers. In some embodiments, the sequencing of step (i) comprises sequencing at least a portion of individual immobilized concatemers to identify the individual captured RNA. In some embodiments, the location of individual sequenced concatemers on the coated support corresponds to the spatial location of the individual RNAs from the cellular sample. In some embodiments, the sequencing of step (i) comprises sequencing at least a portion of individual immobilized concatemers including sequencing at least a portion of the cDNA insert portion of the concatemer that corresponds to the captured RNA. In some embodiments, sequencing the immobilized concatemers on the low non-specific binding coating of step (a) can provide a surface having a high contrast-to-noise (CNR) ratio during sequencing. The high contrast-to-noise (CNR) ratio can provide improved sensitivity and specificity for accurate spatial location of the concatemers on the support which corresponds to accurate location of the RNA transcripts of interest from the cellular sample. In some embodiments, the sequencing of step (i) comprises use of fluorescently labeled nucleotide reagents and sequencing polymerases. In some embodiments, the sequencing of step (i) comprises imaging to detect fluorescent signals emitted from the immobilized concatemers during a sequencing reaction which employs fluorescently labeled nucleotide reagents and sequencing polymerases.

[00118] In some embodiments, the plurality of immobilized concatemers can be sequenced using any nucleic acid sequencing method that employs labeled or non-labeled chain terminating nucleotides, where the chain terminating nucleotides include a 3'-O-azido group

(or 3'-O-methylazido group) or any other type of bulky blocking group at the sugar 3' position. In some embodiments, the concatemer template molecules can be sequenced using a two-stage sequencing method using labeled multivalent molecules and non-labeled chain terminating nucleotides. In some embodiments, the concatemer template molecules can be sequenced using a sequencing-by-synthesis (SBS) method which employs labeled chain-terminating nucleotides. In some embodiments, the concatemer template molecules can be sequenced using a sequencing-by-binding methods (SBB) which employ non-labeled chain-terminating nucleotides. In some embodiments, the concatemer template molecules can be sequenced using phosphate-chain labeled nucleotides. These various sequencing methods are described below.

[00119] In some embodiments, the immobilized concatemers can be detected by contacting the immobilized concatemers with a plurality of oligonucleotide probes that are labeled with a detectable reporter moiety under a condition that is suitable for selective hybridization of the probes to the immobilized concatemers. In some embodiments, the contacting step can further include an imaging step. In some embodiments, the oligonucleotide probes comprise a target-specific sequence such as a sequence that is complementary to any portion of the cDNA insert region. In some embodiments, the oligonucleotide probes comprise a sequence that can hybridize with a universal adaptor sequence including any of the universal binding site for a surface capture primer, a universal binding site for a surface pinning primer, a universal binding sequence for a forward sequencing primer, a universal binding sequence for a reverse sequencing primer, a sample barcode sequence and/or a unique molecular index sequence.

[00120] In some embodiments, the methods for preparing spatially resolved nucleic acids further comprise **step (j)**: determining the location of individual concatemers on the coated support which corresponds to the spatial location of the individual RNAs eluted from the cellular sample. In some embodiments, the sequencing and imaging of step (i) can be used to determine the location of individual concatemers immobilized to the coated support which corresponds to the spatial location of the individual RNAs eluted from the cellular sample.

Cellular Sample

[00121] In any of the methods described herein, the cellular sample refers to a single cell, a whole cell, a plurality of cells, a tissue, an organ, an organism, or section of any of these cellular samples. The cellular sample can be extracted (*e.g.*, biopsied) from an organism, or obtained from a cell culture grown in liquid or in a culture dish. The cellular sample

comprises a sample that is fresh, frozen, fresh frozen, or archived (*e.g.*, formalin-fixed paraffin-embedded; FFPE). The cellular sample can be embedded in a wax, paraffin, resin, epoxy or agar. The cellular sample can be fixed, for example in any one or any combination of two or more of acetone, ethanol, methanol, formaldehyde, paraformaldehyde-Triton or glutaraldehyde. The cellular sample can be sectioned or non-sectioned. The cellular sample can be stained, de-stained or non-stained.

[00122] In some embodiments, the cellular sample can be obtained from a virus, fungus, prokaryote or eukaryote. In some embodiments, the cellular sample can be obtained from an animal, insect or plant. In some embodiments, the cellular sample comprises one or more virally-infected cells.

[00123] In some embodiments, the cellular sample can be obtained from any organism including human, simian, ape, canine, feline, bovine, equine, murine, porcine, caprine, lupine, ranine, piscine, plant, insect or bacteria.

[00124] In some embodiments, the cellular sample can be obtained from any organ including head, neck, brain, breast, ovary, cervix, colon, rectum, endometrium, gallbladder, intestines, bladder, prostate, testicles, liver, lung, kidney, esophagus, pancreas, thyroid, pituitary, thymus, skin, heart, larynx, or other organs.

[00125] In any of the methods described herein, the cellular sample harbors a plurality of RNA which can be eluted onto the plurality of surface capture primers immobilized to the coated support. In some embodiments, the eluted RNA include target RNA and/or non-target RNA. In some embodiments, the eluted RNA comprises wild type RNA, mutant RNA and/or splice variant RNA. In some embodiments, the eluted RNA comprises pre-spliced RNA, partially spliced RNA, and/or fully spliced RNA. In some embodiments, the eluted RNA comprises coding RNA, non-coding RNA, mRNA, tRNA, alternative rRNA, rRNA, microRNA (miRNA), mature microRNA, immature microRNA, lncRNA, ncRNA and/or intron RNA. In some embodiments, the eluted RNA comprises housekeeping RNA, cell-specific RNA, tissue-specific RNA or disease-specific RNA. In some embodiments, the eluted RNA comprises RNA expressed by one or more cells in response to a stimulus such as heat, light, a chemical or a drug. In some embodiments, the eluted RNA comprises RNA found in healthy cells or diseased cells. In some embodiments, the eluted RNA comprises RNA transcribed from transgenic DNA sequences that are introduced into the cellular sample using recombinant DNA procedures. For example, the RNA can be transcribed from a transgenic DNA sequence that is controlled by an inducible or constitutive promoter

sequence. In some embodiments, the eluted RNA comprises RNA that is transcribed from DNA sequences that are not transgenic.

Permeabilization

[00126] In any of the methods described herein, the cellular sample can be permeabilized to permit the nucleic acids within the sample, including the target nucleic acid molecule (*e.g.*, RNA and/or DNA), to migrate from inside the cell(s) to the plurality of capture oligonucleotides that are immobilized on a support. The cellular sample can be contacted with one or more permeabilizing agents, including organic solvents, detergents, chemical compounds, cross-linking agents and/or enzymes. In some embodiments, the organic solvents comprise acetone, ethanol, and methanol. In some embodiments, the detergents comprise saponin, Triton X-100, Tween-20, or sodium dodecyl sulfate (SDS), or N-lauroylsarcosine sodium salt solution. In some embodiments, the cross-linking agent comprises paraformaldehyde. In some embodiments, the enzyme comprises trypsin, pepsin or protease (*e.g.* proteinase K). In some embodiments, the cellular sample can be permeabilized using an alkaline condition, or an acidic condition with a protease enzyme. In some embodiments, the target nucleic acid molecule from the cellular sample is hybridized (captured) on the capture oligonucleotides immobilized on the support (or the support coating) in a manner that preserves spatial location information of the target nucleic acid molecule in the cellular sample.

[00127] In any of the methods described herein, the cellular sample comprises a permeabilized cellular sample. In some embodiments, the methods comprise treating the cellular sample with a permeabilization reagent that alters the cell membrane to permit penetration of experimental reagents into the cells. For example, the permeabilization reagent removes membrane lipids from the cell membrane. In some embodiments, the cellular sample can be treated with a permeabilization reagent which comprises any combination of an organic solvent, detergent, chemical compound, cross-linking agent and/or enzyme. In some embodiments, the organic solvents comprise acetone, ethanol, and methanol. In some embodiments, the detergents comprise saponin, Triton X-100, Tween-20, sodium dodecyl sulfate (SDS), an N-lauroylsarcosine sodium salt solution, or a nonionic polyoxyethylene surfactant (*e.g.*, NP40). In some embodiments, the cross-linking agent comprises paraformaldehyde. In some embodiments, the enzyme comprises trypsin, pepsin or protease (*e.g.* proteinase K). In some embodiments, the cells can be permeabilized using an alkaline

condition, or an acidic condition with a protease enzyme. In some embodiments, the permeabilization reagent comprises water and/or PBS.

[00128] For example, the fixed cells can be permeabilized with 70% ethanol for about 30-60 minutes, and the permeabilizing reagent can be exchanged with PBS-T (*e.g.*, PBS with 0.05% Tween-20). In some embodiments, the cells can be post-fixed with 3% paraformaldehyde and 0.1% glutaraldehyde for about 30-60 minutes, and washed with PBS-T multiple times.

Cell Fixation

[00129] In any of the methods described herein, the cellular sample comprises a fixed cellular sample. In some embodiments, the cellular sample can be treated with a fixation reagent (*e.g.*, a fixing reagent) that preserves the cell and its contents to inhibit degradation and can inhibit cell lysis. For example, the fixation reagent can preserve RNA harbored by the cellular sample. In some embodiments, the fixation reagent inhibits loss of nucleic acids from the cellular sample.

[00130] In some embodiments, the fixation reagent can cross-link the RNA to prevent the RNA from escaping the cellular sample. In some embodiments, a cross-linking fixation reagent comprises any combination of an aldehyde, formaldehyde, paraformaldehyde, formalin, glutaraldehyde, imidoesters, N-hydroxysuccinimide esters (NHS) and/or glyoxal (a bifunctional aldehyde).

[00131] In some embodiments, the fixation reagent comprises at least one alcohol, including methanol or ethanol. In some embodiments, the fixation reagent comprises at least one ketone, including acetone. In some embodiments, the fixation reagent comprises acetic acid, glacial acetic acid and/or picric acid. In some embodiments, the fixation reagent comprises mercuric chloride. In some embodiments, the fixation reagent comprises a zinc salt comprising zinc sulphate or zinc chloride. In some embodiments, the fixation reagent can denature polypeptides.

[00132] In some embodiments, the fixation reagent comprises 4% w/v of paraformaldehyde to water/PBS. In some embodiments, the fixation reagent comprises 10% of 35% formaldehyde at a neutral pH. In some embodiments, the fixation reagent comprises 2% v/v of glutaraldehyde to water/PBS. In some embodiments, the fixation reagent comprises 25% of 37% formaldehyde solution, 70% picric acid and 5% acetic acid.

[00133] In some embodiments, the cellular sample can be fixed on the support with 4% paraformaldehyde for about 30-60 minutes and washed with PBS.

Compaction Oligonucleotides

[00134] In some embodiments, the rolling circle amplification reaction can be conducted in the presence of a plurality of compaction oligonucleotides which, when hybridized to a concatemer molecule, compacts the size and/or shape of the concatemer to form a compact nanoball. In some embodiments, the compaction oligonucleotides comprise single stranded oligonucleotides having a first region at one end that hybridizes to a portion of a concatemer molecule and a second region at the other end that hybridizes to another portion of the same concatemer molecule, where hybridization of the compaction oligonucleotide to a given concatemer compacts the size and/or shape of the concatemer.

[00135] In some embodiments, the compaction oligonucleotides include a 5' region, an optional internal region (intervening region), and a 3' region. The 5' and 3' regions of the compaction oligonucleotide can hybridize to different portions of the same concatemer to pull together distal portions of the concatemer causing compaction of the concatemer to form a DNA nanoball. For example, and without limitation, the 5' region of the compaction oligonucleotide is designed to hybridize to a first portion of the concatemer molecule (*e.g.*, a universal compaction oligonucleotide binding site), and the 3' region of the compaction oligonucleotide is designed to hybridize to a second portion of the concatemer molecule (*e.g.*, a universal compaction oligonucleotide binding site). Inclusion of compaction oligonucleotides during RCA can promote formation of DNA nanoballs having tighter size and shape compared to concatemers generated in the absence of the compaction oligonucleotides. The compact and stable characteristics of the DNA nanoballs improves sequencing accuracy, for example, by increasing signal intensity and the nanoballs retain their shape and size during multiple sequencing cycles.

[00136] In some embodiments, the compaction oligonucleotides comprise single stranded oligonucleotides comprising DNA, RNA, or a combination of DNA and RNA. The compaction oligonucleotides can be any length, including 20-150 nucleotides, *e.g.*, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 nucleotides, or any range therebetween. In some embodiments, the compaction oligonucleotides can be 30-100 nucleotides in length. In some embodiments, the compaction oligonucleotides can be 40-80 nucleotides in length. In some embodiments, the compaction oligonucleotides can be 80-1200 nucleotides in length.

[00137] In some embodiments, the compaction oligonucleotides comprises a 5' region and a 3' region, and optionally an intervening region between the 5' and 3' regions. The

intervening region can be any length, for example about 2-20 nucleotides in length, , *e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20, or any range therebetween.. The intervening region comprises a homopolymer having consecutive identical bases (*e.g.*, AAA, GGG, CCC, TTT, or UUU). The intervening region comprises a non-homopolymer sequence.

[00138] In some embodiments, the 5' region of the compaction oligonucleotides can be wholly complementary or partially complementary along its length to a first portion of a concatemer molecule. In some embodiments, the 3' region of the compaction oligonucleotides can be wholly complementary or partially complementary along its length to a second portion of a concatemer molecule. In some embodiments, the 5' region of the compaction oligonucleotides can hybridize to a first universal sequence portion of a concatemer molecule. In some embodiments, the 3' region of the compaction oligonucleotides can hybridize to a second universal sequence portion of a concatemer molecule.

[00139] In some embodiments, the 5' region of the compaction oligonucleotide can have the same sequence as the 3' region. The 5' region of the compaction oligonucleotide can have a sequence that is different from the 3' region. In some embodiments, the 3' region of the compaction oligonucleotide can have a sequence that is a reverse sequence of the 5' region. In some embodiments, the 5' region of the compaction oligonucleotide can have a sequence that is a reverse sequence of the 3' region.

[00140] In some embodiments, the 3' region of any of the compaction oligonucleotides can include an additional three bases at the terminal 3' end which comprises 2'-O-methyl RNA bases (*e.g.*, designated mUmUmU) or the terminal 3' end lacks additional 2'-O-methyl RNA bases.

[00141] In some embodiments, the compaction oligonucleotides comprise one or more modified bases or linkages at their 5' or 3' ends to confer certain functionalities. In some embodiments, the compaction oligonucleotides comprise at least one phosphorothioate linkages at their 5' and/or 3' ends to confer exonuclease resistance. In some embodiments, at least one nucleotide at or near the 3' end comprises a 2' fluoro base which confers exonuclease resistance. In some embodiments, the 3' end of the compaction oligonucleotides is non-extendible. In some embodiments, the 3' end of the compaction oligonucleotides comprise at least one 2'-O-methyl RNA base which blocks polymerase-catalyzed extension. For example, the 3' end of the compaction oligonucleotide comprises three bases comprising 2'-O-methyl RNA base (*e.g.*, designated mUmUmU). In some embodiments, the compaction oligonucleotides comprise a 3' inverted dT at their 3' ends which blocks polymerase-

catalyzed extension. In some embodiments, the compaction oligonucleotides comprise 3' phosphorylation which blocks polymerase-catalyzed extension. In some embodiments, the internal region of the compaction oligonucleotides comprise at least one locked nucleic acid (LNA) which increases the thermal stability of duplexes formed by hybridizing a compaction oligonucleotide to a concatemer molecule. In some embodiments, the compaction oligonucleotides comprise a phosphorylated 5' end (*e.g.*, using a polynucleotide kinase).

