DIRECT *IN SITU* SEQUENCING

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 63/549,891, filed February 5, 2024, and U.S. Provisional Patent Application No. 63/652,540, filed May 28, 2024, which are herein incorporated by reference in their entireties for all purposes.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0002] The present application is being filed along with a Sequence Listing in electronic format. The Sequence Listing is provided as a file entitled 52933-761.601-PCT SL.XML, created February 4, 2025, which is 6.16 kilobytes in size. The information in the electronic format of the Sequence Listing is incorporated by reference in its entirety.

BACKGROUND

[0003] Nucleic acid sequencing can be used to obtain information in a wide variety of biomedical contexts, including diagnostics, prognostics, biotechnology, and forensic biology. Various sequencing methods have been developed including Maxam-Gilbert sequencing and chain-termination methods, or de novo sequencing methods including shotgun sequencing and bridge PCR, or next-generation methods including polony sequencing, 454 pyrosequencing, Illumina sequencing, SOLiD sequencing, Ion Torrent semiconductor sequencing, HeliScope single molecule sequencing, SMRT® sequencing, and others. Despite advances in DNA sequencing, many challenges still remain unaddressed.

SUMMARY

[0004] Aspects disclosed herein provide methods for identifying a ribonucleic acid (RNA) *in situ*, the method comprising: (a) providing a cell or a tissue that has been permeabilized, wherein the cell or the tissue comprises the RNA; (b) directing a probe to the cell or tissue, wherein the probe comprises (i) an RNA-binding sequence and (ii) an index sequence, under conditions sufficient to couple the RNA-binding sequence to at least a portion of the RNA; (c) amplifying the index sequence to generate one or more amplicons, wherein the one or more amplicons comprises multiple copies of the index sequence or derivative thereof; (d) directing a primer nucleic acid sequence to the cell or the tissue under conditions sufficient to bind the primer nucleic acid sequence to a portion of the one or more amplicons to produce a primed index sequence; (e)

directing the cell or the tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or the derivative thereof, under conditions suitable to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (f) detecting the binding complex within the cell or tissue, thereby identifying the RNA. In some embodiments, the method does not comprise performing a reverse transcription reaction of the RNA. In some embodiments, the method further comprises providing the cell in (a). In some embodiments, the cell is a cancer cell. In some embodiments, the method further comprises providing the tissue in (a). In some embodiments, the tissue is a formalin-fixed paraffin embedded tissue. In some embodiments, the tissue is a fresh-frozen tissue. In some embodiments, the cell or the tissue is extracted from a subject. In some embodiments, the subject is a human subject. In some embodiments, the index sequence comprises a nucleic acid. In some embodiments, the index sequence is 4-50 nucleotides in length. In some embodiments, at least a portion of the index sequence or derivative thereof recognizes the primer nucleic acid sequence. In some embodiments, the primer nucleic acid sequence comprises a blocking group. In some embodiments, the blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence. In some embodiments, the blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group. In some embodiments, the primer nucleic acid sequence does not include a 3' hydroxyl group. In some embodiments, the primer nucleic acid sequence binds to at least a portion of the index sequence. In some embodiments, the primer nucleic acid sequence binds to at least a portion of a reverse complement of the index sequence. In some embodiments, the primed index sequence comprises a double-stranded nucleic acid. In some embodiments, the primed index sequence comprises a single-stranded nucleic acid. In some embodiments, the cell or the tissue is immobilized on a surface of a substrate. In some embodiments, the substrate has a water contact angle of less than or equal to 45 degrees. In some embodiments, the substrate is an interior surface of a flow cell. In some embodiments, the interior surface of the flow cell comprises one or more hydrophilic polymer layers. In some embodiments, the one or more hydrophilic polymer layers comprises a branched polymer. In some embodiments, the one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol. In some embodiments, the nucleotide conjugate comprises a common core. In some embodiments, the one or more nucleotide moieties are coupled to the common core. In some embodiments, the common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot. In some embodiments, the polymer comprises a protein. In some embodiments, in (e) the nucleotide conjugate is included

in a mixture of a plurality of nucleotide conjugates, wherein each of the plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other. In some embodiments, the different types of the nucleotide moieties comprise at least three different types of the nucleotide moieties. In some embodiments, the one or more nucleotide moieties do not comprise a blocking group coupled thereto. In some embodiments, the detecting in (f) comprises imaging the cell or the tissue using one or more image sensors. In some embodiments, the one or more image sensors comprises a photodetector array. In some embodiments, the method further comprises washing the cell or the tissue following (f) to remove the nucleotide conjugate from the cell or the tissue. In some embodiments, the method further comprises repeating (e) and (f) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in the primed index sequence or derivative thereof. In some embodiments, the detectable label comprises a fluorophore. In some embodiments, the detectable label comprises a quantum dot. In some embodiments, the conditions sufficient to couple the target-binding sequence to at least a portion of the target in (b) comprises an incubation temperature, wherein the incubation temperature is about 20-45°C. In some embodiments, the conditions sufficient to couple the target-binding sequence to at least a portion of the target in (b) comprises a salt. In some embodiments, the RNA is a messenger RNA. In some embodiments, the RNA is a micro RNA. In some embodiments, the RNA is a ribosomal RNA. In some embodiments, the RNA is a transfer RNA. In some embodiments, the RNA is a transfer-messenger RNA. In some embodiments, the RNA comprises a modification. In some embodiments, the modification is a 5' phosphorylation. In some embodiments, the modification is an inter-nucleotide linkage. In some embodiments, the probe comprises nucleic acid. In some embodiments, the nucleic acid comprises deoxyribonucleic acid (DNA). In some embodiments, the nucleic acid comprises RNA. In some embodiments, the RNA-binding sequence comprises nucleic acid. In some embodiments, the nucleic acid comprises deoxyribonucleic acid (DNA). In some embodiments, the nucleic acid comprises RNA. In some embodiments, the probe is a padlock probe. In some embodiments, the method further comprises, before (c), ligating a first end of the probe to a second end of the probe using a ligase. In some embodiments, the ligase is a T4 ligase. In some embodiments, the ligase is a SplintR ligase. In some embodiments, performing the amplification reaction in (c) comprises performing a rolling circle amplification reaction using a polymerizing enzyme. In some embodiments, the polymerizing enzyme is a phi29 polymerase. In some embodiments, the cell or the tissue is permeabilized using a solvent. In some embodiments, the solvent is methanol. In some embodiments, the solvent is acetone. In some embodiments, the cell or the tissue is permeabilized using a detergent. In some embodiments, the detergent is saponin. In some embodiments, the detergent is Triton X-100.

[0005] Aspects disclosed herein provide methods for identifying a polypeptide *in situ*, the method comprising: (a) providing a cell or a tissue that has been permeabilized, wherein the cell or the tissue comprises the polypeptide; (b) directing a polypeptide-binding probe to the cell or tissue, wherein the polypeptide-binding probe comprises a polypeptide binding moiety and an oligonucleotide coupled thereto, under conditions sufficient to couple the polypeptide-binding probe to the polypeptide; (c) directing a probe to the cell or the tissue, wherein the with a probe comprising (i) an oligonucleotide-binding probe and (ii) an index sequence under conditions sufficient to couple the oligonucleotide-binding probe to at least a portion of the oligonucleotide of the polypeptide-binding probe; (d) amplifying the index sequence to generate one or more amplicons, wherein the one or more amplicons comprises multiple copies of the index sequence or derivative thereof; (e) directing a primer nucleic acid sequence to the cell or the tissue under conditions sufficient to bind the primer nucleic acid sequence to at least a portion of the one or more amplicons to produce a primed index sequence; (f) directing the cell or the tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or the one or more amplicons thereof, under conditions suitable to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (g) detecting the binding complex within the cell or tissue, thereby identifying the polypeptide. In some embodiment, the method further comprises providing the cell in (a). In some embodiment, the cell is a cancer cell. In some embodiment, the method further comprises providing the tissue in (a). In some embodiment, the tissue is a formalin-fixed paraffin embedded tissue. In some embodiment, the tissue is a fresh-frozen tissue. In some embodiment, the cell or the tissue is extracted from a subject. In some embodiment, the subject is a human subject. In some embodiment, the index sequence comprises a nucleic acid. In some embodiment, the index sequence is 4-50 nucleotides in length. In some embodiment, at least a portion of the index sequence or derivative thereof recognizes the primer nucleic acid sequence. In some embodiment, the primer nucleic acid sequence comprises a blocking group. In some embodiment, the blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence. In some embodiment, the blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group. In some embodiment, the primer nucleic acid sequence does not include a 3' hydroxyl group. In some embodiment, the primer nucleic acid sequence binds to at least a portion of the index sequence. In some embodiment, the primer nucleic acid sequence binds to at least a portion of a reverse complement of the index sequence. In some embodiment, the primed index sequence comprises a double-

stranded nucleic acid. In some embodiment, the primed index sequence comprises a single-stranded nucleic acid. In some embodiment, the cell or the tissue is immobilized on a surface of a substrate. In some embodiment, the surface of the substrate has a water contact angle of less than or equal to 45 degrees. In some embodiment, the surface of the substrate is an interior surface of a flow cell. In some embodiment, the interior surface of the flow cell comprises one or more hydrophilic polymer layers. In some embodiment, the one or more hydrophilic polymer layers comprises a branched polymer. In some embodiment, the one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol. In some embodiment, the nucleotide conjugate comprises a common core. In some embodiment, the one or more nucleotide moieties are coupled to the common core. In some embodiment, the common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot. In some embodiment, the polymer comprises a protein. In some embodiment, in (e) the nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of the plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other. In some embodiment, the different types of the nucleotide moieties comprise at least three different types of the nucleotide moieties. In some embodiment, the one or more nucleotide moieties do not comprise a blocking group coupled thereto. In some embodiment, the detecting in (f) comprises imaging the cell or the tissue using one or more image sensors. In some embodiment, the one or more image sensors comprises a photodetector array. In some embodiments, the method further comprises washing the cell or the tissue following (f) to remove the nucleotide conjugate from the cell or the tissue. In some embodiment, the method further comprises repeating (e) and (f) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in the primed index sequence or derivative thereof. In some embodiment, the detectable label comprises a fluorophore. In some embodiment, the detectable label comprises a quantum dot. In some embodiment, the conditions sufficient to couple the target-binding sequence to at least a portion of the target in (b) comprises an incubation temperature, wherein the incubation temperature is about 20-45°C. In some embodiment, the conditions sufficient to couple the target-binding sequence to at least a portion of the target in (b) comprises a salt. In some embodiment, the polypeptide comprises a protein. In some embodiment, the protein is a cell-surface protein. In some embodiment, the protein is an intracellular protein. In some embodiment, the protein is a signaling protein. In some embodiment, the polypeptide comprises a protein complex. In some embodiment, the polypeptide binding moiety comprises an antibody or antibody fragment. In some embodiment, the polypeptide binding moiety comprises an aptamer. In some embodiment, the polypeptide binding moiety comprises a nanobody. In some embodiment, the oligonucleotide comprises DNA. In some embodiment, the oligonucleotide is 4-

50 nucleotides in length. In some embodiment, the polypeptide binding moiety is covalently attached to the oligonucleotide. In some embodiment, the polypeptide binding moiety is non-covalently attached to the oligonucleotide. In some embodiment, the probe comprises nucleic acid. In some embodiment, the nucleic acid comprises DNA. In some embodiment, the nucleic acid comprises RNA. In some embodiment, the probe is a padlock probe. In some embodiment, before (d), ligating a first end of the probe to a second end of the probe using a ligase. In some embodiment, the ligase is a T4 ligase. In some embodiment, the ligase is a SplintR ligase. In some embodiment, performing the amplification reaction in (d) comprises performing a rolling circle amplification reaction using a polymerizing enzyme. In some embodiment, the polymerizing enzyme is a phi29 polymerase. In some embodiment, the cell or the tissue is permeabilized using a solvent. In some embodiment, the solvent is methanol. In some embodiment, the solvent is acetone. In some embodiment, the cell or the tissue is permeabilized using a detergent. In some embodiment, the detergent is saponin. In some embodiment, the detergent is Triton X-100.

[0006] Aspects disclosed herein provide methods for identifying a DNA *in situ*, the method comprising: (a) providing a cell or a tissue that has been permeabilized, wherein the cell or the tissue comprises the DNA, where the DNA is not cDNA; (b) directing a probe to the cell or the tissue, wherein the probe comprises a (i) DNA-binding sequence and (ii) an index sequence, under conditions sufficient to couple the DNA-binding sequence to at least a portion of the DNA; (c) amplifying the index sequence to generate one or more amplicons, wherein the one or more amplicons comprises multiple copies of the index sequence or derivative thereof; (d) directing a primer nucleic acid sequence to the cell or the tissue under conditions sufficient to bind the primer nucleic acid sequence to at least a portion of the one or more amplicons to produce a primed index sequence; (e) directing the cell or the tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or the derivatives thereof, under conditions suitable to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (f) detecting the binding complex within the cell or tissue, thereby identifying the DNA. In some embodiments, the method further comprises providing the cell in (a). In some embodiments, the cell is a cancer cell. In some embodiments, the method further comprises providing the tissue in (a). In some embodiments, the tissue is a formalin-fixed paraffin embedded tissue. In some embodiments, the tissue is a fresh-frozen tissue. In some embodiments, the cell or the tissue is extracted from a subject. In some embodiments, the subject is a human subject. In some embodiments, the index sequence comprises a nucleic acid. In some embodiments, the index sequence is 4-50 nucleotides in length. In some embodiments, at least a portion of the index

sequence or derivative thereof recognizes the primer nucleic acid sequence. In some embodiments, the primer nucleic acid sequence comprises a blocking group. In some embodiments, the blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence. In some embodiments, the blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group. In some embodiments, the primer nucleic acid sequence does not include a 3' hydroxyl group. In some embodiments, the primer nucleic acid sequence binds to at least a portion of the index sequence. In some embodiments, the primer nucleic acid sequence binds to at least a portion of a reverse complement of the index sequence. In some embodiments, the primed index sequence comprises a double-stranded nucleic acid. In some embodiments, the primed index sequence comprises a single-stranded nucleic acid. In some embodiments, the cell or the tissue is immobilized on a surface of a substrate. In some embodiments, the surface of the substrate has a water contact angle of less than or equal to 45 degrees. In some embodiments, the surface of the substrate is an interior surface of a flow cell. In some embodiments, the interior surface of the flow cell comprises one or more hydrophilic polymer layers. In some embodiments, the one or more hydrophilic polymer layers comprises a branched polymer. In some embodiments, the one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol. In some embodiments, the nucleotide conjugate comprises a common core. In some embodiments, the one or more nucleotide moieties are coupled to the common core. In some embodiments, the common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot. In some embodiments, the polymer comprises a protein. In some embodiments, wherein in (e) the nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of the plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other. In some embodiments, the different types of the nucleotide moieties comprise at least three different types of the nucleotide moieties. In some embodiments, the one or more nucleotide moieties do not comprise a blocking group coupled thereto. In some embodiments, the detecting in (f) comprises imaging the cell or the tissue using one or more image sensors. In some embodiments, the one or more image sensors comprises a photodetector array. In some embodiments, the method further comprises washing the cell or the tissue following (f) to remove the nucleotide conjugate from the cell or the tissue. In some embodiments, the method further comprises repeating (e) and (f) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in the primed index sequence or derivative thereof. In some embodiments, the detectable label comprises a fluorophore. In some embodiments, the detectable label comprises a quantum dot. In some embodiments, the conditions sufficient to couple the target-binding sequence to at least a

portion of the target in (b) comprises an incubation temperature, wherein the incubation temperature is about 20-45°C. In some embodiments, the conditions sufficient to couple the target-binding sequence to at least a portion of the target in (b) comprises a salt. In some embodiments, the DNA comprises a modification. In some embodiments, the modification is a methyl modification. In some embodiments, the modification is a phosphoryl modification. In some embodiments, the probe comprises nucleic acid. In some embodiments, the nucleic acid comprises DNA. In some embodiments, the nucleic acid comprises RNA. In some embodiments, the DNA-binding sequence comprises nucleic acid. In some embodiments, the nucleic acid comprises DNA. In some embodiments, the nucleic acid comprises RNA. In some embodiments, the probe is a padlock probe. In some embodiments, the method further comprises before (c), ligating a first end of the probe In some embodiments, the ligase is a T4 ligase. In some embodiments, the ligase is a SplintR ligase. In some embodiments, performing the amplification reaction in (c) comprises performing a rolling circle amplification reaction using a polymerizing enzyme. In some embodiments, the polymerizing enzyme is a phi29 polymerase. In some embodiments, the cell or the tissue is permeabilized using a solvent. In some embodiments, the solvent is methanol. In some embodiments, the solvent is acetone. In some embodiments, the cell or the tissue is permeabilized using a detergent. In some embodiments, the detergent is saponin. In some embodiments, the detergent is Triton X-100.