[00142] In some embodiments, the compaction oligonucleotides includes an additional three bases at the terminal 3' end which comprises 2'-O-methyl RNA bases (*e.g.*, designated mUmUmU) or the terminal 3' end lacks additional 2'-O-methyl RNA bases.

[00143] In some embodiments, the compaction oligonucleotides can include at least one region having consecutive guanines. For example, the compaction oligonucleotides can include at least one region having 2, 3, 4, 5, 6 or more consecutive guanines. In some embodiments, the compaction oligonucleotides comprise four consecutive guanines which can form a guanine tetrad structure (see **FIG. 20**). The guanine tetrad structure can be stabilized via Hoogsteen hydrogen bonding. The guanine tetrad structure can be stabilized by a central cation including potassium, sodium, lithium, rubidium or cesium.

[00144] At least one compaction oligonucleotide can form a guanine tetrad (**FIG. 20**) and hybridize to the universal binding sequences in a concatemer which can cause the concatemer to fold to form an intramolecular G-quadruplex structure (**FIG. 21**). The concatemers can self-collapse to form compact nanoballs. Formation of the guanine tetrads and G-quadruplexes in the nanoballs may increase the stability of the nanoballs to retain their compact size and shape which can withstand changes in pH, temperature and/or repeated flows of reagents during sequencing inside the cellular sample.

[00145] In some embodiments, the rolling circle amplification reaction comprises a plurality of compaction oligonucleotides having the same sequence. Alternatively, the rolling circle amplification reaction comprises a plurality of compaction oligonucleotides having a mixture of two or more different sequences.

[00146] In some embodiments, the immobilized concatemer template molecule can self-collapse into a compact nucleic acid nanoball. The nanoballs can be imaged and a FWHM measurement can be obtained to give the shape/size of the nanoballs.

[00147] In some embodiments, inclusion of compaction oligonucleotides in the rolling circle amplification reaction can promote collapsing of a concatemer into a DNA nanoball. Conducting RCA with compaction oligonucleotides helps retain the compact size and shape of a DNA nanoball during multiple sequencing cycles which can improve FWHM (full width

half maximum) of a spot image of the DNA nanoball inside a cellular sample. In some embodiments, the DNA nanoball does not unravel during multiple sequencing cycles. In some embodiments, the spot image of the DNA nanoball does not enlarge during multiple sequencing cycles. In some embodiments, the spot image of the DNA nanoball remains a discrete spot during multiple sequencing cycles. The spot image can be represented as a Gaussian spot and the size can be measured as a FWHM. A smaller spot size as indicated by a smaller FWHM typically correlates with an improved image of the spot. In some embodiments, the FWHM of a nanoball spot can be about 10 μm or smaller.

Methods for Sequencing

[00148] The present disclosure provides methods for sequencing a plurality of concatemer template molecules which are immobilized to a support, for example any of the immobilized concatemer molecules described herein. In some embodiments, the sequencing reactions employ a plurality of sequencing primers, a plurality of sequencing polymerases, and nucleotide reagents comprising any one or any combination of nucleotides and/or multivalent molecules. In some embodiments, the nucleotide reagents comprise canonical nucleotides. In some embodiments, the nucleotide reagents comprise nucleotide analogs comprise detectably labeled nucleotides. In some embodiments, the nucleotide reagents comprise nucleotides carrying a removable or non-removable chain terminating moiety. In some embodiments, the nucleotide reagents comprise multivalent molecules each comprising a central core attached to multiple polymer arms each having a nucleotide moiety at the end of the arms. In some embodiments, the sequencing reactions employ binding non-labeled nucleotides without incorporation. In some embodiments, the sequencing reactions employ incorporating non-labeled nucleotide analogs. In some embodiments, the sequencing reactions employ incorporating detectably labeled nucleotides having removable chain terminating moiety. In some embodiments, the sequencing reactions employ a two-stage sequencing reaction comprising binding detectably labeled multivalent molecules without incorporation, and incorporating nucleotide analogs. In some embodiments, the sequencing reactions employ phosphate chain-labeled nucleotides.

Methods for Sequencing using Nucleotide Analogs

[00149] In some aspects, the present disclosure provides methods for sequencing, comprising **step (a)**: contacting a sequencing polymerase to (i) a nucleic acid concatemer molecule and (ii) a nucleic acid primer (*e.g.*, a forward or reverse sequencing primer),

wherein the contacting is conducted under a condition suitable to bind the sequencing polymerase to the nucleic acid concatemer molecule which is hybridized to the nucleic acid primer, wherein the nucleic acid concatemer molecule hybridized to the nucleic acid primer forms the nucleic acid duplex. In some embodiments, the sequencing polymerase comprises a recombinant mutant sequencing polymerase. In some embodiments, the primer comprises a 3' extendible end.

[00150] In some embodiments, the methods for sequencing further comprise **step (b):** contacting the sequencing polymerase with a plurality of nucleotides under a condition suitable for binding at least one nucleotide to the sequencing polymerase which is bound to the nucleic acid duplex and suitable for polymerase-catalyzed nucleotide incorporation. In some embodiments, the sequencing polymerase is contacted with the plurality of nucleotides in the presence of at least one catalytic cation comprising magnesium and/or manganese. In some embodiments, the plurality of nucleotides comprises at least one nucleotide analog having a chain terminating moiety at the sugar 2' or 3' position. In some embodiments, the plurality of nucleotides comprises at least one nucleotide that lacks a chain terminating moiety.

[00151] In some embodiments, the methods for sequencing further comprise **step (c):** incorporating at least one nucleotide into the 3' end of the extendible primer under a condition suitable for incorporating the at least one nucleotide. In some embodiments, the suitable conditions for nucleotide binding the polymerase and for incorporation the nucleotide can be the same or different. In some embodiments, conditions suitable for incorporating the nucleotide comprise inclusion of at least one catalytic cation comprising magnesium and/or manganese. In some embodiments, the at least one nucleotide binds the sequencing polymerase and incorporates into the 3' end of the extendible primer. In some embodiments, the incorporating the nucleotide into the 3' end of the primer in **step (c)** comprises a primer extension reaction.

[00152] In some embodiments, the methods for sequencing further comprise **step (d):** repeating the incorporating at least one nucleotide into the 3' end of the extendible primer of **step (c)** at least once. In some embodiments, the plurality of nucleotides comprises a plurality of nucleotides labeled with detectable reporter moiety. The detectable reporter moiety comprises a fluorophore. In some embodiments, the fluorophore is attached to the nucleotide base. In some embodiments, the fluorophore is attached to the nucleotide base with a linker which is cleavable/removable from the base. In some embodiments, at least one of the nucleotides in the plurality is not labeled with a detectable reporter moiety. In some

embodiments, a particular detectable reporter moiety (*e.g.*, fluorophore) that is attached to the nucleotide can correspond to the nucleotide base (*e.g.*, dATP, dGTP, dCTP, dTTP or dUTP) to permit detection and identification of the nucleotide base. In some embodiments, the method further comprises detecting the at least one incorporated nucleotide at step (c) and/or (d). In some embodiments, the method further comprises identifying the at least one incorporated nucleotide at step (c) and/or (d). In some embodiments, the sequence of the nucleic acid concatemer molecule can be determined by detecting and identifying the nucleotide that binds the sequencing polymerase, thereby determining the sequence of the concatemer molecule. In some embodiments, the sequence of the nucleic acid concatemer molecule can be determined by detecting and identifying the nucleotide that incorporates into the 3' end of the primer, thereby determining the sequence of the concatemer molecule.

[00153] In some embodiments, in the methods for sequencing, the plurality of sequencing polymerases that are bound to the nucleic acid duplexes comprise a plurality of complexed polymerases, having at least a first and second complexed polymerase, wherein (a) the first complexed polymerases comprises a first sequencing polymerase bound to a first nucleic acid duplex comprising a first nucleic acid template sequence which is hybridized to a first nucleic acid primer, (b) the second complexed polymerases comprises a second sequencing polymerase bound to a second nucleic acid duplex comprising a second nucleic acid template sequence which is hybridized to a second nucleic acid primer, (c) the first and second nucleic acid template sequences comprise the same or different sequences, (d) the first and second nucleic acid concatemers are clonally-amplified, (e) the first and second primers comprise extendible 3' ends or non-extendible 3' ends, and (f) the plurality of complexed polymerases are immobilized to a support. In some embodiments, the density of the plurality of complexed polymerases is about $10^2 - 10^{15}$ (*e.g.*, $10^2 - 10^{15}$ or more, *e.g.*, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15}) complexed polymerases per mm^2 that are immobilized to the support.

Two-Stage Sequencing Methods for Nucleic Acids

[00154] In some aspects, the present disclosure provides a two-stage method for sequencing nucleic acid molecules. In some embodiments, the first stage generally comprises binding multivalent molecules to complexed polymerases to form multivalent-complexed polymerases, and detecting the multivalent-complexed polymerases.

[00155] In some embodiments, the first stage comprises **step (a)**: contacting a plurality of a first sequencing polymerase to (i) a plurality of nucleic acid concatemer molecules and (ii) a plurality of nucleic acid primers (*e.g.*, forward or reverse sequencing primers), wherein the contacting is conducted under a condition suitable to bind the plurality of first sequencing polymerases to the plurality of nucleic acid concatemer molecules and the plurality of nucleic acid primers thereby forming a plurality of first complexed polymerases each comprising a first sequencing polymerase bound to a nucleic acid duplex wherein the nucleic acid duplex comprises a nucleic acid concatemer molecule hybridized to a nucleic acid primer. In some embodiments, the first polymerase comprises a recombinant mutant sequencing polymerase.

[00156] In some embodiments, in the methods for sequencing concatemer molecules, the primer comprises a 3' extendible end or a 3' non-extendible end. In some embodiments, the plurality of nucleic acid concatemer molecules comprise amplified template molecules (*e.g.*, clonally amplified template molecules). In some embodiments, the plurality of nucleic acid concatemer molecules comprise one copy of a target sequence of interest. In some embodiments, the plurality of nucleic acid molecules comprise two or more tandem copies of a target sequence of interest (*e.g.*, concatemers). In some embodiments, the nucleic acid concatemer molecules in the plurality of nucleic acid concatemer molecules comprise the same target sequence of interest or different target sequences of interest. In some embodiments, the plurality of nucleic acid concatemer molecules and/or the plurality of nucleic acid primers are in solution or are immobilized to a support. In some embodiments, when the plurality of nucleic acid concatemer molecules and/or the plurality of nucleic acid primers are immobilized to a support, the binding with the first sequencing polymerase generates a plurality of immobilized first complexed polymerases. In some embodiments, the plurality of nucleic acid concatemer molecules and/or nucleic acid primers are immobilized to $10^2 - 10^{15}$ different sites on a support (*e.g.*, $10^2 - 10^{15}$ sites or more, *e.g.*, 10^2 sites, 10^3 sites, 10^4 sites, 10^5 sites, 10^6 sites, 10^7 sites, 10^8 sites, 10^9 sites, 10^{10} sites, 10^{11} sites, 10^{12} sites, 10^{13} sites, 10^{14} sites, 10^{15} sites). In some embodiments, the binding of the plurality of concatemer molecules and nucleic acid primers with the plurality of first sequencing polymerases generates a plurality of first complexed polymerases immobilized to $10^2 - 10^{15}$ different sites on the support (*e.g.*, $10^2 - 10^{15}$ sites or more, *e.g.*, 10^2 sites, 10^3 sites, 10^4 sites, 10^5 sites, 10^6 sites, 10^7 sites, 10^8 sites, 10^9 sites, 10^{10} sites, 10^{11} sites, 10^{12} sites, 10^{13} sites, 10^{14} sites, 10^{15} sites). In some embodiments, the plurality of immobilized first complexed polymerases on the support are immobilized to pre-determined or to random sites on the support. In some embodiments, the plurality of immobilized first complexed polymerases are

in fluid communication with each other to permit flowing a solution of reagents (*e.g.*, enzymes including sequencing polymerases, multivalent molecules, nucleotides, and/or divalent cations) onto the support so that the plurality of immobilized complexed polymerases on the support are reacted with the solution of reagents in a massively parallel manner.

[00157] In some embodiments, the methods for sequencing further comprise **step (b)**: contacting the plurality of first complexed polymerases with a plurality of multivalent molecules to form a plurality of multivalent-complexed polymerases (*e.g.*, binding complexes). In some embodiments, individual multivalent molecules in the plurality of multivalent molecules comprise a core attached to multiple nucleotide arms and each nucleotide arm is attached to a nucleotide (*e.g.*, nucleotide unit) (*e.g.*, FIGS. 2-6). In some embodiments, the contacting of step (b) is conducted under a condition suitable for binding complementary nucleotide units of the multivalent molecules to at least two of the plurality of first complexed polymerases thereby forming a plurality of multivalent-complexed polymerases. In some embodiments, the condition is suitable for inhibiting polymerase-catalyzed incorporation of the complementary nucleotide units into the primers of the plurality of multivalent-complexed polymerases. In some embodiments, the plurality of multivalent molecules comprises at least one multivalent molecule having multiple nucleotide arms (*e.g.*, FIGS. 2-6) each attached with a nucleotide analog (*e.g.*, nucleotide analog unit), where the nucleotide analog includes a chain terminating moiety at the sugar 2' and/or 3' position. In some embodiments, the plurality of multivalent molecules comprises at least one multivalent molecule comprising multiple nucleotide arms each attached with a nucleotide unit that lacks a chain terminating moiety. In some embodiments, at least one of the multivalent molecules in the plurality of multivalent molecules is labeled with a detectable reporter moiety. Any portion of the multivalent molecule can be labeled including the core, nucleotide arm or nucleobase. In some embodiments, the detectable reporter moiety comprises a fluorophore. In some embodiments, the contacting of step (b) is conducted in the presence of at least one non-catalytic cation comprising strontium, barium and/or calcium.

[00158] In some embodiments, the methods for sequencing further comprise **step (c)**: detecting the plurality of multivalent-complexed polymerases. In some embodiments, the detecting includes detecting the multivalent molecules that are bound to the complexed polymerases, where the complementary nucleotide units of the multivalent molecules are bound to the primers but incorporation of the complementary nucleotide units is inhibited. In some embodiments, the multivalent molecules are labeled with a detectable reporter moiety

to permit detection. In some embodiments, the labeled multivalent molecules comprise a fluorophore attached to the core, linker and/or nucleotide unit of the multivalent molecules.

[00159] In some embodiments, the methods for sequencing further comprise **step (d)**: identifying the base of the complementary nucleotide units that are bound to the plurality of first complexed polymerases, thereby determining the sequence of the concatemer molecule. In some embodiments, the multivalent molecules are labeled with a detectable reporter moiety that corresponds to the particular nucleotide units attached to the nucleotide arms to permit identification of the complementary nucleotide units (*e.g.*, nucleotide base adenine, guanine, cytosine, thymine or uracil) that are bound to the plurality of first complexed polymerases.

[00160] In some embodiments, the second stage of the two-stage sequencing method generally comprises nucleotide incorporation. In some embodiments, the methods for sequencing further comprise **step (e)**: dissociating the plurality of multivalent-complexed polymerases and removing the plurality of first sequencing polymerases and their bound multivalent molecules, and retaining the plurality of nucleic acid duplexes.

[00161] In some embodiments, the methods for sequencing further comprises **step (f)**: contacting the plurality of the retained nucleic acid duplexes of step (e) with a plurality of second sequencing polymerases, wherein the contacting is conducted under a condition suitable for binding the plurality of second sequencing polymerases to the plurality of the retained nucleic acid duplexes, thereby forming a plurality of second complexed polymerases each comprising a second sequencing polymerase bound to a nucleic acid duplex. In some embodiments, the second sequencing polymerase comprises a recombinant mutant sequencing polymerase.

[00162] In some embodiments, the plurality of first sequencing polymerases of step (a) have an amino acid sequence that is 100% identical to the amino acid sequence as the plurality of the second sequencing polymerases of step (f). In some embodiments, the plurality of first sequencing polymerases of step (a) have an amino acid sequence that differs from the amino acid sequence of the plurality of the second sequencing polymerases of step (f).

[00163] In some embodiments, the methods for sequencing further comprise **step (g)**: contacting the plurality of second complexed polymerases with a plurality of nucleotides, wherein the contacting is conducted under a condition suitable for binding complementary nucleotides from the plurality of nucleotides to at least two of the second complexed polymerases thereby forming a plurality of nucleotide-complexed polymerases. In some

embodiments, the contacting of step (g) is conducted under a condition that is suitable for promoting polymerase-catalyzed incorporation of the bound complementary nucleotides into the primers of the nucleotide-complexed polymerases thereby forming a plurality of nucleotide-complexed polymerases. In some embodiments, the incorporating the nucleotide into the 3' end of the primer in step (g) comprises a primer extension reaction. In some embodiments, the contacting of step (g) is conducted in the presence of at least one catalytic cation comprising magnesium and/or manganese. In some embodiments, the contacting of step (g) is conducted in the presence of magnesium and/or manganese. In some embodiments, the plurality of nucleotides comprises native nucleotides (*e.g.*, non-analog nucleotides) or nucleotide analogs. In some embodiments, the plurality of nucleotides comprise a 2' and/or 3' chain terminating moiety which is removable or is not removable. In some embodiments, the plurality of nucleotides comprises a plurality of nucleotides labeled with detectable reporter moiety. The detectable reporter moiety may comprise a fluorophore. In some embodiments, the fluorophore is attached to the nucleotide base. In some embodiments, the fluorophore is attached to the nucleotide base with a linker which is cleavable/removable from the base or is not removable from the base. In some embodiments, a particular detectable reporter moiety (*e.g.*, fluorophore) that is attached to the nucleotide can correspond to the nucleotide base (*e.g.*, dATP, dGTP, dCTP, dTTP or dUTP) to permit detection and identification of the nucleotide base. In some embodiments, at least one of the nucleotides in the plurality is not labeled with a detectable reporter moiety. In some embodiments, the plurality of nucleotides is not labeled with a detectable reporter moiety.

[00164] In some embodiments, the methods for sequencing further comprise **step (h)**: when the nucleotides are labeled with a detectable reporter moiety, step (h) comprises detecting the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases. In some embodiments, the plurality of nucleotides is labeled with a detectable reporter moiety to permit detection. In some embodiments, in the methods for sequencing concatemer molecules, when the nucleotides are non-labeled then the detecting step is omitted.

[00165] In some embodiments, the methods for sequencing further comprise **step (i)**: when the nucleotides are labeled with a detectable reporter moiety, step (i) comprises identifying the bases of the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases. In some embodiments, the identification of the incorporated complementary nucleotides in step (i) can be used to confirm the identity of the complementary nucleotides of the multivalent molecules that are bound to the plurality

of first complexed polymerases in step (d). In some embodiments, the identifying of step (i) can be used to determine the sequence of the nucleic acid concatemer molecules. In some embodiments, in the methods for sequencing concatemer molecules, when the nucleotides are non-labeled then the identifying step is omitted.

[00166] In some embodiments, the methods for sequencing further comprise **step (j)**: removing the chain terminating moiety from the incorporated nucleotide when step (g) is conducted by contacting the plurality of second complexed polymerases with a plurality of nucleotides that comprise at least one nucleotide having a 2' and/or 3' chain terminating moiety.

[00167] In some embodiments, the methods for sequencing further comprise **step (k)**: repeating steps (a) – (j) at least once. In some embodiments, the sequence of the nucleic acid concatemer molecules can be determined by detecting and identifying the multivalent molecules that bind the sequencing polymerases but do not incorporate into the 3' end of the primer at steps (c) and (d). In some embodiments, the sequence of the nucleic acid concatemer molecule can be determined (or confirmed) by detecting and identifying the nucleotide that incorporates into the 3' end of the primer at steps (h) and (i).

[00168] In some embodiments, in any of the methods for sequencing nucleic acid molecules, the binding of the plurality of first complexed polymerases with the plurality of multivalent molecules forms at least one avidity complex, the method comprising the steps: **(a)** binding a first nucleic acid primer, a first sequencing polymerase, and a first multivalent molecule to a first portion of a concatemer template molecule thereby forming a first binding complex, wherein a first nucleotide unit of the first multivalent molecule binds to the first sequencing polymerase; and **(b)** binding a second nucleic acid primer, a second sequencing polymerase, and the first multivalent molecule to a second portion of the same concatemer template molecule thereby forming a second binding complex, wherein a second nucleotide unit of the first multivalent molecule binds to the second sequencing polymerase, wherein the first and second binding complexes which include the same multivalent molecule forms an avidity complex. In some embodiments, the first sequencing polymerase comprises any wild type or mutant polymerase described herein. In some embodiments, the second sequencing polymerase comprises any wild type or mutant polymerase described herein. The concatemer template molecule comprises tandem repeat sequences of a sequence of interest and at least one universal sequencing primer binding site. The first and second nucleic acid primers can bind to a sequencing primer binding site along the concatemer template molecule. Exemplary multivalent molecules are shown in **FIGS. 2-6**.

[00169] In some embodiments, in any of the methods for sequencing nucleic acid molecules, wherein the method includes binding the plurality of first complexed polymerases with the plurality of multivalent molecules to form at least one avidity complex, the method comprising the steps: **(a)** contacting the plurality of sequencing polymerases and the plurality of nucleic acid primers with different portions of a concatemer nucleic acid concatemer molecule to form at least first and second complexed polymerases on the same concatemer template molecule; **(b)** contacting a plurality of multivalent molecules to the at least first and second complexed polymerases on the same concatemer template molecule, under conditions suitable to bind a single multivalent molecule from the plurality to the first and second complexed polymerases, wherein at least a first nucleotide unit of the single multivalent molecule is bound to the first complexed polymerase which includes a first primer hybridized to a first portion of the concatemer template molecule thereby forming a first binding complex (*e.g.*, first ternary complex), and wherein at least a second nucleotide unit of the single multivalent molecule is bound to the second complexed polymerase which includes a second primer hybridized to a second portion of the concatemer template molecule thereby forming a second binding complex (*e.g.*, second ternary complex), wherein the contacting is conducted under a condition suitable to inhibit polymerase-catalyzed incorporation of the bound first and second nucleotide units in the first and second binding complexes, and wherein the first and second binding complexes which are bound to the same multivalent molecule forms an avidity complex; and **(c)** detecting the first and second binding complexes on the same concatemer template molecule, and **(d)** identifying the first nucleotide unit in the first binding complex thereby determining the sequence of the first portion of the concatemer template molecule, and identifying the second nucleotide unit in the second binding complex thereby determining the sequence of the second portion of the concatemer template molecule. In some embodiments, the plurality of sequencing polymerases comprise any wild type or mutant sequencing polymerase described herein. The concatemer template molecule comprises tandem repeat sequences of a sequence of interest and at least one universal sequencing primer binding site. The plurality of nucleic acid primers can bind to a sequencing primer binding site along the concatemer template molecule. Exemplary multivalent molecules are shown in **FIGS. 2-6**.

[00170] In some embodiments, the two-stage sequencing can employ multivalent molecules that are labeled with fluorophores and the detecting and/or identifying steps comprise use of fluorescence imaging. In some embodiments, the fluorescence imaging comprises dual wavelength excitation/four wavelength emission fluorescence imaging. In

some embodiments, four different types of multivalent molecules are employed, each comprising a different nucleotide unit (or nucleotide unit analog). For example, a first type of multivalent molecules comprise dATP nucleotide units, a second type of multivalent molecules comprise dGTP nucleotide units, a third type of multivalent molecules comprise dCTP nucleotide units, and a fourth type of multivalent molecules comprise dTTP nucleotide units. In some embodiments, the four different types of multivalent molecules are labeled with a different fluorophore that corresponds to the nucleotide units attached to a given multivalent molecule to permit identification of the nucleotide units. In some embodiments, the detecting step comprises simultaneous or single excitation at a wavelength sufficient to excite all four fluorophores and imaging of fluorescence emission at wavelengths sufficient to detect each respective fluorophore. In some embodiments, the four labeled multivalent molecules are used to determine the identity of a terminal nucleotide in the nucleic acid template molecule. In some embodiments, the four types of multivalent molecules are labeled with different fluorophores, including for example fluorophores that emit different visible colors such as blue, green, yellow, orange or red. In some embodiments, the four types of multivalent molecules are labeled with different fluorophores, including for example Cy2 or a dye or fluorophore similar in excitation or emission properties, Cy3 or a dye or fluorophore similar in excitation or emission properties, Cy3.5 or a dye or fluorophore similar in excitation or emission properties, Cy5 or a dye or fluorophore similar in excitation or emission properties, Cy5.5 or a dye or fluorophore similar in excitation or emission properties, and Cy7 or a dye or fluorophore similar in excitation or emission properties. In some embodiments, the detecting step comprises simultaneous excitation at any two of 532 nm, 568 nm and 633 nm, and imaging of fluorescence emission at about 570 nm, 592 nm, 670 nm, and 702 nm, respectively. In some embodiments, the fluorescence imaging comprises dual wavelength excitation/dual wavelength emission fluorescence imaging. In some embodiments, the four different types of multivalent molecules are labeled with distinguishable fluorophores (or a set of fluorophores), and the detecting step comprises simultaneous or single excitation at a wavelength sufficient to excite one, two, three, or four fluorophores or sets of fluorophores, and imaging of fluorescence emission at wavelengths is sufficient to detect each respective fluorophore.

[00171] In some embodiments, the two-stage sequencing methods can be conducted with three different types of labeled multivalent molecules and one type of unlabeled multivalent molecule (*e.g.*, a “dark” multivalent molecule), where the labeled multivalent molecules are labeled with a different fluorophore that corresponds to the nucleotide units attached to a

given multivalent molecule to permit identification of the nucleotide units. In some embodiments, the detecting step comprises simultaneous excitation at a wavelength sufficient to excite the three types of fluorophores and imaging of fluorescence emission at wavelengths is sufficient to detect each respective fluorophore, and detection of the fourth type of multivalent molecule is determined or determinable with reference to the location of “dark” or unlabeled spots.

[00172] In some embodiments, the fluorophores comprise a FRET donor and acceptor pair, such that multiple detections and identifications can be performed under a single excitation and imaging step. In some embodiments, a sequencing cycle comprises forming a plurality of complexed polymerases, contacting the complexed polymerases with a plurality of different types of fluorescently-labeled multivalent molecules, and detecting the fluorescently-labeled multivalent molecules that are bound to the complexed polymerases. In some embodiments, a sequencing cycle can be conducted in less than 30 minutes, in less than 20 minutes, or in less than 10 minutes. In some embodiments, conducting sequencing reactions with labeled multivalent molecules gives an average Q-score for base calling accuracy over a sequencing run which is greater than or equal to 30, and/or greater than or equal to 40. In some embodiments, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the base calls have a Q-score of greater than 30 and/or greater than or equal to 40. In some embodiments, the present disclosure provides the method, herein at least 95% of the base calls have a Q-score of greater than 30.

Sequencing-by-Binding Methods

[00173] In some aspects, the present disclosure provides methods for sequencing any of the immobilized concatemer molecules described herein, wherein the sequencing methods comprise a sequencing-by-binding (SBB) procedure which employs non-labeled chain-terminating nucleotides. In some embodiments, the sequencing-by-binding (SBB) method comprises the steps of **(a)** sequentially contacting a primed template nucleic acid with at least two separate mixtures under ternary complex stabilizing conditions, wherein the at least two separate mixtures each include a polymerase and a nucleotide, whereby the sequentially contacting results in the primed template nucleic acid being contacted, under the ternary complex stabilizing conditions, with nucleotide cognates for first, second and third base type base types in the template; **(b)** examining the at least two separate mixtures to determine whether a ternary complex formed; and **(c)** identifying the next correct nucleotide for the primed template nucleic acid molecule, wherein the next correct nucleotide is identified as a

cognate of the first, second or third base type if ternary complex is detected in step (b), and wherein the next correct nucleotide is imputed to be a nucleotide cognate of a fourth base type based on the absence of a ternary complex in step (b); **(d)** adding a next correct nucleotide to the primer of the primed template nucleic acid after step (b), thereby producing an extended primer; and **(e)** repeating steps (a) through (d) at least once on the primed template nucleic acid that comprises the extended primer. Exemplary sequencing-by-binding methods are described in U.S. patent Nos. 10,246,744 and 10,731,141 (where the contents of both patents are hereby incorporated by reference in their entireties).

Methods for Sequencing using Phosphate-Chain Labeled Nucleotides

[00174] The present disclosure provides methods for sequencing any of the immobilized concatemer molecules described herein, wherein the sequencing methods comprise sequencing with phosphate-chain labeled nucleotides, comprising **step (a):** contacting (i) a plurality of sequencing polymerases, (ii) a plurality of concatemer template molecules immobilized to a support and (iii) a plurality of nucleic acid sequencing primers (*e.g.*, forward or reverse sequencing primers), where the contacting is conducted under a condition suitable to form a plurality of complexed sequencing polymerases each complex comprising a sequencing polymerase bound to a nucleic acid duplex wherein the nucleic acid duplex comprises a portion of a concatemer template molecule hybridized to a nucleic acid sequencing primer. In some embodiments, the sequencing polymerases comprise a recombinant mutant sequencing polymerase that can bind and incorporate nucleotide analogs. In some embodiments, the sequencing polymerases are immobilized to the same support as the concatemers. In some embodiments, the sequencing polymerases are not immobilized to the a support. In some embodiments, the sequencing primers comprise 3' extendible ends or 3' blocked end that can be converted into a 3' extendible end.