[0007] Aspects disclosed herein provide methods for identifying a ribonucleic acid in situ, the method comprising: (a) providing a cell or tissue that has been permeabilized, wherein the cell or the tissue comprises the ribonucleic acid; (b) directing a nucleic acid molecule to the cell or the tissue, wherein the nucleic acid molecule comprises a target-binding sequence, under conditions sufficient to couple the target-binding sequence to at least a portion of the ribonucleic acid, wherein the nucleic acid molecule further comprises an index sequence; (c) directing a primer nucleic acid sequence to the cell or the tissue under conditions sufficient to bind the primer nucleic acid sequence to at least a portion of the one or more amplicons to produce a primed index sequence; (d) directing the cell or the tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or derivative thereof, under conditions suitable to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; (e) detecting the binding complex within the cell or tissue, thereby identifying the ribonucleic acid. In some embodiments, the method further comprises providing the cell in (a). In some embodiments, the cell is a cancer cell. In some embodiments, the method further comprises providing the tissue in (a). In some embodiments, the tissue is a formalin-fixed paraffin embedded

tissue. In some embodiments, the tissue is a fresh-frozen tissue. In some embodiments, the cell or the tissue is extracted from a subject. In some embodiments, the subject is a human subject. In some embodiments, the ribonucleic acid is a messenger ribonucleic acid. In some embodiments, the ribonucleic acid is a micro ribonucleic acid. In some embodiments, the ribonucleic acid is a ribosomal ribonucleic acid. In some embodiments, the ribonucleic acid is a transfer ribonucleic acid. In some embodiments, the ribonucleic acid is a transfer-messenger ribonucleic acid. In some embodiments, the ribonucleic acid comprises a modification. In some embodiments, the modification comprises a methyl group. In some embodiments, the nucleic acid molecule is a deoxyribonucleic acid. In some embodiments, the nucleic acid molecule is about 10-80 nucleotides in length. In some embodiments, the nucleic acid molecule is about 20-60 nucleotides in length. In some embodiments, the nucleic acid molecule is about 30-50 nucleotides in length. In some embodiments, the nucleic acid molecule comprises a modification. In some embodiments, the modification is a 5' phosphorylation. In some embodiments, the modification is an inter-nucleotide linkage. In some embodiments, the target-binding sequence is about 5-25 nucleotides in length. In some embodiments, the method further comprises introducing an index-binding molecule under conditions sufficient to couple the index-binding molecule to the index sequence. In some embodiments, the index-binding molecule comprises a nucleic acid. In some embodiments, the index-binding molecule is a padlock probe. In some embodiments, the method further comprises ligating the index-binding molecule using a ligase. In some embodiments, the ligase is a T4 ligase. In some embodiments, the ligase is a SplintR ligase. In some embodiments, the method further comprises amplifying the index sequence. In some embodiments, the amplifying comprises performing rolling circle amplification using a polymerizing enzyme. In some embodiments, the polymerizing enzyme is a phi29 polymerase. In some embodiments, the amplifying generates multiple copies of the index sequence or a reverse complement of the index sequence. In some embodiments, the conditions sufficient to couple the target-binding sequence to at least a portion of the ribonucleic acid in (b) comprises a salt. In some embodiments, the conditions sufficient to couple the target-binding sequence to at least a portion of the ribonucleic acid in (b) comprises an incubation temperature, wherein the incubation temperature is about 20-45°C. In some embodiments, the conditions sufficient to couple the target-binding sequence to at least a portion of the ribonucleic acid in (b) comprises a chaotropic reagent. In some embodiments, the index sequence is uniquely associated with the ribonucleic acid. In some embodiments, the index sequence is associated with the ribonucleic acid and at least one other analyte. In some embodiments, the other analyte is a nucleic acid. In some embodiments, the nucleic acid is a ribonucleic acid. In some embodiments, the nucleic acid is a deoxyribonucleic acid. In some embodiments, the other analyte is a polypeptide. In some embodiments, the polypeptide is a

protein. In some embodiments, the index sequence is 4-50 nucleotides in length. In some embodiments, at least a portion of the index sequence recognizes the primer nucleic acid sequence. In some embodiments, the primer nucleic acid sequence comprises a blocking group. In some embodiments, the blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence. In some embodiments, the blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group. In some embodiments, the primer nucleic acid sequence does not include a 3' hydroxyl group. In some embodiments, the primer nucleic acid sequence binds to at least a portion of the index sequence. In some embodiments, the primer nucleic acid sequence binds to at least a portion of the reverse complement of the index sequence. In some embodiments, the primed index sequence comprises a double-stranded nucleic acid. In some embodiments, the primed index sequence comprises a single-stranded nucleic acid. In some embodiments, the cell or the tissue is immobilized on a surface of a substrate. In some embodiments, the surface has a water contact angle of less than or equal to 45 degrees. In some embodiments, the surface is an interior surface of a flow cell. In some embodiments, the interior surface of the flow cell comprises one or more hydrophilic polymer layers. In some embodiments, the one or more hydrophilic polymer layers comprises a branched polymer. In some embodiments, the one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol. In some embodiments, the nucleotide conjugate comprises a common core. In some embodiments, the one or more nucleotide moieties are coupled to the common core. In some embodiments, the common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot. In some embodiments, the polymer comprises a protein. In some embodiments, in (d) the nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of the plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other. In some embodiments, the different types of the nucleotide moieties comprise at least three different types of the nucleotide moieties. In some embodiments, the one or more nucleotide moieties do not comprise a blocking group coupled thereto. In some embodiments, the detecting in (e) comprises imaging the cell or the tissue using one or more image sensors. In some embodiments, the one or more image sensors comprises a photodetector array. In some embodiments, the method further comprises washing the cell or the tissue following (e) to remove the nucleotide conjugate from the cell or the tissue. In some embodiments, the method further comprises repeating steps (d) and (e) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in the primed index sequence or derivative thereof. In some embodiments, the detectable label comprises a fluorophore. In some embodiments, the detectable label comprises a quantum dot.

[0008] Aspects disclosed herein provide systems for identifying a RNA *in situ*, the system comprising: a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising any one of the methods described herein. In some embodiments, the system further comprises the RNA.

[0009] Aspects disclosed herein provide systems for identifying a polypeptide *in situ*, the system comprising: a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising any one of the methods described herein. In some embodiments, the system further comprises the polypeptide. In some embodiments, the system further comprises the polypeptide-binding probe.

[0010] Aspects disclosed herein provide systems for identifying a DNA *in situ*, the system comprising: a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising any one of the methods described herein. In some embodiments, the system further comprises the DNA. In some embodiments, the system further comprises a cell. In some embodiments, the cell is a cancer cell. In some embodiments, the system further comprises a tissue. In some embodiments, the tissue is a formalin-fixed paraffin embedded tissue. In some embodiments, the tissue is a fresh-frozen tissue. In some embodiments, the cell or tissue is extracted from a subject. In some embodiments, the subject is a human subject. In some embodiments, the system further comprises the probe. In some embodiments, the probe is a padlock probe. In some embodiments, the padlock probe comprises a nucleic acid. In some embodiments, the system further comprises the primer nucleic acid molecule sequence. In some embodiments, the system further comprises a ligase. In some embodiments, the ligase is a T4 ligase. In some embodiments, the ligase is a SplintR ligase. In some embodiments, the system further comprises a polymerizing enzyme. In some embodiments, the polymerizing enzyme is a phi29 polymerizing enzyme. In some embodiments, the system further comprises a salt. In some embodiments, the salt comprises NaCl, $MgCl_2$, or $CaCl_2$. In some embodiments, the system further comprises a solvent. In some embodiments, the solvent is a polar solvent. In some embodiments, the solvent is methanol. In some embodiments, the solvent is acetone. In some embodiments, the system further comprises a detergent. In some embodiments, the detergent is saponin. In some embodiments, the detergent is Triton X-100. In some embodiments, the system further comprises a substrate. In some embodiments, the substrate is an interior surface of a flow cell. In some embodiments, the one or more computer processors comprises a central processing unit. In some embodiments, the one or more computer processors comprises a graphic processing unit. In some

embodiments, the one or more computer processors comprises a general-purpose processing unit. In some embodiments, the one or more processors comprises a field-programmable gate array. In some embodiments, the system further comprises a fluidic system. In some embodiments, the fluidic system is configured to provide fluid flow control capability for delivering reagents to a sample in a flow cell.

[0011] Aspects disclosed herein provide systems for identifying a ribonucleic acid in situ, the system comprising: a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising any one of the methods disclosed herein. In some embodiments, the system further comprises the cell. In some embodiments, the cell is a cancer cell. In some embodiments, the system further comprises the tissue. In some embodiments, the tissue is a formalin-fixed paraffin embedded tissue. In some embodiments, the tissue is a fresh-frozen tissue. In some embodiments, the cell or the tissue is extracted from a subject. In some embodiments, the subject is a human subject. In some embodiments, the system further comprises the ribonucleic acid. In some embodiments, the system further comprises the nucleic acid molecule. In some embodiments, the system further comprises a primer nucleic acid sequence. In some embodiments, the system further comprises an index-binding molecule. In some embodiments, the index-binding molecule comprises a nucleic acid. In some embodiments, the index-binding molecule is a padlock probe. In some embodiments, the index-binding molecule is configured to bind to the index sequence. In some embodiments, the system further comprises a ligase. In some embodiments, the ligase is a T4 ligase. In some embodiments, the ligase is a SplintR ligase. In some embodiments, the system further comprises a polymerizing enzyme. In some embodiments, the polymerizing enzyme is a phi29 polymerase. In some embodiments, the system further comprises a salt. In some embodiments, the salt comprises NaCl, $MgCl_2$, or $CaCl_2$. In some embodiments, the system further comprises a solvent. In some embodiments, the solvent is a polar solvent. In some embodiments, the system further comprises a chaotropic reagent. In some embodiments, the system further comprises a substrate. In some embodiments, the substrate is an interior surface of a flow cell. In some embodiments, the one or more computer processors comprises a central processing unit. In some embodiments, the one or more computer processors comprises a graphic processing unit. In some embodiments, the one or more computer processors comprises a general-purpose processing unit. In some embodiments, the one or more processors comprises a field-programmable gate array. In some embodiments, the system further comprises a fluidic system. In some embodiments, the fluidic system is configured to provide fluid flow control capability for delivering reagents to a sample in a flow cell.