[00175] In some embodiments, the methods for sequencing concatemer template molecules further comprise **step (b):** contacting the plurality of complexed sequencing polymerases with a plurality of phosphate chain-labeled nucleotides under a condition suitable for binding at least one phosphate chain-labeled nucleotide to one of the complexed sequencing polymerases, and the condition is suitable for promoting polymerase-catalyzed nucleotide incorporation. In some embodiments, the complexed sequencing polymerase is contacted with the plurality of nucleotides in the presence of at least one catalytic cation comprising magnesium and/or manganese. In some embodiments, individual phosphate chain-labeled nucleotides in the plurality comprise an aromatic base, a five carbon sugar

(*e.g.*, ribose or deoxyribose), and phosphate chain comprising 3-20 (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 phosphate groups) phosphate groups, where the terminal phosphate group is linked to a detectable reporter moiety (*e.g.*, a fluorophore). The first, second and third phosphate groups can be referred to as alpha, beta, and gamma phosphate groups. In some embodiments, a particular detectable reporter moiety which is attached to the terminal phosphate group corresponds to the nucleotide base (*e.g.*, dATP, dGTP, dCTP, dTTP or dUTP) to permit detection and identification of the nucleotide base. In some embodiments, the sequencing polymerases are capable of binding a complementary phosphate chain labeled nucleotide and incorporating the complementary nucleotide opposite a nucleotide in a template molecule. In some embodiments, the polymerase-catalyzed nucleotide incorporation reaction cleaves between the alpha and beta phosphate groups thereby releasing a multi-phosphate chain linked to the detectable reporter moiety. In some embodiments, the plurality of phosphate chain-labeled nucleotides comprises one type or a mixture of any two or more types of nucleotides comprising dATP, dGTP, dCTP, dTTP and/or dUTP.

[00176] In some embodiments, the sequencing method further comprises **step (c)**: detecting the fluorescent signal emitted by the phosphate chain labeled nucleotide that is bound by the sequencing polymerase, and incorporated into the terminal end of the sequencing primer. In some embodiments, step (c) further comprises identifying the phosphate chain labeled nucleotide that is bound by the sequencing polymerase, and incorporated into the terminal end of the sequencing primer.

[00177] In some embodiments, the sequencing method further comprises **step (d)**: repeating steps (b) – (c) at least once. In some embodiments, sequencing methods that employ phosphate chain labeled nucleotides can be conducted according to the methods described in U.S. patent Nos. 7,170,050; 7,302,146; and/or 7,405,281.

[00178] In some embodiments, in step (a), the plurality of concatemer template molecules are immobilized to a support which comprises a plurality of separate compartments. In some embodiments, the plurality of sequencing polymerases are in solution in a compartment. In some embodiments, at least one sequencing polymerase is immobilized to the bottom of individual compartments. In some embodiments, the separate compartments comprise a silica bottom through which light can penetrate. In some embodiments, the separate compartments comprise a silica bottom configured with a nanophotonic confinement structure comprising a hole in a metal cladding film (*e.g.*, aluminum cladding film). In some embodiments, the hole in the metal cladding has a small aperture, for example, approximately 70 nm. In some

embodiments, the height of the nanophotonic confinement structure is approximately 100 nm. In some embodiments, the nanophotonic confinement structure comprises a zero mode waveguide (ZMW). In some embodiments, the nanophotonic confinement structure contains a liquid.

[00179] In some embodiments, the covalently closed circular library molecule (600) can serve as a non-immobilized template molecule. In some embodiments, the sequencing methods comprise **step (a)**: providing a support having a plurality of sequencing polymerases immobilized thereon. In some embodiments, the sequencing polymerase comprises a processive DNA polymerase. In some embodiments, the sequencing polymerase comprises a wild type or mutant DNA polymerase, including for example a Phi29 DNA polymerase. In some embodiments, the support comprise a plurality of separate compartments and a sequencing polymerase is immobilized to the bottom of a compartment. In some embodiments, the separate compartments comprise a silica bottom through which light can penetrate. In some embodiments, the separate compartments comprise a silica bottom configured with a nanophotonic confinement structure comprising a hole in a metal cladding film (*e.g.*, aluminum cladding film). In some embodiments, the hole in the metal cladding has a small aperture, for example, approximately 70 nm. In some embodiments, the height of the nanophotonic confinement structure is approximately 100 nm. In some embodiments, the nanophotonic confinement structure comprises a zero mode waveguide (ZMW). In some embodiments, the nanophotonic confinement structure contains a liquid.

[00180] In some embodiments, the sequencing method further comprises **step (b)**: contacting the plurality of immobilized sequencing polymerases with a plurality of single stranded circular nucleic acid template molecules (*e.g.*, covalently closed circular library molecules (600)) and a plurality of oligonucleotide sequencing primers, under a condition suitable for individual immobilized sequencing polymerases to bind a single stranded circular template molecule, and suitable for individual sequencing primers to hybridize to individual single stranded circular template molecules, thereby generating a plurality of polymerase/template/primer complexes. In some embodiments, the individual sequencing primers hybridize to a universal sequencing primer binding site on the single stranded circular template molecule.

[00181] In some embodiments, the sequencing method further comprises **step (c)**: contacting the plurality of polymerase/template/primer complexes with a plurality of phosphate chain labeled nucleotides each comprising an aromatic base, a five carbon sugar (*e.g.*, ribose or deoxyribose), and phosphate chain comprising 3-20 phosphate groups, where

the terminal phosphate group is linked to a detectable reporter moiety (*e.g.*, a fluorophore). The first, second and third phosphate groups can be referred to as alpha, beta and gamma phosphate groups. In some embodiments, a particular detectable reporter moiety which is attached to the terminal phosphate group corresponds to the nucleotide base (*e.g.*, dATP, dGTP, dCTP, dTTP or dUTP) to permit detection and identification of the nucleotide base. In some embodiments, the plurality of polymerase/template/primer complexes are contacted with the plurality of phosphate chain labeled nucleotides under a condition suitable for polymerase-catalyzed nucleotide incorporation. In some embodiments, the sequencing polymerases are capable of binding a complementary phosphate chain labeled nucleotide and incorporating the complementary nucleotide opposite a nucleotide in a template molecule. In some embodiments, the polymerase-catalyzed nucleotide incorporation reaction cleaves between the alpha and beta phosphate groups thereby releasing a multi-phosphate chain linked to a fluorophore.

[00182] In some embodiments, the sequencing method further comprises **step (d)**: detecting the fluorescent signal emitted by the phosphate chain labeled nucleotide that is bound by the sequencing polymerase, and incorporated into the terminal end of the sequencing primer. In some embodiments, step (d) further comprises identifying the phosphate chain labeled nucleotide that is bound by the sequencing polymerase, and incorporated into the terminal end of the sequencing primer.

[00183] In some embodiments, the sequencing method further comprises **step (d)**: repeating steps (c) – (d) at least once. In some embodiments, sequencing methods that employ phosphate chain labeled nucleotides can be conducted according to the methods described in U.S. patent Nos. 7,170,050; 7,302,146; and/or 7,405,281.

Sequencing Polymerases

[00184] In any of the methods described herein, sequencing polymerases can be used for conducting sequencing reactions. In some embodiments, the sequencing polymerase(s) is/are capable of binding and incorporating a complementary nucleotide opposite a nucleotide in a concatemer template molecule. In some embodiments, the sequencing polymerase(s) is/are capable of binding a complementary nucleotide unit of a multivalent molecule opposite a nucleotide in a concatemer template molecule. In some embodiments, the plurality of sequencing polymerases comprise recombinant mutant polymerases.

[00185] Examples of suitable polymerases for use in sequencing with nucleotides and/or multivalent molecules include but are not limited to: Klenow DNA polymerase; *Thermus aquaticus* DNA polymerase I (Taq polymerase); KlenTaq polymerase; *Candidatus altiarthaeales* archaeon; *Candidatus Hadarchaeum Yellowstone*; *Hadesarchaea* archaeon; *Euryarchaeota* archaeon; *Thermoplasma* archaeon; *Thermococcus* polymerases such as *Thermococcus litoralis*, bacteriophage T7 DNA polymerase; human alpha, delta and epsilon DNA polymerases; bacteriophage polymerases such as T4, RB69 and phi29 bacteriophage DNA polymerases; *Pyrococcus furiosus* DNA polymerase (Pfu polymerase); *Bacillus subtilis* DNA polymerase III; *E. coli* DNA polymerase III alpha and epsilon; 9 degree N polymerase; reverse transcriptases such as HIV type M or O reverse transcriptases; avian myeloblastosis virus reverse transcriptase; Moloney Murine Leukemia Virus (MMLV) reverse transcriptase; or telomerase. Further non-limiting examples of DNA polymerases include those from various Archaea genera, such as, *Aeropyrum*, *Archaeoglobus*, *Desulfurococcus*, *Pyrobaculum*, *Pyrococcus*, *Pyrolobus*, *Pyrodictum*, *Staphylothermus*, *Stetteria*, *Sulfolobus*, *Thermococcus*, and *Vulcanisaeta* and the like or variants thereof, including such polymerases as are known in the art such as 9 degrees N, VENT, DEEP VENT, THERMINATOR, Pfu, KOD, Pfx, Tgo and RB69 polymerases.

Nucleotides and Chain-Terminating Nucleotides

[00186] In any of the methods described herein, any of the sequencing methods described herein can employ at least one nucleotide. The nucleotides comprise a base, sugar and at least one phosphate group. In some embodiments, at least one nucleotide in the plurality comprises an aromatic base, a five carbon sugar (*e.g.*, ribose or deoxyribose), and one or more phosphate groups (*e.g.*, 1-10 phosphate groups). The plurality of nucleotides can comprise at least one type of nucleotide selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP. The plurality of nucleotides can comprise a mixture of any combination of two or more types of nucleotides selected from the group consisting of dATP, dGTP, dCTP, dTTP and/or dUTP. In some embodiments, at least one nucleotide in the plurality is not a nucleotide analog. In some embodiments, at least one nucleotide in the plurality comprises a nucleotide analog.

[00187] In some embodiments, in any of the methods for sequencing described herein, at least one nucleotide in the plurality of nucleotides comprise a chain of one, two or three phosphorus atoms where the chain is typically attached to the 5' carbon of the sugar moiety

via an ester or phosphoramidate linkage. In some embodiments, at least one nucleotide in the plurality is an analog having a phosphorus chain in which the phosphorus atoms are linked together with intervening O, S, NH, methylene or ethylene. In some embodiments, the phosphorus atoms in the chain include substituted side groups including O, S or BH₃. In some embodiments, the chain includes phosphate groups substituted with analogs including phosphoramidate, phosphorothioate, phosphordithioate, and O-methylphosphoroamidite groups.

[00188] In some embodiments, in any of the methods for sequencing described herein, at least one nucleotide in the plurality of nucleotides comprises a terminator nucleotide analog having a chain terminating moiety (*e.g.*, blocking moiety) at the sugar 2' position, at the sugar 3' position, or at the sugar 2' and 3' position. In some embodiments, the chain terminating moiety can inhibit polymerase-catalyzed incorporation of a subsequent nucleotide unit or free nucleotide in a nascent strand during a primer extension reaction. In some embodiments, the chain terminating moiety is attached to the 3' sugar hydroxyl position where the sugar comprises a ribose or deoxyribose sugar moiety. In some embodiments, the chain terminating moiety is removable/cleavable from the 3' sugar hydroxyl position to generate a nucleotide having a 3'OH sugar group which is extendible with a subsequent nucleotide in a polymerase-catalyzed nucleotide incorporation reaction. In some embodiments, the chain terminating moiety comprises an alkyl group, alkenyl group, alkynyl group, allyl group, aryl group, benzyl group, azide group, amine group, amide group, keto group, isocyanate group, phosphate group, thio group, disulfide group, carbonate group, urea group, acetal group or silyl group. In some embodiments, the chain terminating moiety is cleavable/removable from the nucleotide, for example by reacting the chain terminating moiety with a chemical agent, pH change, light or heat. In some embodiments, the chain terminating moieties alkyl, alkenyl, alkynyl and allyl are cleavable with tetrakis(triphenylphosphine)palladium(0) (Pd(PPh₃)₄) with piperidine, or with 2,3-Dichloro-5,6-dicyano-1,4-benzo-quinone (DDQ). In some embodiments, the chain terminating moieties aryl and benzyl are cleavable with H₂ Pd/C. In some embodiments, the chain terminating moieties amine, amide, keto, isocyanate, phosphate, thio, disulfide are cleavable with phosphine or with a thiol group including beta-mercaptoethanol or dithiothritol (DTT). In some embodiments, the chain terminating moiety carbonate is cleavable with potassium carbonate (K₂CO₃) in MeOH, with triethylamine in pyridine, or with Zn in acetic acid (AcOH). In some embodiments, the chain terminating moieties urea and silyl are cleavable

with tetrabutylammonium fluoride, pyridine-HF, with ammonium fluoride, or with triethylamine trihydrofluoride.

[00189] In some embodiments, in any of the methods for sequencing described herein, at least one nucleotide in the plurality of nucleotides comprises a terminator nucleotide analog having a chain terminating moiety (*e.g.*, blocking moiety) at the sugar 2' position, at the sugar 3' position, or at the sugar 2' and 3' position. In some embodiments, the chain terminating moiety comprises an azide, azido or azidomethyl group. In some embodiments, the chain terminating moiety comprises a 3'-O-azido or 3'-O-azidomethyl group. In some embodiments, the chain terminating moieties azide, azido and azidomethyl group are cleavable/removable with a phosphine compound. In some embodiments, the phosphine compound comprises a derivatized tri-alkyl phosphine moiety or a derivatized tri-aryl phosphine moiety. In some embodiments, the phosphine compound comprises Tris(2-carboxyethyl)phosphine (TCEP) or bis-sulfo triphenyl phosphine (BS-TPP) or Tri(hydroxypropyl)phosphine (THPP). In some embodiments, the cleaving agent comprises 4-dimethylaminopyridine (4-DMAP).

[00190] In some embodiments, in any of the methods for sequencing described herein, the nucleotide comprises a chain terminating moiety which is selected from a group consisting of 3'-deoxy nucleotides, 2',3'-dideoxynucleotides, 3'-methyl, 3'-azido, 3'-azidomethyl, 3'-O-azidoalkyl, 3'-O-ethynyl, 3'-O-aminoalkyl, 3'-O-fluoroalkyl, 3'-fluoromethyl, 3'-difluoromethyl, 3'-trifluoromethyl, 3'-sulfonyl, 3'-malonyl, 3'-amino, 3'-O-amino, 3'-sulfhydryl, 3'-aminomethyl, 3'-ethyl, 3'-butyl, 3'-*tert* butyl, 3'-Fluorenylmethyloxycarbonyl, 3'-*tert*-Butyloxycarbonyl, 3'-O-alkyl hydroxylamino group, 3'-phosphorothioate, and 3-O-benzyl, or derivatives thereof.

[00191] In some embodiments, in any of the methods for sequencing described herein, the plurality of nucleotides comprises a plurality of nucleotides labeled with detectable reporter moiety. The detectable reporter moiety comprises a fluorophore. In some embodiments, the fluorophore is attached to the nucleotide base. In some embodiments, the fluorophore is attached to the nucleotide base with a linker which is cleavable/removable from the base. In some embodiments, at least one of the nucleotides in the plurality is not labeled with a detectable reporter moiety. In some embodiments, a particular detectable reporter moiety (*e.g.*, fluorophore) that is attached to the nucleotide can correspond to the nucleotide base (*e.g.*, dATP, dGTP, dCTP, dTTP or dUTP) to permit detection and identification of the nucleotide base.

[00192] In some embodiments, in any of the methods for sequencing nucleic acid molecules described herein, the cleavable linker on the nucleotide base comprises a cleavable moiety comprising an alkyl group, alkenyl group, alkynyl group, allyl group, aryl group, benzyl group, azide group, amine group, amide group, keto group, isocyanate group, phosphate group, thio group, disulfide group, carbonate group, urea group, acetal group or silyl group. In some embodiments, the cleavable linker on the base is cleavable/removable from the base by reacting the cleavable moiety with a chemical agent, pH change, light or heat. In some embodiments, the cleavable moieties alkyl, alkenyl, alkynyl and allyl are cleavable with tetrakis(triphenylphosphine)palladium(0) ($\text{Pd(PPh}_3)_4$) with piperidine, or with 2,3-Dichloro-5,6-dicyano-1,4-benzo-quinone (DDQ). In some embodiments, the cleavable moieties aryl and benzyl are cleavable with H_2 Pd/C. In some embodiments, the cleavable moieties amine, amide, keto, isocyanate, phosphate, thio, disulfide are cleavable with phosphine or with a thiol group including beta-mercaptoethanol or dithiothritol (DTT). In some embodiments, the cleavable moiety carbonate is cleavable with potassium carbonate (K_2CO_3) in MeOH, with triethylamine in pyridine, or with Zn in acetic acid (AcOH). In some embodiments, the cleavable moieties urea and silyl are cleavable with tetrabutylammonium fluoride, pyridine-HF, with ammonium fluoride, or with triethylamine trihydrofluoride.

[00193] In some embodiments, in any of the methods for sequencing described herein, the cleavable linker on the nucleotide base comprises cleavable moiety including an azide, azido or azidomethyl group. In some embodiments, the cleavable moieties azide, azido and azidomethyl group are cleavable/removable with a phosphine compound. In some embodiments, the phosphine compound comprises a derivatized tri-alkyl phosphine moiety or a derivatized tri-aryl phosphine moiety. In some embodiments, the phosphine compound comprises Tris(2-carboxyethyl)phosphine (TCEP) or bis-sulfo triphenyl phosphine (BS-TPP) or Tri(hydroxypropyl)phosphine (THPP). In some embodiments, the cleaving agent comprises 4-dimethylaminopyridine (4-DMAP).

[00194] In some embodiments, in any of the methods for sequencing described herein, the chain terminating moiety (*e.g.*, at the sugar 2' and/or sugar 3' position) and the cleavable linker on the nucleotide base have the same or different cleavable moieties. In some embodiments, the chain terminating moiety (*e.g.*, at the sugar 2' and/or sugar 3' position) and the detectable reporter moiety linked to the base are chemically cleavable/removable with the same chemical agent. In some embodiments, the chain terminating moiety (*e.g.*, at the sugar 2' and/or sugar 3' position) and the detectable reporter moiety linked to the base are chemically cleavable/removable with different chemical agents.

Multivalent Molecules

[00195] In any of the methods described herein, the sequencing employs at least one multivalent molecule which comprises a plurality of nucleotide arms attached to a core and having any configuration including a starburst, helter skelter, or bottle brush configuration (*e.g.*, **FIG. 2**). In some embodiments, the multivalent molecule comprises: (1) a core; and (2) a plurality of nucleotide arms which comprise (i) a core attachment moiety, (ii) a spacer comprising a PEG moiety, (iii) a linker, and (iv) a nucleotide unit, wherein the core is attached to the plurality of nucleotide arms, wherein the spacer is attached to the linker, wherein the linker is attached to the nucleotide unit. In some embodiments, the nucleotide unit comprises a base, sugar and at least one phosphate group, and the linker is attached to the nucleotide unit through the base. In some embodiments, the linker comprises an aliphatic chain or an oligo ethylene glycol chain where both linker chains having 2-6 subunits. In some embodiments, the linker also includes an aromatic moiety. An exemplary nucleotide arm is shown in **FIG. 6**. Exemplary multivalent molecules are shown in **FIGS. 2-5**. An exemplary spacer is shown in **FIG. 7** (top) and exemplary linkers are shown in **FIG. 7** (bottom) and **FIG. 8**. Exemplary nucleotides attached to a linker are shown in **FIGS. 9A-9D**. An exemplary biotinylated nucleotide arm is shown in **FIG. 10**.

[00196] In some embodiments, a multivalent molecule comprises a core attached to multiple nucleotide arms, and wherein the multiple nucleotide arms have the same type of nucleotide unit which is selected from a group consisting of dATP, dGTP, dCTP, dTTP and dUTP.

[00197] In some embodiments, a multivalent molecule comprises a core attached to multiple nucleotide arms, where each arm includes a nucleotide unit. The nucleotide unit comprises an aromatic base, a five carbon sugar (*e.g.*, ribose or deoxyribose), and one or more phosphate groups (*e.g.*, 1-10 phosphate groups). The plurality of multivalent molecules can comprise one type multivalent molecule having one type of nucleotide unit selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP. The plurality of multivalent molecules can comprise a mixture of any combination of two or more types of multivalent molecules, where individual multivalent molecules in the mixture comprise nucleotide units selected from the group consisting of dATP, dGTP, dCTP, dTTP and/or dUTP.

[00198] In some embodiments, the nucleotide unit comprises a chain of one, two or three phosphorus atoms where the chain is typically attached to the 5' carbon of the sugar moiety via an ester or phosphoramidate linkage. In some embodiments, at least one nucleotide unit is

a nucleotide analog having a phosphorus chain in which the phosphorus atoms are linked together with intervening O, S, NH, methylene or ethylene. In some embodiments, the phosphorus atoms in the chain include substituted side groups including O, S or BH₃. In some embodiments, the chain includes phosphate groups substituted with analogs including phosphoramidate, phosphorothioate, phosphordithioate, and O-methylphosphoroamidite groups.

[00199] In some embodiments, the multivalent molecule comprises a core attached to multiple nucleotide arms, and wherein individual nucleotide arms comprise a nucleotide unit which is a nucleotide analog having a chain terminating moiety (*e.g.*, blocking moiety) at the sugar 2' position, at the sugar 3' position, or at the sugar 2' and 3' position. In some embodiments, the nucleotide unit comprises a chain terminating moiety (*e.g.*, blocking moiety) at the sugar 2' position, at the sugar 3' position, or at the sugar 2' and 3' position. In some embodiments, the chain terminating moiety can inhibit polymerase-catalyzed incorporation of a subsequent nucleotide unit or free nucleotide in a nascent strand during a primer extension reaction. In some embodiments, the chain terminating moiety is attached to the 3' sugar hydroxyl position where the sugar comprises a ribose or deoxyribose sugar moiety. In some embodiments, the chain terminating moiety is removable/cleavable from the 3' sugar hydroxyl position to generate a nucleotide having a 3'OH sugar group which is extendible with a subsequent nucleotide in a polymerase-catalyzed nucleotide incorporation reaction. In some embodiments, the chain terminating moiety comprises an alkyl group, alkenyl group, alkynyl group, allyl group, aryl group, benzyl group, azide group, amine group, amide group, keto group, isocyanate group, phosphate group, thio group, disulfide group, carbonate group, urea group, acetal group or silyl group. In some embodiments, the chain terminating moiety is cleavable/removable from the nucleotide unit, for example by reacting the chain terminating moiety with a chemical agent, pH change, light or heat. In some embodiments, the chain terminating moieties alkyl, alkenyl, alkynyl and allyl are cleavable with tetrakis(triphenylphosphine)palladium(0) (Pd(PPh₃)₄) with piperidine, or with 2,3-Dichloro-5,6-dicyano-1,4-benzo-quinone (DDQ). In some embodiments, the chain terminating moieties aryl and benzyl are cleavable with H₂ Pd/C. In some embodiments, the chain terminating moieties amine, amide, keto, isocyanate, phosphate, thio, disulfide are cleavable with phosphine or with a thiol group including beta-mercaptoethanol or dithiothritol (DTT). In some embodiments, the chain terminating moiety carbonate is cleavable with potassium carbonate (K₂CO₃) in MeOH, with triethylamine in pyridine, or with Zn in acetic acid (AcOH). In some embodiments, the chain terminating moieties urea

and silyl are cleavable with tetrabutylammonium fluoride, pyridine-HF, with ammonium fluoride, or with triethylamine trihydrofluoride.

[00200] In some embodiments, the nucleotide unit comprises a chain terminating moiety (*e.g.*, blocking moiety) at the sugar 2' position, at the sugar 3' position, or at the sugar 2' and 3' position. In some embodiments, the chain terminating moiety comprises an azide, azido or azidomethyl group. In some embodiments, the chain terminating moiety comprises a 3'-O-azido or 3'-O-azidomethyl group. In some embodiments, the chain terminating moieties azide, azido and azidomethyl group are cleavable/removable with a phosphine compound. In some embodiments, the phosphine compound comprises a derivatized tri-alkyl phosphine moiety or a derivatized tri-aryl phosphine moiety. In some embodiments, the phosphine compound comprises Tris(2-carboxyethyl)phosphine (TCEP) or bis-sulfo triphenyl phosphine (BS-TTP) or Tri(hydroxypropyl)phosphine (THPP). In some embodiments, the cleaving agent comprises 4-dimethylaminopyridine (4-DMAP).

[00201] In some embodiments, the nucleotide unit comprising a chain terminating moiety which is selected from the group consisting of 3'-deoxy nucleotides, 2',3'-dideoxynucleotides, 3'-methyl, 3'-azido, 3'-azidomethyl, 3'-O-azidoalkyl, 3'-O-ethynyl, 3'-O-aminoalkyl, 3'-O-fluoroalkyl, 3'-fluoromethyl, 3'-difluoromethyl, 3'-trifluoromethyl, 3'-sulfonyl, 3'-malonyl, 3'-amino, 3'-O-amino, 3'-sulfhydryl, 3'-aminomethyl, 3'-ethyl, 3'-butyl, 3'-*tert* butyl, 3'-Fluorenylmethyloxycarbonyl, 3'-*tert*-Butyloxycarbonyl, 3'-O-alkyl hydroxylamino group, 3'-phosphorothioate, and 3-O-benzyl, or derivatives thereof.

[00202] In some embodiments, the multivalent molecule comprises a core attached to multiple nucleotide arms, wherein the nucleotide arms comprise a spacer, linker and nucleotide unit, and wherein the core, linker and/or nucleotide unit is labeled with detectable reporter moiety. In some embodiments, the detectable reporter moiety comprises a fluorophore. In some embodiments, a particular detectable reporter moiety (*e.g.*, fluorophore) that is attached to the multivalent molecule can correspond to the base (*e.g.*, dATP, dGTP, dCTP, dTTP or dUTP) of the nucleotide unit to permit detection and identification of the nucleotide base.

[00203] In some embodiments, at least one nucleotide arm of a multivalent molecule has a nucleotide unit that is attached to a detectable reporter moiety. In some embodiments, the detectable reporter moiety is attached to the nucleotide base. In some embodiments, the detectable reporter moiety comprises a fluorophore. In some embodiments, a particular detectable reporter moiety (*e.g.*, fluorophore) that is attached to the multivalent molecule can

correspond to the base (*e.g.*, dATP, dGTP, dCTP, dTTP or dUTP) of the nucleotide unit to permit detection and identification of the nucleotide base.

[00204] In some embodiments, the core of a multivalent molecule comprises an avidin-like or streptavidin-like moiety and the core attachment moiety comprises biotin. In some embodiments, the core comprises an streptavidin-type or avidin-type moiety which includes an avidin protein, as well as any derivatives, analogs and other non-native forms of avidin that can bind to at least one biotin moiety. Other forms of avidin moieties include native and recombinant avidin and streptavidin as well as derivatized molecules, *e.g.* non-glycosylated avidin and truncated streptavidins. For example, an avidin moiety may include de-glycosylated forms of avidin, bacterial streptavidin produced by *Streptomyces* (*e.g.*, *Streptomyces avidinii*), as well as derivatized forms, for example, N-acyl avidins, *e.g.*, N-acetyl, N-phthalyl and N-succinyl avidin, and the commercially-available products EXTRAVIDIN™, CAPTAVIDIN™, NEUTRAVIDIN™ and NEUTRALITE AVIDIN™.

[00205] In some embodiments, any of the methods for sequencing nucleic acid molecules described herein can include forming a binding complex, where the binding complex comprises (i) a polymerase, a nucleic acid concatemer molecule duplexed with a primer, and a nucleotide, or the binding complex comprises (ii) a polymerase, a nucleic acid concatemer molecule duplexed with a primer, and a nucleotide unit of a multivalent molecule. In some embodiments, the binding complex has a persistence time of greater than about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 or 1 second. The binding complex has a persistence time of greater than about 0.1-0.25 seconds, or about 0.25-0.5 seconds, or about 0.5-0.75 seconds, or about 0.75-1 second, or about 1-2 seconds, or about 2-3 seconds, or about 3-4 second, or about 4-5 seconds, and/or wherein the method is or may be carried out at a temperature of at or above 15 °C, at or above 20 °C, at or above 25 °C, at or above 35 °C, at or above 37 °C, at or above 42 °C at or above 55 °C at or above 60 °C, or at or above 72 °C, or at or above 80 °C, or within a range defined by any of the foregoing. The binding complex (*e.g.*, ternary complex) remains stable until subjected to a condition that causes dissociation of interactions between any of the polymerase, template molecule, primer and/or the nucleotide unit or the nucleotide. For example, a dissociating condition comprises contacting the binding complex with any one or any combination of a detergent, EDTA and/or water. In some embodiments, the present disclosure provides said method wherein the binding complex is deposited on, attached to, or hybridized to, a surface showing a contrast to noise ratio in the detecting step of greater than 20. In some embodiments, the present disclosure provides said method

wherein the contacting is performed under a condition that stabilizes the binding complex when the nucleotide or nucleotide unit is complementary to a next base of the template nucleic acid, and destabilizes the binding complex when the nucleotide or nucleotide unit is not complementary to the next base of the template nucleic acid.

Automated Mode

[00206] In any of the methods described herein, any combination of the steps for preparing spatially resolved nucleic acids can be performed in an automated mode using the fluid dispensing system, including cell seeding, cell fixation, cell permeabilization, reverse transcription reactions, circularization oligonucleotide hybridization, circularization oligonucleotide ligation reaction, rolling circle amplification, and sequencing.

[00207] In some embodiments, the cellular sample is deposited on the flowcell (*e.g.*, support). The flowcell can be coated with a reagent that promotes cell adhesion or sticking to the flowcell. The flowcell, having a cell sample adhered thereon, can be placed onto a sequencing apparatus having a flowcell holder/cradle which is fluidically connected to an automated fluid dispensing system and configured on a fluorescent microscope. In some embodiments, the sequencing apparatus can be configured with at least one fluidic delivery device, at least one fluidics device (*e.g.*, microfluidics device), at least one imaging device and/or at least one sensor to detect signals from the sequencing reactions.

[00208] In some embodiments, the automated fluid dispensing system can be used to deliver fixation reagents to the cellular sample on the flowcell, and the cellular sample can be incubated under conditions suitable for cell fixation.

[00209] In some embodiments, the automated fluid dispensing system can be used to deliver permeabilization reagents to the fixed cellular sample on the flowcell, and the cellular sample can be incubated under conditions suitable for cell permeabilization.

[00210] In some embodiments, the automated fluid dispensing system can be used to deliver reagents for conducting reverse transcription of RNA on the flowcell under a condition suitable for generating a plurality of cDNA immobilized to the flowcell.

[00211] In some embodiments, the automated fluid dispensing system can be used to deliver reagents for conducting circularization oligonucleotide hybridization and ligation under a condition suitable for generating a plurality of covalently closed circular molecules on the flowcell.