[0012] Aspects disclosed herein provide kits for identifying a RNA *in situ*, the kit comprising: (a) a probe comprising (i) an index sequence, and (ii) an RNA-binding sequence, wherein the RNA-binding sequence recognizes to at least a portion of the target; (b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein the nucleotide conjugate is configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (b) instructions for identifying the RNA *in situ*, wherein the instructions comprise any one of the methods disclosed herein.

[0013] Aspects disclosed herein provide kits for identifying a polypeptide *in situ*, the kit comprising: (a) a polypeptide-binding probe comprising (i) a polypeptide binding moiety, and (ii) an oligonucleotide coupled thereto; (b) a probe comprising (i) an index sequence, and (ii) an oligonucleotide-binding probe, wherein the oligonucleotide-binding probe recognizes to at least a portion of the oligonucleotide; (c) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein the nucleotide conjugate is configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (d) instructions for identifying the polypeptide *in situ*, wherein the instructions comprise any one of the methods disclosed herein.

[0014] Aspects disclosed herein provide kits for identifying a DNA *in situ*, the kit comprising: (a) a probe comprising (i) an index sequence, and (ii) an RNA-binding sequence, wherein the RNA-binding sequence recognizes to at least a portion of the target; (b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein the nucleotide conjugate is configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (c) instructions for identifying the RNA *in situ*, wherein the instructions comprise any one of the methods disclosed herein. In some embodiments, the kit further comprises a primer nucleic acid sequence configured to bind to at least a portion of the index sequence or derivative thereof to produce the primed index sequence. In some embodiments, the probe is a padlock probe. In some embodiments, the padlock probe comprises a nucleic acid. In some embodiments, the kit further comprises a ligase. In some embodiments, the ligase is a T4 ligase. In some embodiments, the ligase is a SplintR ligase. In some embodiments, the kit further comprises a polymerizing enzyme. In some embodiments, the

polymerizing enzyme is a phi29 polymerase. In some embodiments, the kit further comprises a buffer. In some embodiments, the buffer comprises a salt. In some embodiments, the salt comprises NaCl, $MgCl_2$, or $CaCl_2$. In some embodiments, the buffer comprises a solvent. In some embodiments, the solvent is polar. In some embodiments, the buffer comprises a chaotropic agent. In some embodiments, the kit further comprises a solvent. In some embodiments, the solvent is methanol. In some embodiments, the solvent is acetone. In some embodiments, the kit further comprises a detergent. In some embodiments, the detergent is saponin. In some embodiments, the detergent is Triton X-100. In some embodiments, the kit further comprises a flow cell. In some embodiments, the flow cell comprises a substrate configured to be coupled to the cell or the tissue.

[0015] Aspects disclosed herein provide kits for identifying a ribonucleic acid in situ, the kit comprising: (a) a nucleic acid molecule comprising (i) an index sequence, and (ii) a target-binding sequence, wherein the target-binding sequence is complementary to at least a portion of the ribonucleic acid sequence; (b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence, wherein the nucleotide conjugate is configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (c) instructions for identifying the ribonucleic acid in situ, wherein the instructions comprise any one of the methods disclosed herein. In some embodiments, the kit further comprises a primer nucleic acid sequence configured to bind to at least a portion of the index sequence or derivative thereof to produce a primed index sequence. In some embodiments, the kit further comprises an index-binding molecule. In some embodiments, the index-binding molecule comprises a nucleic acid. In some embodiments, the index-binding molecule is a padlock probe. In some embodiments, the index-binding molecule is configured to bind to the index sequence. In some embodiments, the kit further comprises a ligase. In some embodiments, the ligase is a T4 ligase. In some embodiments, the ligase is a SplintR ligase. In some embodiments, the kit further comprises a polymerizing enzyme. In some embodiments, the polymerizing enzyme is a phi29 polymerase. In some embodiments, the kit further comprises a buffer. In some embodiments, the buffer comprises a salt. In some embodiments, the salt comprises NaCl, $MgCl_2$, or $CaCl_2$. In some embodiments, the buffer comprises a solvent. In some embodiments, the solvent is polar. In some embodiments, the buffer comprises a chaotropic agent. In some embodiments, the kit further comprises a flow cell. In some embodiments, the flow cell comprises a substrate configured to be coupled to the cell or the tissue.

[0016] Aspects disclosed herein provide methods for *in situ* sequencing, the methods comprising a) providing a biological sample deposited on a solid support, wherein the biological

sample harbors a plurality of RNA including at least a first target RNA; b) providing a plurality of padlock probes wherein individual padlock probes in the plurality comprise a linear oligonucleotide comprising a first binding arm at one end and a second binding arm at the other end and a universal sequence region between the first and second binding arms, wherein the first and second binding arms of a padlock probe can hybridize to portions of a target RNA molecule to form an RNA-padlock probe complex, wherein individual padlock probes comprises any one or any combination of two or more sequences including: (i) a universal binding site for a sequencing primer; (ii) a universal binding site for an amplification primer; (iii) a universal binding site for a compaction oligonucleotide; and/or (iv) an optional target barcode sequence which uniquely identifies a target RNA; c) contacting the biological sample with the plurality of padlock probes wherein the contacting is conducted under a condition suitable for moving the plurality of the padlock probes into the biological sample and suitable for hybridizing individual padlock probes to at least a portion of their cognate target RNA inside the biological sample thereby forming a plurality of RNA-padlock probe complexes, wherein the first and second binding arms of individual padlock probes hybridize to proximal positions on a target RNA molecule to form a circularized padlock probe having a nick or gap between the hybridized first and second binding arms; d) contacting the biological sample with a padlock probe closing reagent, wherein the contacting is conducted under a condition suitable for moving the padlock probe closing reagent into the biological sample, and the condition is suitable for generating a plurality of covalently closed padlock probes; e) contacting the biological sample with a rolling circle amplification reagent, wherein the contacting is conducted under a condition suitable for moving the rolling circle amplification reagent into the biological sample; f) conducting rolling circle amplification inside the biological sample thereby generating a plurality of concatemers inside the biological sample, wherein the rolling circle amplification reaction is conducted in the presence of a plurality of compaction oligonucleotides, wherein individual compaction oligonucleotides comprise a 5′ region that binds a first portion of a concatemer molecule and a 3′ region that binds a second portion of the same concatemer molecule thereby pulling together distal portions of the concatemer molecule causing compaction of the concatemer molecule to form a DNA nanoball; and g)sequencing the plurality of concatemer molecules inside the biological sample by contacting the plurality of concatemers inside the biological sample with (i) a plurality of universal sequencing primers, (ii) a plurality of sequencing polymerases, and (iii) a plurality of detectably labeled multivalent molecules, under a condition suitable for hybridizing the plurality of universal sequencing primers to their respective universal sequencing primer binding sites on the concatemers and conducting a plurality of polymerase-catalyzed sequencing cycles thereby generating a plurality of sequencing read products inside the biological sample, wherein individual

detectably labeled multivalent molecules comprise a core attached to a plurality of nucleotide-arms, wherein the nucleotide-arms are attached to a nucleotide moiety.

[0017] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0018] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] The novel features of the inventive concepts are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present inventive concepts will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the inventive concepts are utilized, and the accompanying drawings of which:

[0020] FIGS. 1A-1H illustrate the steps utilizing a non-limiting examples of multivalent binding composition (e.g. nucleotide conjugate) for sequencing a target nucleic acid: **FIG. 1A** illustrates a non-limiting example 4-of attaching target nucleic acid to a surface; **FIG. 1B** illustrates clonally the target nucleic acid to form clusters of amplified target nucleic acid molecules; **FIG. 1C** illustrates a non-limiting example of priming the target nucleic acid to produce a primed target nucleic acid; **FIG. 1D** illustrates a non-limiting example of contacting the primed target nucleic acid to the binding complex and polymerase to form a binding complex; **FIG. 1E** illustrates a non-limiting example of the images of the binding complex captured on the surface; **FIG. 1F** illustrates a non-limiting example of extending the primer strand by one nucleotide; **FIG. 1G** illustrates a non-limiting example of another cycle of contacting the primed target nucleic acid to the binding complex and polymerase to form a binding complex; and **FIG.**

1H illustrates non-limiting examples of the images of binding complex captured on the surface in subsequent sequencing cycles.

[0021] FIG. 2A-2B illustrate a non-limiting example of detecting target nucleic acid using the nucleotide conjugates (e.g. detectable nucleotide conjugates). **FIG. 2A** shows the step of contacting the polymerase and nucleotide conjugates (e.g. detectable nucleotide conjugates) to some nucleic acid molecules; **FIG 2B** shows the binding complex formed between the polymerase, nucleotide conjugates (e.g. detectable nucleotide conjugates), and the target nucleic acid molecules.