[00212] In some embodiments, the automated fluid dispensing system can be used to deliver reagents for conducting rolling circle amplification under a condition suitable for generating a plurality of concatemer molecules immobilized to the flowcell.

[00213] In some embodiments, the automated fluid dispensing system can be used to deliver sequencing reagents for conducting sequencing cycles of the concatemer molecules under a condition suitable for generating a plurality of sequencing read products on the flowcell. In some embodiments, individual cycle times can be achieved in less than 30 minutes. In some embodiments, the field of view (FOV) can exceed 1 mm² and the cycle time for scanning large area (> 10 mm²) can be less than 5 minutes.

[00214] In some embodiments, the automated fluid dispensing system can be used to deliver reagents for removing the plurality of sequencing read products from the concatemer molecules and retaining the concatemer molecules on the flowcell.

Supports and Low Non-Specific Coatings

[00215] In some aspects, the present disclosure provides pairwise sequencing compositions and methods which employ a support comprising a plurality of oligonucleotide surface primers immobilized thereon. In some embodiments, the support is passivated with a low non-specific binding coating. In some embodiments, the surface coatings described herein exhibit very low non-specific binding to reagents typically used for nucleic acid capture, amplification and sequencing workflows, such as dyes, nucleotides, enzymes, and nucleic acid primers. The surface coatings can exhibit low background fluorescence signals or high contrast-to-noise (CNR) ratios compared to conventional surface coatings.

[00216] In some embodiments, the low non-specific binding coating comprises one layer or multiple layers (**FIG. 1**). In some embodiments, the plurality of surface primers are immobilized to the low non-specific binding coating. In some embodiments, at least one surface primer is embedded within the low non-specific binding coating. In some embodiments, the low non-specific binding coating enables improved nucleic acid hybridization and amplification performance. In some embodiments, the supports comprise a substrate (or support structure), one or more layers of a covalently or non-covalently attached low-binding, chemical modification layers, *e.g.*, silane layers, polymer films, and one or more covalently or non-covalently attached surface primers that can be used for tethering single-stranded nucleic acid library molecules to the support. In some embodiments, the formulation of the coating, *e.g.*, the chemical composition of one or more layers, the coupling chemistry used to cross-link the one or more layers to the support and/or to each other, and

the total number of layers, may be varied such that non-specific binding of proteins, nucleic acid molecules, and other hybridization and amplification reaction components to the coating is minimized or reduced relative to a comparable monolayer. The formulation of the coating described herein may be varied such that non-specific hybridization on the coating is minimized or reduced relative to a comparable monolayer. The formulation of the coating may be varied such that non-specific amplification on the coating is minimized or reduced relative to a comparable monolayer. The formulation of the coating may be varied such that specific amplification rates and/or yields on the coating are maximized. Amplification levels suitable for detection are achieved in no more than 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, or more than 30 amplification cycles in some cases disclosed herein.

[00217] The support structure that comprises the one or more chemically-modified layers, *e.g.*, layers of a low non-specific binding polymer, may be independent or integrated into another structure or assembly. For example, in some embodiments, the support structure may comprise one or more surfaces within an integrated or assembled microfluidic flow cell. The support structure may comprise one or more surfaces within a microplate format, *e.g.*, the bottom surface of the wells in a microplate. In some embodiments, the support structure comprises the interior surface (such as the lumen surface) of a capillary. In some embodiments, the support structure comprises the interior surface (such as the lumen surface) of a capillary etched into a planar chip.

[00218] The attachment chemistry used to graft a first chemically-modified layer to the surface of the support will generally be dependent on both the material from which the surface is fabricated and the chemical nature of the layer. In some embodiments, the first layer may be covalently attached to the surface. In some embodiments, the first layer may be non-covalently attached, *e.g.*, adsorbed to the support through non-covalent interactions such as electrostatic interactions, hydrogen bonding, or van der Waals interactions between the support and the molecular components of the first layer. In either case, the support may be treated prior to attachment or deposition of the first layer. Any of a variety of surface preparation techniques known to those of skill in the art may be used to clean or treat the surface. For example, glass or silicon surfaces may be acid-washed using a Piranha solution (a mixture of sulfuric acid (H_2SO_4) and hydrogen peroxide (H_2O_2)), base treatment in KOH and NaOH, and/or cleaned using an oxygen plasma treatment method.

[00219] Silane chemistries constitute non-limiting approaches for covalently modifying the silanol groups on glass or silicon surfaces to attach more reactive functional groups (*e.g.*, amines or carboxyl groups), which may then be used in coupling linker molecules (*e.g.*, linear

hydrocarbon molecules of various lengths, such as C6, C12, C18 hydrocarbons, or linear polyethylene glycol (PEG) molecules) or layer molecules (*e.g.*, branched PEG molecules or other polymers) to the surface. Examples of suitable silanes that may be used in creating any of the disclosed low binding coatings include, but are not limited to, (3-Aminopropyl) trimethoxysilane (APTMS), (3-Aminopropyl) triethoxysilane (APTES), any of a variety of PEG-silanes (*e.g.*, comprising molecular weights of 1K, 2K, 5K, 10K, 20K, etc.), amino-PEG silane (*i.e.*, comprising a free amino functional group), maleimide-PEG silane, biotin-PEG silane, and the like.

[00220] Any of a variety of molecules known to those of skill in the art including, but not limited to, amino acids, peptides, nucleotides, oligonucleotides, other monomers or polymers, or combinations thereof may be used in creating the one or more chemically-modified layers on the support, where the choice of components used may be varied to alter one or more properties of the layers, *e.g.*, the surface density of functional groups and/or tethered oligonucleotide primers, the hydrophilicity/hydrophobicity of the layers, or the three dimensional nature (*i.e.*, “thickness”) of the layer. Examples of polymers that may be used to create one or more layers of low non-specific binding material in any of the disclosed coatings include, but are not limited to, polyethylene glycol (PEG) of various molecular weights and branching structures, streptavidin, polyacrylamide, polyester, dextran, poly-lysine, and poly-lysine copolymers, or any combination thereof. Examples of conjugation chemistries that may be used to graft one or more layers of material (*e.g.* polymer layers) to the surface and/or to cross-link the layers to each other include, but are not limited to, biotin-streptavidin interactions (or variations thereof), his tag – Ni/NTA conjugation chemistries, methoxy ether conjugation chemistries, carboxylate conjugation chemistries, amine conjugation chemistries, NHS esters, maleimides, thiol, epoxy, azide, hydrazide, alkyne, isocyanate, and silane.

[00221] The low non-specific binding surface coating may be applied uniformly across the support. Alternatively, the surface coating may be patterned, such that the chemical modification layers are confined to one or more discrete regions of the support. For example, the coating may be patterned using photolithographic techniques to create an ordered array or random pattern of chemically-modified regions on the support. Alternately or in combination, the coating may be patterned using, *e.g.*, contact printing and/or ink-jet printing techniques. In some embodiments, an ordered array or random pattern of chemically-modified regions may comprise at least 1, 5, 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, or 10,000 or more discrete regions.

[00222] In some embodiments, the low nonspecific binding coatings comprise hydrophilic polymers that are non-specifically adsorbed or covalently grafted to the support. Typically, passivation is performed utilizing poly(ethylene glycol) (PEG, also known as polyethylene oxide (PEO) or polyoxyethylene) or other hydrophilic polymers with different molecular weights and end groups that are linked to a support using, for example, silane chemistry. The end groups distal from the surface can include, but are not limited to, biotin, methoxy ether, carboxylate, amine, NHS ester, maleimide, and bis-silane. In some embodiments, two or more layers of a hydrophilic polymer, *e.g.*, a linear polymer, branched polymer, or multi-branched polymer, may be deposited on the surface. In some embodiments, two or more layers may be covalently coupled to each other or internally cross-linked to improve the stability of the resulting coating. In some embodiments, surface primers with different nucleotide sequences and/or base modifications (or other biomolecules, *e.g.*, enzymes or antibodies) may be tethered to the resulting layer at various surface densities. In some embodiments, for example, both surface functional group density and surface primer concentration may be varied to attain a desired surface primer density range. Additionally, surface primer density can be controlled by diluting the surface primers with other molecules that carry the same functional group. For example, and without limitation, amine-labeled surface primers can be diluted with amine-labeled polyethylene glycol in a reaction with an NHS-ester coated surface to reduce the final primer density. Surface primers with different lengths of linker between the hybridization region and the surface attachment functional group can also be applied to control surface density. Example of suitable linkers include poly-T and poly-A strands at the 5' end of the primer (*e.g.*, 0 to 20 bases, *e.g.*, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20), PEG linkers (*e.g.*, 3 to 20 monomer units, *e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 monomer units), and carbon-chain (*e.g.*, C6, C12, C18, etc.). To measure the primer density, fluorescently-labeled primers may be tethered to the surface and a fluorescence reading then compared with that for a dye solution of known concentration.

[00223] In some embodiments, the low nonspecific binding coatings comprise a functionalized polymer coating layer covalently bound at least to a portion of the support via a chemical group on the support, a primer grafted to the functionalized polymer coating, and a water-soluble protective coating on the primer and the functionalized polymer coating. In

some embodiments, the functionalized polymer coating comprises a poly(N-(5-azidoacetamidylpentyl)acrylamide-co-acrylamide (PAZAM).

[00224] In order to scale primer surface density and add additional dimensionality to hydrophilic or amphoteric coatings, supports comprising multi-layer coatings of PEG and other hydrophilic polymers have been developed. By using hydrophilic and amphoteric surface layering approaches that include, but are not limited to, the polymer/co-polymer materials described below, it is possible to increase primer loading density on the support significantly. Traditional PEG coating approaches use monolayer primer deposition, which have been generally reported for single molecule applications, but do not yield high copy numbers for nucleic acid amplification applications. As described herein “layering” can be accomplished using traditional crosslinking approaches with any compatible polymer or monomer subunits such that a surface comprising two or more highly crosslinked layers can be built sequentially. Examples of suitable polymers include, but are not limited to, streptavidin, poly acrylamide, polyester, dextran, poly-lysine, and copolymers of poly-lysine and PEG. In some embodiments, the different layers may be attached to each other through any of a variety of conjugation reactions including, but not limited to, biotin-streptavidin binding, azide-alkyne click reaction, amine-NHS ester reaction, thiol-maleimide reaction, and ionic interactions between positively charged polymer and negatively charged polymer. In some embodiments, high primer density materials may be constructed in solution and subsequently layered onto the surface in multiple steps.

[00225] Examples of materials from which the support structure may be fabricated include, but are not limited to, glass, fused-silica, silicon, a polymer (*e.g.*, polystyrene (PS), macroporous polystyrene (MPPS), polymethylmethacrylate (PMMA), polycarbonate (PC), polypropylene (PP), polyethylene (PE), high density polyethylene (HDPE), cyclic olefin polymers (COP), cyclic olefin copolymers (COC), polyethylene terephthalate (PET)), or any combination thereof. Various compositions of both glass and plastic support structures are contemplated.

[00226] The support structure may be rendered in any of a variety of geometries and dimensions known to those of skill in the art, and may comprise any of a variety of materials known to those of skill in the art. For example, the support structure may be locally planar (*e.g.*, comprising a microscope slide or the surface of a microscope slide). Globally, the support structure may be cylindrical (*e.g.*, comprising a capillary or the interior surface of a capillary), spherical (*e.g.*, comprising the outer surface of a non-porous bead), or irregular (*e.g.*, comprising the outer surface of an irregularly-shaped, non-porous bead or particle). In

some embodiments, the surface of the support structure used for nucleic acid hybridization and amplification may be a solid, non-porous surface. In some embodiments, the surface of the support structure used for nucleic acid hybridization and amplification may be porous, such that the coatings described herein penetrate the porous surface, and nucleic acid hybridization and amplification reactions performed thereon may occur within the pores.

[00227] The support structure that comprises the one or more chemically-modified layers, *e.g.*, layers of a low non-specific binding polymer, may be independent or integrated into another structure or assembly. For example, the support structure may comprise one or more surfaces within an integrated or assembled microfluidic flow cell. The support structure may comprise one or more surfaces within a microplate format, *e.g.*, the bottom surface of the wells in a microplate. In some embodiments, the support structure comprises the interior surface (such as the lumen surface) of a capillary. In some embodiments the support structure comprises the interior surface (such as the lumen surface) of a capillary etched into a planar chip.

[00228] As described herein, the low non-specific binding supports of the present disclosure exhibit reduced non-specific binding of proteins, nucleic acids, and other components of the hybridization and/or amplification formulation used for solid-phase nucleic acid amplification. The degree of non-specific binding exhibited by a given support surface may be assessed either qualitatively or quantitatively. For example, exposure of the surface to fluorescent dyes (*e.g.*, cyanins such as Cy3, or Cy5, etc., fluoresceins, coumarins, rhodamines, etc. or other dyes disclosed herein), fluorescently-labeled nucleotides, fluorescently-labeled oligonucleotides, and/or fluorescently-labeled proteins (*e.g.* polymerases) under a standardized set of conditions, followed by a specified rinse protocol and fluorescence imaging may be used as a qualitative tool for comparison of non-specific binding on supports comprising different surface formulations. In some embodiments, exposure of the surface to fluorescent dyes, fluorescently-labeled nucleotides, fluorescently-labeled oligonucleotides, and/or fluorescently-labeled proteins (*e.g.* polymerases) under a standardized set of conditions, followed by a specified rinse protocol and fluorescence imaging may be used as a quantitative tool for comparison of non-specific binding on supports comprising different surface formulations—provided that care has been taken to ensure that the fluorescence imaging is performed under conditions where fluorescence signal is linearly related (or related in a predictable manner) to the number of fluorophores on the support surface (*e.g.*, under conditions where signal saturation and/or self-quenching of the fluorophore is not an issue) and suitable calibration standards are used. In some

embodiments, other techniques known to those of skill in the art, for example, radioisotope labeling and counting methods may be used for quantitative assessment of the degree to which non-specific binding is exhibited by the different support surface formulations of the present disclosure.

[00229] Some surfaces disclosed herein exhibit a ratio of specific to nonspecific binding of a fluorophore such as Cy3 of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 75, 100, or greater than 100, or any intermediate value spanned by the range herein. Some surfaces disclosed herein exhibit a ratio of specific to nonspecific fluorescence of a fluorophore such as Cy3 of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 75, 100, or greater than 100, or any intermediate value spanned by the range herein.