[0022] FIG. 3 shows a flow chart outlining the steps for sequencing a target nucleic acid and extending the primer strand through a single base addition.

[0023] FIG. 4 shows a flow chart outlining the steps for sequencing a target nucleic acid and extending the primer strand through incorporating the nucleotide on the particle-nucleotide conjugate.

[0024] FIG. 5A-5C shows schematic representations of non-limiting examples of varying configurations of the nucleotide conjugates (e.g. detectable nucleotide conjugates): **FIG. 5A** shows nucleotide conjugates (e.g. detectable nucleotide conjugates) having various multi-arm configurations; **FIG. 5B** shows a polymer-nucleotide conjugate having the polymer branch radiating from the center; and **FIG. 5C** shows nucleotide conjugates (e.g. detectable nucleotide conjugates) having the binding moiety biotin.

[0025] FIG. 6 shows a generalized graphical depiction of the increase in signal intensity that has been observed during binding, persistence, and washing and removal of multivalent substrates.

[0026] FIGS. 7A-7G show fluorescence images of multivalent polyethylene glycol (PEG) polymer-nucleotide (base-labeled) conjugates, having an effective nucleotide concentration of 500 nanomolar (nM) and varying PEG branch length, after contacting to a support surface comprising DNA templates (comprising G or A as the first base; prepared using rolling circle amplification (RCA)) in an exposure buffer comprising 20 nM klenow polymerase and 2.5 mM Sr^{+2}. Images were acquired after washing with an imaging buffer having the same composition as the exposure buffer but lacking nucleotides and polymerase. Panels show images obtained using multivalent PEG-nucleotide ligands with arm lengths as follows. **FIG. 7A**: 1K PEG. **FIG. 7B**: 2K PEG. **FIG. 7C**: 3K PEG. **FIG. 7D**: 5K PEG. **FIG. 7E**: 10K PEG. **FIG. 7F**: 20K PEG. **FIG. 7G** shows images obtained using 10K PEG and an inactive klenow polymerase comprising the mutation D882H. FIG. 7H shows images obtained using 10K PEG and an inactive klenow polymerase comprising the mutation D882E. FIG. 7I shows images obtained using 10K PEG and

an inactive klenow polymerase comprising the mutation D882A. FIG. 7G shows images obtained using 10K PEG and an active wild type klenow polymerase.

[0027] FIG. 8 shows a quantitative representation of the fluorescence intensities in the images shown in **FIGS. 7A-7F,** separated by color value, with orange trace corresponding to the red label (Cy3 label; A bases) and blue trace corresponding to the green label (Cy5 label; G bases).

[0028] FIGS. 9A-9J show fluorescence images of the steps in a sequencing reaction using multivalent PEG-substrate compositions. **FIG. 9A**. Red and green fluorescent images post exposure of DNA RCA templates (G and A first base) to 500 nM base labeled nucleotides (A-Cy3 and G-Cy5) in exposure buffer containing 20 nM Klenow polymerase and 2.5 mM Sr^{+2}. Images were collected after washing with imaging buffer with the same composition as the exposure buffer but containing no nucleotides or polymerase. Contrast was scaled to maximize visualization of the dimmest signals, but no signals persisted following washing with imaging buffer (a. inset). **FIGS. 9B-9E**: fluorescence images showing multivalent PEG-nucleotide (base-labeled) ligands PB1 (**FIG. 9B**), PB2 (**FIG. 9C**), PB3 (**FIG. 9D**), and PB5 (**FIG. 9E**) having an effective nucleotide concentration of 500 nM after mixing in the exposure buffer and imaging in the imaging buffer as described above. **FIG. 9F**: fluorescence image showing multivalent PEG-nucleotide (base-labeled) ligand PB5 at 2.5uM after mixing in the exposure buffer and imaging in the imaging buffer as above. **FIGS. 9G-9I**. Fluorescence images showing further base discrimination by exposure of the binding complex to inactive mutants of klenow polymerase (**FIG. 9G**. D882H; **FIG. 9H**. D882E; **FIG. 9I**. D882A) vs. the wild type Klenow (control) enzyme (**FIG. 9J**).

[0029] FIGS. 10A-10B show the efficacy of the multivalent reporter compositions in determining the base sequence of a DNA sequence over 5 sequencing cycles: **FIG. 10A** shows images and expected sequences for templates taken after each sequencing cycle; and **FIG. 10B** shows aligned sequencing results utilizing the images taken in **FIG. 10A**.

[0030] FIG. 11 shows normalized fluorescence from multivalent substrates bound to DNA clusters as in FIG. 9, with the substrate complexes formed in the presence (condition B) and absence (condition A) of Triton-X100 (0.016%).

[0031] FIG. 12 shows normalized fluorescence of multivalent substrates and free nucleotides. (Top) Two replicates of a multivalent substrate bound to DNA clusters (Conditions A and B) vs. binding complexes formed using labeled free nucleotides (Condition C) after 1 minute; (Bottom) Time course of fluorescence from multivalent substrate complexes over the course of 60 min.

[0032] FIG. 13 shows a non-limiting example of a computing device; in this case, a device with one or more processors, memory, storage, and a network interface.

[0033] FIG. 14 shows a non-limiting example of a web/mobile application provision system; in this case, a system providing browser-based and/or native mobile user interfaces.

[0034] FIG. 15 shows a non-limiting example of a cloud-based web/mobile application provision system; in this case, a system comprising an elastically load balanced, auto-scaling web server and application server resources as well synchronously replicated databases.

[0035] FIG. 16 shows a schematic of a multivalent molecule comprising a core attached to a plurality of nucleotide-arms

[0036] FIG. 17 shows a schematic of a multivalent molecule comprising a dendrimer core attached to a plurality of nucleotide arms.

[0037] FIG. 18 shows a schematic of a multivalent molecule comprising a core attached to a plurality of nucleotide arms, where the nucleotide arms comprise a biotin, a spacer, a linker, and a nucleotide moiety.

[0038] FIG. 19 shows a schematic of a nucleotide-arm comprising a core attachment moiety, a spacer, a linker, and a nucleotide moiety.

[0039] FIG. 20 shows the chemical structure for a spacer (top), and various linkers.

[0040] FIG. 21 shows chemical structures of various linkers.

[0041] FIG. 22 shows chemical structures of various linker attached to nucleotide moieties.

[0042] FIG. 23 shows chemical structures of various linker attached to nucleotide moieties.

[0043] FIG. 24 shows chemical structures of various linker attached to nucleotide moieties.

[0044] FIG. 25 shows the chemical structure of a biotinylated nucleotide arm. In this example, the nucleotide moiety 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.

[0045] FIG. 26 shows a schematic of a guanine tetrad (e.g. G-tetrad).

[0046] FIG. 27 shows a schematic of an intramolecular G-quadruplex structure.

[0047] FIG. 28 shows a schematic of a workflow for generating circularized padlock probes, comprising hybridizing first and second padlock probes to the first and second target RNA molecules (respectively) to generate first and second circularized padlock probes (respectively).

[0048] FIG. 29 shows schematic of a workflow that is a continuation of the schematic shown in FIG. 28.

[0049] FIG. 30 shows a schematic of a workflow for generating circularized padlock probes, comprising hybridizing first and second padlock probes to the first and second target RNA molecules (respectively) to generate first and second circularized padlock probes (respectively).

[0050] FIG. 31 shows a schematic of a workflow that is a continuation of the workflow shown in FIG. 30.

[0051] FIG. 32 shows a schematic of a workflow for generating circularized padlock probes, comprising hybridizing first and second padlock probes to the first and second target RNA molecules (respectively) to generate first and second circularized padlock probes (respectively).

[0052] FIG. 33 shows a schematic of a workflow that is a continuation of the workflow shown in FIG. 32.

[0053] FIG. 34 shows three graphs of data from a sequencing run wherein padlock probes were hybridized to RNA inside a biological sample, subjected to gap fill-in and ligation to generate circular nucleic acids which were subjected to rolling circle amplification to generate concatemer molecules inside the biological sample. The concatemers were subjected to a two-stage sequencing workflow for 100 sequencing cycles. The top graph shows intensity measurements. The middle graph shows background measurements. The bottom graph shows phasing and pre-phasing measurements.

[0054] FIG. 35 shows a schematic of a workflow for sequencing a concatemer that is generated inside a cell.

[0055] FIG. 36 shows a schematic of a workflow for sequencing a concatemer that is generated inside the cell.

[0056] FIG. 37 shows a schematic of a workflow for sequencing a concatemer that is generated inside the cell.

[0057] FIG. 38 shows a schematic of a workflow for sequencing a concatemer that is generated inside the cell.

[0058] FIG. 39 shows a schematic of a workflow for detecting a target *in situ* involving direct detection of the target and formation of a binding complex.

[0059] FIG. 40 shows a schematic of a workflow for detecting a target *in situ* involving direct detection of the target, performing an amplification reaction to form an amplicon, and formation of a binding complex with the amplicon.