[00230] In some embodiments, the degree of non-specific binding exhibited by the disclosed low-binding supports may be assessed using a standardized protocol for contacting the surface with a labeled protein (*e.g.*, bovine serum albumin (BSA), streptavidin, a DNA polymerase, a reverse transcriptase, a helicase, a single-stranded binding protein (SSB), etc., or any combination thereof), a labeled nucleotide, a labeled oligonucleotide, etc., under a standardized set of incubation and rinse conditions, followed by detection of the amount of label remaining on the surface and comparison of the signal resulting therefrom to an appropriate calibration standard. In some embodiments, the label may comprise a fluorescent label. In some embodiments, the label may comprise a radioisotope. In some embodiments, the label may comprise any other detectable label known to one of skill in the art. In some embodiments, the degree of non-specific binding exhibited by a given support surface formulation may thus be assessed in terms of the number of non-specifically bound protein molecules (or nucleic acid molecules or other molecules) per unit area. In some embodiments, the low-binding supports of the present disclosure may exhibit non-specific protein binding (or non-specific binding of other specified molecules, (*e.g.*, cyanins such as Cy3, or Cy5, etc., fluoresceins, coumarins, rhodamines, etc. or other dyes disclosed herein)) of less than 0.001 molecule per μm^2 , less than 0.01 molecule per μm^2 , less than 0.1 molecule per μm^2 , less than 0.25 molecule per μm^2 , less than 0.5 molecule per μm^2 , less than 1 molecule per μm^2 , less than 10 molecules per μm^2 , less than 100 molecules per μm^2 , or less than 1,000 molecules per μm^2 . Those of skill in the art will realize that a given support surface of the present disclosure may exhibit non-specific binding falling anywhere within this range, for example, of less than 86 molecules per μm^2 . For example, some modified surfaces disclosed herein exhibit nonspecific protein binding of less than 0.5

molecule/ μm^2 following contact with a 1 μM solution of Cy3 labeled streptavidin (GE Amersham) in phosphate buffered saline (PBS) buffer for 15 minutes, followed by 3 rinses with deionized water. Some modified surfaces disclosed herein exhibit nonspecific binding of Cy3 dye molecules of less than 0.25 molecules per μm^2 . In independent nonspecific binding assays, 1 μM labeled Cy3 SA (ThermoFisher), 1 μM Cy5 SA dye (ThermoFisher), 10 μM Aminoallyl-dUTP-ATTO-647N (Jena Biosciences), 10 μM Aminoallyl-dUTP-ATTO-Rho 1 (Jena Biosciences), 10 μM Aminoallyl-dUTP-ATTO-Rho 1 (Jena Biosciences), 10 μM 7-Propargylamino-7-deaza-dGTP-Cy5 (Jena Biosciences, and 10 μM 7-Propargylamino-7-deaza-dGTP-Cy3 (Jena Biosciences) were incubated on the low binding coated supports at 37° C. for 15 minutes in a 384 well plate format. Each well was rinsed 2-3 $\times$ with 50 μL deionized RNase/DNase Free water and 2-3 $\times$ with 25 mM ACES buffer pH 7.4. The 384 well plates were imaged on a GE Typhoon instrument using the Cy3, AF555, or Cy5 filter sets (according to dye test performed) as specified by the manufacturer at a PMT gain setting of 800 and resolution of 50-100 μm . For higher resolution imaging, images were collected on an Olympus IX83 microscope (*e.g.*, inverted fluorescence microscope) (Olympus Corp., Center Valley, Pa.) with a total internal reflectance fluorescence (TIRF) objective (100 $\times$, 1.5 NA, Olympus), a CCD camera (*e.g.*, an Olympus EM-CCD monochrome camera, Olympus XM-10 monochrome camera, or an Olympus DP80 color and monochrome camera), an illumination source (*e.g.*, an Olympus 100W Hg lamp, an Olympus 75W Xe lamp, or an Olympus U-HGLGPS fluorescence light source), and excitation wavelengths of 532 nm or 635 nm. Dichroic mirrors were purchased from Semrock (IDEX Health & Science, LLC, Rochester, N.Y.), *e.g.*, 405, 488, 532, or 633 nm dichroic reflectors/beamsplitters, and band pass filters were chosen as 532 LP or 645 LP concordant with the appropriate excitation wavelength. Some modified surfaces disclosed herein exhibit nonspecific binding of dye molecules of less than 0.25 molecules per μm^2 . In some embodiments, the coated support was immersed in a buffer (*e.g.*, 25 mM ACES, pH 7.4) while the image was acquired.

[00231] In some embodiments, the surfaces disclosed herein exhibit a ratio of specific to nonspecific binding of a fluorophore such as Cy3 of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 75, 100, or greater than 100, or any intermediate value spanned by the range herein. In some embodiments, the surfaces disclosed herein exhibit a ratio of specific to nonspecific fluorescence signals for a fluorophore such as Cy3 of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 75, 100, or greater than 100, or any intermediate value spanned by the range herein.

[00232] The low-background surfaces consistent with the disclosure herein may exhibit specific dye attachment (*e.g.*, Cy3 attachment) to non-specific dye adsorption (*e.g.*, Cy3 dye adsorption) ratios of at least 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 15:1, 20:1, 30:1, 40:1, 50:1, or more than 50 specific dye molecules attached per molecule nonspecifically adsorbed. Similarly, when subjected to an excitation energy, low-background surfaces consistent with the disclosure herein to which fluorophores, *e.g.*, Cy3, have been attached may exhibit ratios of specific fluorescence signal (*e.g.*, arising from Cy3-labeled oligonucleotides attached to the surface) to non-specific adsorbed dye fluorescence signals of at least 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 15:1, 20:1, 30:1, 40:1, 50:1, or more than 50:1.

[00233] In some embodiments, the degree of hydrophilicity (or “wettability” with aqueous solutions) of the disclosed support surfaces may be assessed, for example, through the measurement of water contact angles in which a small droplet of water is placed on the surface and its angle of contact with the surface is measured using, *e.g.*, an optical tensiometer. In some embodiments, a static contact angle may be determined. In some embodiments, an advancing or receding contact angle may be determined. In some embodiments, the water contact angle for the hydrophilic, low-binding support surfaced disclosed herein may range from about 0 degrees to about 30 degrees. In some embodiments, the water contact angle for the hydrophilic, low-binding support surfaced disclosed herein may no more than 50 degrees, 40 degrees, 30 degrees, 25 degrees, 20 degrees, 18 degrees, 16 degrees, 14 degrees, 12 degrees, 10 degrees, 8 degrees, 6 degrees, 4 degrees, 2 degrees, or 1 degree. In many cases the contact angle is no more than 40 degrees. Those of skill in the art will realize that a given hydrophilic, low-binding support surface of the present disclosure may exhibit a water contact angle having a value of anywhere within this range.

[00234] In some embodiments, the hydrophilic surfaces disclosed herein facilitate reduced wash times for bioassays, often due to reduced nonspecific binding of biomolecules to the low-binding surfaces. In some embodiments, adequate wash steps may be performed in less than 60, 50, 40, 30, 20, 15, 10, or less than 10 seconds. For example, adequate wash steps may be performed in less than 30 seconds.

[00235] Some low-binding surfaces of the present disclosure may exhibit significant improvement in stability or durability to prolonged exposure to solvents and elevated temperatures, or to repeated cycles of solvent exposure or changes in temperature. For example, the stability of the disclosed surfaces may be tested by fluorescently labeling a functional group on the surface, or a tethered biomolecule (*e.g.*, an oligonucleotide primer) on the surface, and monitoring fluorescence signal before, during, and after prolonged

exposure to solvents and elevated temperatures, or to repeated cycles of solvent exposure or changes in temperature. In some embodiments, the degree of change in the fluorescence used to assess the quality of the surface may be less than 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, or 25% over a time period of 1 minute, 2 minutes, 3 minutes, 4 minutes, 5 minutes, 10 minutes, 20 minutes, 30 minutes, 40 minutes, 50 minutes, 60 minutes, 2 hours, 3 hours, 4 hours, 5 hours, 6 hours, 7 hours, 8 hours, 9 hours, 10 hours, 15 hours, 20 hours, 25 hours, 30 hours, 35 hours, 40 hours, 45 hours, 50 hours, or 100 hours of exposure to solvents and/or elevated temperatures (or any combination of these percentages as measured over these time periods). In some embodiments, the degree of change in the fluorescence used to assess the quality of the surface may be less than 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, or 25% over 5 cycles, 10 cycles, 20 cycles, 30 cycles, 40 cycles, 50 cycles, 60 cycles, 70 cycles, 80 cycles, 90 cycles, 100 cycles, 200 cycles, 300 cycles, 400 cycles, 500 cycles, 600 cycles, 700 cycles, 800 cycles, 900 cycles, or 1,000 cycles of repeated exposure to solvent changes and/or changes in temperature (or any combination of these percentages as measured over this range of cycles).

[00236] In some embodiments, the surfaces disclosed herein may exhibit a high ratio of specific signal to nonspecific signal or other background. For example, when used for nucleic acid amplification, some surfaces may exhibit an amplification signal that is at least 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 75, 100, or greater than 100 fold greater than a signal of an adjacent unpopulated region of the surface. Similarly, some surfaces exhibit an amplification signal that is at least 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 75, 100, or greater than 100 fold greater than a signal of an adjacent amplified nucleic acid population region of the surface.

[00237] In some embodiments, fluorescence images of the disclosed low background surfaces when used in nucleic acid hybridization or amplification applications to create colonies of hybridized or clonally-amplified nucleic acid molecules (*e.g.*, that have been directly or indirectly labeled with a fluorophore) exhibit contrast-to-noise ratios (CNRs) of at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, or greater than 250.

[00238] One or more types of primer may be attached or tethered to the support surface. In some embodiments, the one or more types of adaptors or primers may comprise spacer sequences, adaptor sequences for hybridization to adaptor-ligated target library nucleic acid sequences, forward amplification primers, reverse amplification primers, sequencing primers, and/or molecular barcoding sequences, or any combination thereof. In some embodiments, 1 primer or adaptor sequence may be tethered to at least one layer of the surface. In some

embodiments, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10 different primer or adaptor sequences may be tethered to at least one layer of the surface.

[00239] In some embodiments, the tethered adaptor and/or primer sequences may range in length from about 10 nucleotides to about 100 nucleotides. In some embodiments, the tethered adaptor and/or primer sequences may be at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, or at least 100 nucleotides in length. In some embodiments, the tethered adaptor and/or primer sequences may be at most 100, at most 90, at most 80, at most 70, at most 60, at most 50, at most 40, at most 30, at most 20, or at most 10 nucleotides in length. Any of the lower and upper values described in this paragraph may be combined to form a range included within the present disclosure, for example, in some embodiments the length of the tethered adaptor and/or primer sequences may range from about 20 nucleotides to about 80 nucleotides. Those of skill in the art will recognize that the length of the tethered adaptor and/or primer sequences may have any value within this range, *e.g.*, about 24 nucleotides.

[00240] In some embodiments, the resultant surface density of primers (*e.g.*, capture primers) on the low binding support surfaces of the present disclosure may range from about 100 primer molecules per μm^2 to about 100,000 primer molecules per μm^2 . In some embodiments, the resultant surface density of primers on the low binding support surfaces of the present disclosure may range from about 1,000 primer molecules per μm^2 to about 1,000,000 primer molecules per μm^2 . In some embodiments, the surface density of primers may be at least 1,000, at least 10,000, at least 100,000, or at least 1,000,000 molecules per μm^2 . In some embodiments, the surface density of primers may be at most 1,000,000, at most 100,000, at most 10,000, or at most 1,000 molecules per μm^2 . Any of the lower and upper values described in this paragraph may be combined to form a range included within the present disclosure, for example, in some embodiments the surface density of primers may range from about 10,000 molecules per μm^2 to about 100,000 molecules per μm^2 . Those of skill in the art will recognize that the surface density of primer molecules may have any value within this range, *e.g.*, about 455,000 molecules per μm^2 . In some embodiments, the surface density of target library nucleic acid sequences initially hybridized to adaptor or primer sequences on the support surface may be less than or equal to that indicated for the surface density of tethered primers. In some embodiments, the surface density of clonally-amplified target library nucleic acid sequences hybridized to adaptor or primer sequences on the support surface may span the same range as that indicated for the surface density of tethered primers.

[00241] Local densities as listed above do not preclude variation in density across a surface, such that a surface may comprise a region having an oligo density of, for example, $500,000/\mu\text{m}^2$, while also comprising at least a second region having a substantially different local density.

[00242] In some embodiments, the performance of nucleic acid hybridization and/or amplification reactions using the disclosed reaction formulations and low-binding supports may be assessed using fluorescence imaging techniques, where the contrast-to-noise ratio (CNR) of the images provides a key metric in assessing amplification specificity and non-specific binding on the support. CNR is commonly defined as: $\text{CNR} = (\text{Signal} - \text{Background}) / \text{Noise}$. The background term is commonly taken to be the signal measured for the interstitial regions surrounding a particular feature (diffraction limited spot, DLS) in a specified region of interest (ROI). While signal-to-noise ratio (SNR) is often considered to be a benchmark of overall signal quality, it can be shown that improved CNR can provide a significant advantage over SNR as a benchmark for signal quality in applications that require rapid image capture (*e.g.*, sequencing applications for which cycle times must be minimized), as shown in the example below. At high CNR the imaging time required to reach accurate discrimination (and thus accurate base-calling in the case of sequencing applications) can be drastically reduced even with moderate improvements in CNR. Improved CNR in imaging data on the imaging integration time provides a method for more accurately detecting features such as clonally-amplified nucleic acid colonies on the support surface.

[00243] In most ensemble-based sequencing approaches, the background term is typically measured as the signal associated with 'interstitial' regions. In addition to "interstitial" background (B_{inter}), "intrastitial" background (B_{intra}) exists within the region occupied by an amplified DNA colony. The combination of these two background signals dictates the achievable CNR, and subsequently directly impacts the optical instrument requirements, architecture costs, reagent costs, run-times, cost/genome, and ultimately the accuracy and data quality for cyclic array-based sequencing applications. The B_{inter} background signal arises from a variety of sources; a few examples include auto-fluorescence from consumable flow cells, non-specific adsorption of detection molecules that yield spurious fluorescence signals that may obscure the signal from the ROI, the presence of non-specific DNA amplification products (*e.g.*, those arising from primer dimers). In typical next generation sequencing (NGS) applications, this background signal in the current field-of-view (FOV) is averaged over time and subtracted. The signal arising from individual DNA colonies (*i.e.*, $(\text{Signal}) - B(\text{interstitial})$ in the FOV) yields a discernable feature that can be classified. In some

embodiments, the intrastitial background (B(intrastitial)) can contribute a confounding fluorescence signal that is not specific to the target of interest, but is present in the same ROI thus making it far more difficult to average and subtract.

[00244] Nucleic acid amplification on the low-binding coated supports described herein may decrease the B(interstitial) background signal by reducing non-specific binding, may lead to improvements in specific nucleic acid amplification, and may lead to a decrease in non-specific amplification that can impact the background signal arising from both the interstitial and intrastitial regions. In some embodiments, the disclosed low-binding coated supports, optionally used in combination with the disclosed hybridization and/or amplification reaction formulations, may lead to improvements in CNR by a factor of 2, 5, 10, 100, 250, 500 or 1000-fold over those achieved using conventional supports and hybridization, amplification, and/or sequencing protocols. Although described here in the context of using fluorescence imaging as the read-out or detection mode, the same principles apply to the use of the disclosed low-binding coated supports and nucleic acid hybridization and amplification formulations for other detection modes as well, including both optical and non-optical detection modes.

INCORPORATION BY REFERENCE

[00245] Throughout this application various publications, patents, and/or patent applications are referenced. The disclosures of the publications, patents and/or patent applications are hereby incorporated by reference in their entireties into this application in order to more fully describe the state of the art to which this disclosure pertains.

EQUIVALENTS

[00246] The details of one or more embodiments of the disclosure are set forth in the accompanying description above. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, the preferred methods and materials are now described. Other features, objects, and advantages of the disclosure will be apparent from the description and from the claims. In the specification and the appended claims, the singular forms include plural referents unless the context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. All patents and publications cited in this specification are incorporated by reference.