[0060] FIG. 41 shows a schematic of a workflow for detecting an RNA *in situ* involving direct detection of the RNA, performing an amplification reaction to form an amplicon, and formation of a binding complex with the amplicon.

[0061] FIG. 42 shows a schematic of a workflow for detecting a DNA *in situ* involving direct detection of the DNA, performing an amplification reaction to form an amplicon, and formation of a binding complex with the amplicon.

[0062] FIG. 43 shows a schematic of a workflow for detecting a polypeptide *in situ* involving direct detection of the polypeptide, performing an amplification reaction to form an amplicon, and formation of a binding complex with the amplicon.

DETAILED DESCRIPTION

[0063] Provided herein are kits, systems, and methods related to detecting a target nucleic acid *in situ* in a biological sample. The present disclosure provides novel solutions and approaches to addressing many of the shortcomings of existing technologies. Multiple components may be used to facilitate the detection of the target nucleic acid, including a nucleic acid molecule, a detectable nucleotide conjugate, and a primer nucleic acid sequence. The nucleic acid molecule may comprise an index sequence that corresponds with the target nucleic acid. The detecting may comprise binding the nucleic acid molecule to the target nucleic acid and contacting the nucleic acid molecule with a primer nucleic acid sequence to form a primed index sequence. The nucleotide conjugate (e.g. detectable nucleotide conjugate)may be added to the sample and a binding complex may be formed between at least two copies of the primed index sequence and the detectable nucleotide conjugate. The binding complex may be detected to reveal the identity of the terminal nucleotide of the primed index sequence. This process may be repeated to reveal the nucleotide identities of each of the nucleotides within the index sequence, thereby identifying the target nucleic acid.

[0064] The methods described herein may relate to detecting one or more binding complexes as described above and thereby identifying the target nucleic acid. The systems described herein may comprise a computing device comprising a computer program with instructions to perform any one of the methods described herein. The computing device may comprise a number of components, including but not limited to one or more processors, an operating system, and a memory. The kits described herein may be used for performing any one of the methods described herein. The kits may include a nucleic acid molecule, a detectable nucleotide conjugated, and instructions for use.

DEFINITIONS

[0065] Unless defined otherwise, all terms of art, notations and other technical and scientific terms or terminology used herein are intended to have the same meaning as is commonly understood by one of ordinary skill in the art to which the claimed subject matter pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art.

[0066] Throughout this application, various embodiments may be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the disclosure. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range.

[0067] As used in the specification and claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, the term "a sample" includes a plurality of samples, including mixtures thereof.

[0068] As used herein, the term "about" a number refers to that number plus or minus 10% of that number. The term "about" a range refers to that range minus 10% of its lowest value and plus 10% of its greatest value.

[0069] As used herein, a "detectable nucleotide conjugate" refers to a composition comprising a plurality of nucleotides and a detectable label attached to a common core. The one or more nucleotides can include unlabeled and labeled nucleotides. The nucleotides can include a blocking group. The detectable label can be a fluorophore. The plurality of nucleotides and the detectable label can be attached to the common core covalently or non-covalently.

[0070] The terms "determining," "measuring," "evaluating," "assessing," "assaying," and "analyzing" are often used interchangeably herein to refer to forms of measurement. The terms include determining if an element is present or not (for example, detection). These terms can include quantitative, qualitative or quantitative and qualitative determinations. Assessing can be relative or absolute. "Detecting the presence of" can include determining the amount of something present in addition to determining whether it is present or absent depending on the context.

[0071] As used herein, "nucleic acid" (also referred to as a "polynucleotide" or "oligonucleotide") is a polymer of two or more nucleotides joined by covalent internucleosidic linkages, or variants or functional fragments thereof. In naturally occurring examples of nucleic acids, the internucleoside linkage is typically a phosphodiester bond. However, other examples optionally comprise other internucleoside linkages, such as phosphorothiolate linkages and may or may not comprise a phosphate group. Nucleic acids include double- and single-stranded deoxyribonucleic acid (DNA), as well as double- and single-stranded ribonucleic acid (RNA), DNA/RNA hybrids, peptide-nucleic acids (PNAs), hybrids between PNAs and DNA or RNA, and may also include other types of nucleic acid modifications.

[0072] As used herein, a "nucleotide" refers to a nucleotide, nucleoside, or analog thereof. The nucleotide refers to both naturally occurring and chemically modified nucleotides and can include but are not limited to a nucleoside, a ribonucleotide, a deoxyribonucleotide, a protein-nucleic acid residue, or derivatives. Examples of the nucleotide includes an adenine, a thymine, a uracil, a cytosine, a guanine, or residue thereof; a deoxyadenine, a deoxythymine, a deoxyuracil, a deoxycytosine, a deoxyguanine, or residue thereof; a adenine PNA, a thymine PNA, a uracil PNA, a cytosine PNA, a guanine PNA, or residue or equivalents thereof, an N- or C-glycoside of a purine or pyrimidine base (*e.g.*, a deoxyribonucleoside containing 2-deoxy-D-ribose or ribonucleoside containing D-ribose).

[0073] "Complementary" as used herein, refers to the topological compatibility or matching together of interacting surfaces of a ligand molecule and its receptor. Thus, the receptor and its ligand can be described as complementary, and furthermore, the contact surface characteristics are complementary to each other. With respect to nucleic acid sequences, complementary refers the sequence of a nucleic acid having the propensity to hybridize to another nucleic acid sequence by virtue of the nucleotides that makeup the nucleic acid sequence. Nucleic acids bind to each other through the formation of non-covalent interactions between nucleotides based on base pairing interactions between adenine (A) and thymine (T) and between guanine (G) and cytosine (C). When two nucleic acids are aligned antiparallel to each other, the nucleotide bases at each position in the sequences will be complementary.

[0074] "Branched polymer", as used herein, refers to a polymer having a plurality of functional groups that help conjugate a biologically active molecule such as a nucleotide, and the functional group can be either on the side chain of the polymer or directly attaches to a central core or central backbone of the polymer. The branched polymer can have linear backbone with one or more functional groups coming off the backbone for conjugation. The branched polymer can also be a polymer having one or more sidechains, wherein the side chain has a site suitable for conjugation. Examples of the functional group include but are limited to hydroxyl, ester, amine, carbonate, acetal, aldehyde, aldehyde hydrate, alkenyl, acrylate, methacrylate, acrylamide, active sulfone, hydrazide, thiol, alkanoic acid, acid halide, isocyanate, isothiocyanate, maleimide, vinylsulfone, dithiopyridine, vinylpyridine, iodoacetamide, epoxide, glyoxal, dione, mesylate, tosylate, and tresylate.

[0075] "Polymerizing enzyme" as used herein, refers to an enzyme that contains a nucleotide binding moiety and helps formation of a binding complex between a target 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. The

polymerase can include catalytically inactive polymerase, catalytically active polymerase, reverse transcriptase, and other enzymes containing a nucleotide binding moiety.

[0076] "Persistence time," as used herein, refers to the length of time that a binding complex remains stable without any binding component dissociates from the binding complex. The binding complex may comprise a target nucleic acid, a polymerase, and a detectable nucleotide conjugate. The binding complex may comprise the target nucleic acid, a polymerase and nucleotide (e.g., a detectable nucleotide). 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 example label is a fluorescent label.

[0077] 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.

[0078] 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).

[0079] 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.

[0080] 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., ±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.

[0081] 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.

[0082] 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 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).

[0083] 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 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 comprise non-natural internucleoside linkages, including phosphorothioate, phosphorothiolate, or peptide nucleic acid (PNA) linkages. In some embodiments, nucleic acids comprise a one type of polynucleotides or a mixture of two or more different types of polynucleotides.

[0084] The term "operably linked" and "operably joined" or related terms as used herein refers to juxtaposition of components. The juxtaposition 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.

[0085] 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 are 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).

[0086] The term "primer" and related terms 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 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).

[0087] 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 describe 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.

[0088] 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. In some embodiments, a barcode sequence uniquely identifies a target RNA. 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.

[0089] In some embodiments, any of the amplification primer sequences, sequencing primer sequences, capture primer sequences, target capture sequences, circularization anchor sequences, sample barcode sequences, target 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.

[0090] 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).

[0091] 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 Hoogstein 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.

[0092] 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.

[0093] 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.

[0094] 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 Hoogstein hydrogen bonds with an appropriate complementary base. Example bases include, but are not limited to, purines and pyrimidines such as: 2-aminopurine, 2,6-diaminopurine, adenine (A), ethenoadenine, N^6-Δ^2-isopentenyladenine (6iA), N^6-Δ^2-isopentenyl-2-methylthioadenine (2ms6iA), N^6-methyladenine, guanine (G), isoguanine, N^2-dimethylguanine (dmG), 7-methylguanine (7mG), 2-thiopyrimidine, 6-thioguanine (6sG), hypoxanthine and O^6-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, O^4-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-methycytosines; base (Y); as well as methylated, glycosylated, and acylated base moieties; and the like. Additional example bases can be found in Fasman, 1989, in "Practical Handbook of Biochemistry and Molecular Biology", pp. 385-394, CRC Press, Boca Raton, Fla.

[0095] 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.

[0096] 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 phosphoramide 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 BH_3. In some embodiments, the chain includes phosphate groups substituted with analogs including phosphoramidate, phosphorothioate, phosphordithioate, and O-methylphosphoroamidite groups.

[0097] 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).

[0098] A reporter moiety (or label) comprises a fluorescent label or a fluorophore. Example 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, carboxynapthofluorescein, 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.

[0099] In some embodiments, the reporter moiety can be a FRET pair, 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.

[00100] 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.

[00101] 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.

[00102] 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.

[00103] 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.

[00104] 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.

[00105] 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.

[00106] The support can have a plurality (e.g., two or more) of nucleic acid templates immobilized thereon. The plurality of immobilized nucleic acid templates have the same sequence or have 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.

[00107] 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) 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. 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.

[00108] 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. The location of the randomly located sites on the support are not pre-determined. 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) 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. 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.

[00109] In some embodiment, 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.

[00110] In some embodiment, 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.

[00111] 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.

[00112] 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 includes nucleic acid concatemers (e.g., nucleic acid clusters).

[00113] 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.

[00114] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

METHODS

[00115] Provided herein are methods related to detecting one or more targets (e.g. RNAs, DNAs, polypeptides) *in situ* in a biological sample. The methods described herein do not comprise cDNA synthesis. A nucleic acid molecule containing a target-binding sequence and an index sequence may be introduced to the biological sample such that it couples to the target (e.g. the RNA, the DNA, or the polypeptide). A primer nucleic acid sequence may be added to the sample and binds to the index sequence to form a primed index sequence. A nucleotide conjugate (e.g. a detectable nucleotide conjugate), which includes a detectable label and one or more nucleotide moieties may be added to the sample to form a binding complex. The binding complex

may be detected, which corresponds to the sequence associated with the index sequence and thereby the identity of the target (e.g. the RNA, the DNA, or the polypeptide).

[00116] The binding complex involved in the methods disclosed herein can include at least one nucleotide conjugate (e.g. detectable nucleotide conjugate), which may include a plurality of copies of the same nucleotide conjugated thereto. When the nucleotide is complementary to the index sequence, the nucleotide conjugate may form a binding complex with a polymerase and the primed index sequence, and the binding complex exhibits increased stability and longer persistence time than the binding complex formed using a single unconjugated or untethered nucleotide.

[00117] The methods described herein offer several advantages over other *in situ* workflows. For example, the biological sample may be placed on a solid support, for example a planar support comprising glass or plastic which can be fabricated into any shape and size. Assembly of a hybridization chamber on the support may not be needed. Preparation of chemically-washed glass beads for cell adherence may also not be needed. The support can be passivated with a coating that promotes cell adhesion to the support. The coating does not need to be formulated to include tethered capture primers. The support, having a cell sample adhered thereon, can be easily adapted to fit into an existing flow cell holder/cradle which is fluidically connected to an automated fluid dispensing system and configured on a fluorescent microscope. Any combination of the steps for conducting *in situ* sequencing can be performed in an automated mode using the fluid dispensing system, including cell seeding, cell fixation, cell permeabilization, probe hybridization, probe ligation reaction, amplification, and sequencing.

[00118] Another advantage of the methods described herein, is the formation of amplicons (e.g. concatemers) inside the biological sample. The amplicons may be single-stranded and collapse into compact DNA nanoballs, where each nanoball carries numerous tandem copies of a polynucleotide unit along their lengths, where the polynucleotide unit includes a target sequence-of-interest and at least a universal sequencing primer binding site. Each polynucleotide unit can bind a sequencing primer, a sequencing polymerase and a nucleotide reagent (e.g., detectably labeled multivalent molecules), to form a detectable sequencing complex (e.g., a detectable ternary complex). Each nanoball may carry numerous detectable sequencing complexes. The compact nature of the nanoballs increases the local concentration of detectably-labeled nucleotide reagents that are used during the sequencing workflow which increases the signal intensity emitted from a nanoball to give a discrete detectable signal which can be imaged as a fluorescent spot inside the biological sample. Each spot may correspond to an amplicon and each amplicon corresponds to a target molecule in the biological sample. Multiple spots can be

detected and imaged simultaneously in the biological sample. Alternatively, multiple spots can be detected and imaged in separate batches in the biological sample.

[00119] An aspect of the disclosure provides a method for identifying a target (e.g. an RNA, a DNA, or a polypeptide) *in situ* using the components shown in FIG. 39. The method comprises (a) providing a cell or tissue that has been permeabilized, wherein the cell or tissue comprises the target (3901); (b) directing a nucleic acid molecule (3902) to the cell or tissue (3902), wherein the nucleic acid molecule comprises a target-binding sequence under conditions sufficient to couple the target-binding sequence to at least a portion of the target (e.g. an RNA, or a DNA), wherein the nucleic acid molecule comprises an index sequence; (c) directing (3904) a primer nucleic acid sequence (3905) to the cell or tissue under conditions sufficient to bind the primer nucleic acid sequence (3905) to at least a portion of the one or more amplicons to produce a primed index sequence; (d) directing (3906) the cell of tissue with a nucleotide conjugate (e.g. a detectable nucleotide conjugate) comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or derivative thereof, under conditions suitable to form a binding complex (3907) between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (e) detecting the binding complex within the cell or tissue, thereby identifying the target.

[00120] Another aspect of the disclosure provides a method for An aspect of the disclosure provides a method for identifying a target (e.g. an RNA, a DNA, or a polypeptide) *in situ* using the components shown in FIG. 40. The method comprises (a) providing a cell or tissue that has been permeabilized, wherein the cell or tissue comprises the target (4001); (b) directing a nucleic acid molecule (4003) to the cell or tissue (4002), wherein the nucleic acid molecule comprises a target-binding sequence under conditions sufficient to couple the target-binding sequence to at least a portion of the target (e.g. an RNA, a DNA or a polypeptide), wherein the nucleic acid molecule comprises an index sequence; (c) directing (4004) a probe (e.g. a padlock probe) (4005) to the cell or tissue under conditions sufficient to binding the probe to the nucleic acid molecule; (d) amplifying (4006) (e.g. performing a rolling circle amplification reaction) of the index sequence to generate one or more amplicons, wherein the one or more amplicons comprises multiple copies of the index sequence or derivative thereof to form one or more amplicons (4007); (e) directing (4008) a primer nucleic acid sequence (4009) to the cell or tissue under conditions sufficient to bind the primer nucleic acid sequence (4007) to at least a portion of the one or more amplicons to produce a primed index sequence; (f) directing the cell of tissue with a nucleotide conjugate (e.g. a detectable nucleotide conjugate) comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index

sequence or derivative thereof, under conditions suitable to form a binding complex (4010) between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (g) detecting the binding complex within the cell or tissue, thereby identifying the target.

[00121] Another aspect of the disclosure provides a method for identifying a, RNA *in situ* using the components shown in FIG. 41. The method comprises (a) providing a cell or tissue that has been permeabilized, wherein the cell or tissue comprises the RNA (4101); (b) directing a probe (4103) to the cell or tissue (4102), wherein the probe (4103) comprises an RNA-binding sequence under conditions sufficient to couple the RNA-binding sequence to at least a portion of the RNA, wherein the probe further comprises an index sequence; (c) amplifying (4104) (e.g. performing a rolling circle amplification reaction) of the index sequence to generate one or more amplicons, wherein the one or more amplicons comprises multiple copies of the index sequence or derivative thereof to form one or more amplicons (4105); (d) directing (4106) a primer nucleic acid sequence (4107) to the cell or tissue under conditions sufficient to bind the primer nucleic acid sequence to at least a portion of the one or more amplicons to produce a primed index sequence; (e) directing the cell or tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or the derivative thereof, under conditions suitable to form a binding complex (4108) between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (f) detecting the binding complex within the cell or tissue, thereby identifying the RNA.

[00122] Another aspect of the disclosure provides a method for identifying a, DNA *in situ* using the components shown in FIG. 42. The method comprises (a) providing a cell or tissue that has been permeabilized, wherein the cell or tissue comprises the DNA (4201), wherein the DNA is not cDNA; (b) directing a DNA-binding sequence (4203) to the cell or tissue (4202), wherein the probe comprises an index sequence, under conditions sufficient to couple the DNA-binding sequence to at least a portion of the DNA; (c) amplifying (4204) (e.g. performing a rolling circle amplification reaction) of the index sequence to generate one or more amplicons, wherein the one or more amplicons comprises multiple copies of the index sequence or derivative thereof to form one or more amplicons (4205); (d) directing (4206) a primer nucleic acid sequence (4207) to the cell or tissue under conditions sufficient to bind the primer nucleic acid sequence to at least a portion of the one or more amplicons to produce a primed index sequence; (e) directing the cell or tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or the derivative thereof, under conditions suitable to form a binding complex (4208) between the one or more

nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (f) detecting the binding complex within the cell or tissue, thereby identifying the DNA.