[00247] The foregoing description has been presented only for the purposes of illustration and is not intended to limit the disclosure to the precise form disclosed, but by the claims appended hereto.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for preparing spatially resolved nucleic acids, comprising:
 - a) providing a support comprising (i) a low non-specific binding coating comprising at least one hydrophilic polymer, wherein the low non-specific binding coating has a water contact angle of no more than 45 degrees, and (ii) a plurality of immobilized surface capture primers covalently tethered to the low non-specific binding coating wherein each individual surface capture primer comprises a universal surface capture primer sequence (110), a universal binding site for a reverse sequencing primer (120), and an RNA capture sequence;
 - b) positioning a cellular sample on the coated support under a condition suitable for the cellular sample to remain in a fixed position on the coated support, disrupting the cellular sample, releasing a plurality of RNA molecules of the cellular sample under a condition that is suitable for preserving spatial location information of the RNA molecules from the cellular sample, and collecting the plurality of RNA molecules on the plurality of immobilized surface capture primers of the support, under a condition suitable for hybridizing individual RNA molecules to individual immobilized surface capture primers to generate a plurality of capture primer-RNA duplexes, wherein each primer-RNA duplex comprises an immobilized surface capture primer hybridized to an RNA molecule;
 - c) conducting a reverse transcription reaction on the coated support under a condition suitable for extending the 3' end of the immobilized surface capture primers and using the hybridized RNA as a template strand, under a condition suitable for generating a non-template polyC tail at the 3' end, thereby generating a plurality of first strand cDNA molecules with a non-template polyC tail at the 3' end, wherein the reverse transcription reaction comprises a plurality of template switching oligonucleotides, wherein each of the template switching oligonucleotides comprises a polyG region that is capable of hybridizing to the non-template polyC tail of the first strand cDNA molecules, a universal binding site for a forward sequencing primer (130), and a universal binding site for a surface pinning primer (140);

- d)** conducting a primer extension reaction from the 3' terminal end of the non-template polyC tail of the plurality of first strand cDNA molecules and using the template switching oligonucleotide as a template strand, thereby generating a plurality of immobilized full length first strand cDNAs, wherein each full length first strand cDNA comprises a universal surface capture primer sequence (110); a universal binding site for a reverse sequencing primer (120); an RNA capture sequence; a cDNA insert region; a polyC tail region; a universal binding site for a forward sequencing primer (130); and a universal binding site for a surface pinning primer (140);
- e)** removing the RNA molecule while retaining the plurality of immobilized full length first strand cDNA;
- f)** contacting the plurality of retained immobilized full length first strand cDNAs with a plurality of single stranded circularization oligonucleotides under a condition suitable for hybridizing individual circularization oligonucleotides to an immobilized full length first strand cDNA, thereby forming a single stranded circular molecule with a gap, wherein individual single stranded circularization oligonucleotides comprise: (i) a sequence that is complementary to the universal binding site for a reverse sequencing primer (120), (ii) a sequence that is complementary to the universal surface capture primer sequence (110), (iii) a linker region, (iv) a sequence that is complementary to the universal binding site for a surface pinning primer (140), (v) and a sequence that is complementary to the universal binding site for a forward sequencing primer (130);
- g)** conducting a polymerase-catalyzed extension reaction to fill in the gap, using the polyC tail region, the cDNA insert region and the RNA capture sequence of the immobilized full length first strand cDNA as a template strand, wherein the polymerase-catalyzed extension reaction forms a single stranded circularized molecule with a nick, and wherein the nick is closed by conducting an enzymatic ligation reaction to generate a single stranded covalently closed circular molecule which is hybridized to the immobilized full length first strand cDNA;
- h)** conducting a rolling circle amplification reaction using the terminal 3' end of the immobilized full length first strand cDNA as an initiation site and the covalently closed circular molecule as a template strand, thereby generating a

- plurality of immobilized nucleic acid concatemers that are spatially resolved on the support;
- i) sequencing the plurality of individual immobilized nucleic acid concatemers, wherein the sequencing comprises at least a portion of the cDNA insert region of the individual nucleic acid concatemers, which correspond to the individual RNA molecules eluted from the cellular sample; and
 - j) determining the location of individual nucleic acid concatemers on the coated support, which correspond to the spatial location of the individual RNA molecules eluted from the cellular sample.
2. The method of claim 1, wherein the cellular sample of step **b)** comprises a single cell, a plurality of cells, a tissue, an organ, an organism, or a sectioned cellular sample.
 3. The method of claim 1, wherein the cellular sample of step **b)** comprises a fresh sample, a frozen sample, a fresh frozen sample, or a formalin-fixed paraffin-embedded sample.
 4. The method of claim 1, wherein the plurality of template switching oligonucleotides of step **c)** comprises chimeric DNA and/or RNA oligonucleotides.
 5. The method of claim 1, wherein the linker region of the single stranded circularization oligonucleotides of step **f)** comprise at least one sample index sequence for multiplexing, at least one unique molecular index (UMI) sequence for molecular tagging and/or at least one universal binding site for a compaction oligonucleotide.
 6. The method of claim 1, wherein the rolling circle amplification reaction of step **h)** is conducted in the presence of a plurality of compaction oligonucleotides, wherein the compaction oligonucleotides comprise single stranded oligonucleotides each having a first region at a first distal end that hybridizes to one portion of a concatemer molecule and a second region at the second distal end that hybridizes to a second portion of the same concatemer molecule, wherein the first portion of the concatemer molecule and the second portion of the concatemer molecule are brought in proximity, thereby causing compaction of the concatemer molecule to form a DNA nanoball.

7. The method of claim 1, wherein the sequencing of step **i)** comprises:
- contacting the plurality of concatemer molecules with a plurality of sequencing polymerases and a plurality of nucleic acid sequencing primers, wherein the contacting is conducted under a condition suitable to form a plurality of complexed sequencing polymerases, wherein each complexed sequencing polymerase comprises a sequencing polymerase bound to a nucleic acid duplex, and wherein the nucleic acid duplex comprises a portion of a concatemer molecule hybridized to the nucleic acid sequencing primer;
 - contacting the plurality of complexed sequencing polymerases with a plurality of detectably labeled nucleotides comprising a blocking moiety at the 2' or 3' sugar position, wherein the contacting is conducted under a condition suitable for binding at least one nucleotide to at least one of the complexed sequencing polymerases, and wherein the condition is suitable for promoting polymerase-catalyzed nucleotide incorporation;
 - incorporating a nucleotide into the 3' end of a sequencing primer of at least one complexed sequencing polymerase;
 - detecting the incorporated nucleotide and identifying the nucleobase of the incorporated nucleotide;
 - removing the blocking moiety from the incorporated nucleotide; and
 - repeating steps b. to e. at least once.
8. The method of claim 1, wherein the sequencing of step **i)** comprises:
- contacting the plurality of concatemer molecules with a plurality of sequencing polymerases and a plurality of nucleic acid sequencing primers, wherein the contacting is conducted under a condition suitable to form a plurality of complexed sequencing polymerases, wherein each complexed sequencing polymerase comprises a sequencing polymerase bound to a nucleic acid duplex, and wherein the nucleic acid duplex comprises a portion of a concatemer molecule hybridized to the nucleic acid sequencing primer;
 - contacting the plurality of complexed sequencing polymerases with a plurality of nucleotides each comprising a detectable label attached to a phosphate moiety of the phosphate chain, wherein the contacting is conducted under a condition suitable for binding at least one nucleotide to at least one of the complexed

- sequencing polymerases, and wherein the condition is suitable for promoting polymerase-catalyzed nucleotide incorporation;
- c. incorporating a nucleotide into the 3' end of a sequencing primer of at least one complexed sequencing polymerase;
 - d. detecting the incorporated nucleotide and identifying the nucleo-base of the incorporated nucleotide; and
 - e. repeating steps b. to d. at least once.
9. The method of claim 1, wherein the sequencing of step i) comprises:
- a. contacting the plurality of concatemer molecules with a plurality of a first sequencing polymerase and a plurality of nucleic acid sequencing primers, wherein the contacting is conducted under a condition suitable to bind the plurality of first polymerases to the plurality of nucleic acid template molecules and the plurality of nucleic acid primers, thereby forming a plurality of first complexed polymerases, wherein each first complexed polymerase comprises a first polymerase bound to a nucleic acid duplex, wherein the nucleic acid duplex comprises a nucleic acid template molecule hybridized to a nucleic acid primer;
 - b. contacting the plurality of first complexed polymerases with a plurality of detectably labeled multivalent molecules to form a plurality of multivalent-binding complexes, wherein individual detectably labeled multivalent molecules in the plurality comprise a core attached to multiple nucleotide arms, and wherein each nucleotide arm is attached to a nucleotide unit, wherein the contacting is conducted under a condition suitable for binding complementary nucleotide units of the multivalent molecules to at least two of the plurality of first complexed polymerases thereby forming a plurality of multivalent-binding complexes, and wherein the condition is suitable for inhibiting incorporation of the complementary nucleotide units into the primers of the plurality of multivalent-binding complexes;
 - c. detecting the plurality of multivalent-binding complexes; and
 - d. identifying the nucleo-base of the complementary nucleotide units in the plurality of multivalent-binding complexes, thereby determining the sequence of the nucleic acid template molecules.

10. The method of claim 9, wherein the sequencing of step **i)** comprises:
 - a. dissociating the plurality of multivalent-binding complexes, by removing the plurality of first sequencing polymerases and their bound multivalent molecules, and retaining the plurality of nucleic acid duplexes;
 - b. contacting the plurality of the retained nucleic acid duplexes of step **e)** with a plurality of second sequencing polymerases under a condition suitable for binding the plurality of second polymerases to the plurality of the retained nucleic acid duplexes, thereby forming a plurality of second complexed polymerases each comprising a second polymerase bound to a nucleic acid duplex; and
 - c. contacting the plurality of second complexed polymerases with a plurality of nucleotides, wherein the contacting is conducted under a condition suitable for binding complementary nucleotides from the plurality of nucleotides to at least two of the second complexed polymerases, thereby forming a plurality of nucleotide-binding complexes, and wherein the condition is suitable for promoting nucleotide incorporation of the bound complementary nucleotides into the primers of the nucleotide-binding complexes.
11. The method of claim 10 further comprising:
 - d. detecting the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases.
12. The method of claim 10 further comprising:
 - d. detecting the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases; and
 - e. identifying the nucleo-bases of the complementary nucleotides which are incorporated into the primers of the nucleotide-complexed polymerases.
13. The method of claim 9, wherein the contacting the plurality of first complexed polymerases with the plurality of multivalent molecules of step **b.** is conducted in the presence of a non-catalytic divalent cation that inhibits polymerase-catalyzed nucleotide incorporation, wherein the non-catalytic divalent cation comprises strontium, barium or calcium.

14. The method of claim 9, wherein individual multivalent molecules in the plurality of multivalent molecules comprise: **(a)** a core; and **(b)** a plurality of nucleotide arms which comprise: **(i)** a core attachment moiety, **(ii)** a spacer, **(iii)** a linker, and **(iv)** a nucleotide unit, wherein the core is attached to the plurality of nucleotide arms via their core attachment moiety, wherein the spacer is attached to the linker, and wherein the linker is attached to the nucleotide unit.
15. The method of claim 14, wherein the linker comprises an aliphatic chain having 2-6 subunits or an oligo ethylene glycol chain having 2-6 subunits.
16. The method of claim 14, wherein the plurality of nucleotide arms attached to a given core have the same type of nucleotide units, and wherein the types of nucleotide units comprise dATP, dGTP, dCTP, dTTP or dUTP.
17. The method of claim 9, wherein the plurality of multivalent molecules comprise one type of a multivalent molecule wherein each multivalent molecule in the plurality has the same type of nucleotide unit selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP.
18. The method of claim 9, wherein the plurality of multivalent molecules comprise a mixture of any combination of two or more types of multivalent molecules each type having nucleotide units selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP.
19. The method of claim 9, further comprising forming a plurality of binding complexes, comprising the steps of:
 - a) binding a first nucleic acid sequencing primer, a first sequencing polymerase, and a first multivalent molecule to a first portion of a concatemer molecule, thereby forming a first binding complex, wherein a first nucleotide unit of the first multivalent molecule binds to the first polymerase; and
 - b) binding a second nucleic acid sequencing primer, a second sequencing polymerase, and the first multivalent molecule to a second portion of the same concatemer molecule, thereby forming a second binding complex, wherein a second nucleotide unit of the first multivalent molecule binds to the second polymerase,

wherein the first and the second binding complexes which include the same multivalent molecule forms an avidity complex.

20. The method of claim 9, further comprising forming an avidity complex comprising the steps of:
- a) contacting the plurality of the first sequencing polymerase and the plurality of nucleic acid sequencing primers with different portions of a concatemer nucleic acid template molecule to form at least first complexed polymerase and a second complexed polymerase on the same concatemer molecule; and
 - b) contacting a plurality of detectably labeled multivalent molecules to the at least first complexed polymerase and the second complexed polymerase on the same concatemer molecule, under conditions suitable to bind a single multivalent molecule from the plurality to the first complexed polymerase and the second complexed polymerase, wherein at least a first nucleotide unit of the single multivalent molecule is bound to the first complexed polymerase, which comprises a first primer hybridized to a first portion of the concatemer molecule thereby forming a first binding complex, and wherein at least a second nucleotide unit of the single multivalent molecule is bound to the second complexed polymerase which comprises a second primer hybridized to a second portion of the concatemer molecule, thereby forming a second binding complex,
 - wherein the contacting is conducted under a condition suitable to inhibit polymerase-catalyzed incorporation of the bound first and second nucleotide units in the first and second binding complexes, and
 - wherein the first and second binding complexes which are bound to the same multivalent molecule forms an avidity complex;
 - c) detecting the first and second binding complexes on the same concatemer molecule; and
 - d) identifying the first nucleotide unit in the first binding complex, thereby determining the sequence of the first portion of the concatemer molecule, and identifying the second nucleotide unit in the second binding complex thereby determining the sequence of the second portion of the concatemer molecule.

21. The method of claim 10, wherein the contacting the plurality of second complexed polymerases with the plurality of nucleotides of step c. is conducted in the presence of a catalytic divalent cation that promotes polymerase-catalyzed nucleotide incorporation, wherein the catalytic divalent cation comprises magnesium or manganese.
22. The method of claim 10, wherein individual nucleotides in the plurality of nucleotides in step c. comprise an aromatic base, a five-carbon sugar, and 1-10 phosphate groups.
23. The method of claim 10, wherein the plurality of nucleotides of step c. comprise one type of nucleotide selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP, or comprise a mixture of any combination of two or more types of nucleotides selected from the group consisting of dATP, dGTP, dCTP, dTTP and dUTP.
24. The method of claim 10, wherein at least one of the nucleotides in the plurality of nucleotides in step c. is labeled with a fluorophore.
25. The method of claim 10, wherein the plurality of nucleotides in step c. lack a fluorophore label.
26. The method of claim 10, wherein at least one of the nucleotides in the plurality of nucleotides of step c. comprises a removable chain terminating moiety attached to the 3' carbon position of the sugar group, wherein the removable chain terminating moiety comprises an alkyl group, alkenyl group, alkynyl group, allyl group, aryl group, benzyl group, azide group, azido group, O-azidomethyl group, amine group, amide group, keto group, isocyanate group, phosphate group, thio group, disulfide group, carbonate group, urea group, or silyl group, and wherein the removable chain terminating moiety is cleavable with a chemical compound to generate an extendible 3'OH moiety on the sugar group.