[00123] Another aspect of the disclosure provides a method for identifying a polypeptide (e.g. a protein) *in situ* using the components shown in FIG. 43. The method comprises (a) providing a cell or tissue that has been permeabilized, wherein the cell or tissue comprises the polypeptide (4301); (b) directing a polypeptide-binding probe (4303) to the cell or tissue (4302), wherein the polypeptide-binding probe comprises an polypeptide-binding moiety and an oligonucleotide (4304) coupled thereto, under conditions sufficient to couple the polypeptide-binding probe to at least a portion of the polypeptide; (c) directing (4305) the cell or tissue with a probe (4307) comprising an oligonucleotide binding probe under conditions sufficient to couple the oligonucleotide-binding probe to at least a portion of the oligonucleotide or the polypeptide-binding probe, wherein the probe further comprises an index sequence; (c) amplifying (4307) (e.g. performing a rolling circle amplification reaction) of the index sequence to generate one or more amplicons, wherein the one or more amplicons comprises multiple copies of the index sequence or derivative thereof to form one or more amplicons (4308); (e) directing (4309) a primer nucleic acid sequence (4310) to the cell or tissue under conditions sufficient to bind the primer nucleic acid sequence to at least a portion of the one or more amplicons to produce a primed index sequence; (f) directing the cell or tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence or the derivative thereof, under conditions suitable to form a binding complex (4311) between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (f) detecting the binding complex within the cell or tissue, thereby identifying the DNA.

i. Biological samples

[00124] The methods described herein comprise detecting one or more targets (e.g. RNAs, DNAs, and/or polypeptides) in a biological sample. The biological sample may comprise a tissue or be a tissue. In some embodiments, the biological sample may comprise a cell. In some embodiments, the biological sample may be a cell. The biological sample may be from an organism. The biological sample may be from a cell culture. In some embodiments, the biological sample may be preserved, having been harvested previously. The biological sample may comprise a whole cell, a plurality of whole cells, an intact tissue or an intact tumor. In some embodiments, the biological sample may comprise a fresh biological sample, a freshly-frozen biological sample,

a sectioned biological sample, or an FFPE biological sample. In some embodiments, the biological sample may comprise one or more living cells or non-living cells. In some embodiments, the biological sample comprises an intact cell or tissue. In some embodiments, the biological sample comprises a cell or tissue that is/are not lysed. In some embodiments, the biological sample comprises a single cell, a plurality of cells, a tissue, an organ, an organism, or section of these biological samples.

[00125] In some embodiments, the biological sample is derived from eukaryotes (such as animals, plants, fungi, protista), archaebacteria, or eubacteria. The biological sample may be derived from prokaryotic or eukaryotic cells, such as adherent or non-adherent eukaryotic cells. The biological sample may be derived from a primary or immortalized cell line from a rodent, porcine, feline, canine, bovine, equine, primate, or human cell lines. In some embodiments, the biological sample may be obtained from a virus, fungus, or prokaryote. In some embodiments, the biological sample can be obtained from an animal, insect or plant. In some embodiments, the biological sample comprises one or more virally-infected cells. In some embodiments, the biological 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. In some embodiments, the biological sample can be obtained from an organ including but not limited to 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.

[00126] The biological sample may be a solid sample, such as a tissue biopsy. The biological sample may be a fluid sample, such as blood or a component of blood (e.g., serum or plasma). In some embodiments, the biological sample is obtained from skin, heart, lung, kidney, breath, bone marrow, stool, semen, vaginal fluid, interstitial fluids derived from tumorous tissue, breast, pancreas, cerebral spinal fluid, tissue, throat swab, biopsy, placental fluid, amniotic fluid, liver, muscle, smooth muscle, bladder, gall bladder, colon, intestine, brain, cavity fluids, sputum, pus, micropiota, meconium, breast milk, prostate, esophagus, thyroid, serum, saliva, urine, gastric and digestive fluid, tears, ocular fluids, sweat, mucus, earwax, oil, glandular secretions, spinal fluid, hair, fingernails, skin cells, plasma, nasal swab or nasopharyngeal wash, spinal fluid, cord blood, emphatic fluids, and/or other excretions or body tissues. The biological sample may be a cell-free sample.

[00127] The biological sample may be a tissue. The tissue may be extracted from a subject. In some embodiments, the subject may be a human subject. In some embodiments the tissue may be formalin-fixed paraffin embedded. In some other embodiments, the tissue may be fresh-frozen. The tissue may be sliced into slices, wherein the thickness of the slice may be about

1-30 μm, about 2-29 μm, about 3-28 μm, about 4-27 μm, about 5-26 μm, about 6-25 μm, about 7-24 μm, about 8-23 μm, about 9-22 μm, about 10-21 μm, about 11-20 μm, about 12-19 μm, about 13-18 μm, about 14-17 μm, about 15-16 μm, about 1 μm, about 2 μm, about 3 μm, about 4 μm, about 5 μm, about 6 μm, about 7 μm, about 8 μm, about 9 μm, about 10 μm, about 11 μm, about 12 μm, about 13 μm, about 14 μm, about 15 μm, about 16 μm, about 17 μm, about 18 μm, about 19 μm, about 20 μm, about 21 μm, about 22 μm, about 23 μm, about 24 μm, about 25 μm, about 26 μm, about 27 μm, about 28 μm, about 29 μm, or about 30 μm thick.

[00128] In some embodiments, the biological sample can be grown/cultured on a support and analysis of the cultured biological sample can be conducted on the support. Nucleic acids in the biological sample may not be released from the cells and are not captured on the support. In some cases, the nucleic acids in the biological sample may be retained in the cells and sequenced in the cells (e.g., *in situ* sequencing). In some embodiments, the nucleic acids in the biological sample may be subjected to an optional library prep workflow by appending at least one adaptor sequence, for example a universal adaptor sequence. The nucleic acids in the biological sample may be amplified in the cells by conducting rolling circle amplification.

[00129] The biological sample may be processed and/or analyzed in a flow cells. For example, flow cell, having cultured cells attached thereon, can be placed on a sequencing apparatus 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. The sequencing reactions can be conducted with soluble sequencing primers. The sequencing primers can be designed to hybridize to a target sequence or a universal adaptor sequence. The sequencing primers can include or lack a sample index sequence (e.g., sample barcode sequence). The nucleic acids in the cultured cells can be sequenced by employing any of the nucleotide conjugates (e.g. detectable nucleotide conjugates) described herein. In some embodiments, the sequencing reactions can employ labeled or non-labeled nucleotide analogs. 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 mm^2 and the cycle time for scanning large area (> 10 mm^2) can be less than 5 minutes.

[00130] In some embodiments, the biological sample can be placed on a support. Provided herein are solid supports (e.g., low non-specific binding supports) further comprising a low non-specific binding surface

[00131] The biological sample may comprise cells. The cells described herein may be white blood cells, red blood cells, platelets, epithelial cells, endothelial cells, neurons, glial cells, astrocytes, fibroblasts, skeletal muscle cells, smooth muscle cells, gametes, or cells from the heart, lungs, brain, liver, kidney, spleen, pancreas, thymus, bladder, stomach, colon, or small

intestine. The cells may be normal or healthy cells. Alternately or in combination, the cells may be diseased cells, such as cancerous cells, or from pathogenic cells that are infecting a host. In some embodiments, the cell belongs to a subset of cells, such as immune cell (e.g., T cells, cytotoxic (killer) T cells, helper T cells, alpha beta T cells, gamma delta T cells, T cell progenitors, B cells, B-cell progenitors, lymphoid stem cells, myeloid progenitor cells, lymphocytes, granulocytes, Natural Killer cells, plasma cells, memory cells, neutrophils, eosinophils, basophils, mast cells, monocytes, dendritic cells, and/or macrophages, or any combination thereof), undifferentiated human stem cells, human stem cells that have been induced to differentiate, or rare cells (e.g., circulating tumor cells (CTCs), circulating epithelial cells, circulating endothelial cells, circulating endometrial cells, bone marrow cells, progenitor cells, foam cells, mesenchymal cells, or trophoblasts). Other cells are contemplated and consistent with the disclosure herein.

[00132] The biological sample can be embedded in a wax, resin, epoxy or agar. The biological 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 biological sample can be sectioned or non-sectioned. The biological sample can be stained, de-stained or non-stained.

[00133] In some embodiments, the biological sample may be permeabilized in the presence of at least about 60%, 70%, 80%, 90% or more methanol (or ethanol) and incubated on ice for a period of time. The period of time for incubation can be at least about 10, 15, 20, 25, 30, 35, 40, 50, 60 or more minutes.

[00134] In some embodiments, the methods comprise treating the biological sample with a permeabilization reagent that alters the cell membrane to permit penetration of experimental reagents into the cells. For example, the permeabilization reagent may remove membrane lipids from the cell membrane. In some embodiments, the biological 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, or sodium dodecyl sulfate (SDS), or N-lauroylsarcosine sodium salt solution. In some embodiments, the cross-linking agent comprises paraformaldehyde, a nonionic polyoxyethylene surfactant (e.g., NP40), or a combination thereof. 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. 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.

[00135] The biological sample can be used to generate a three-dimensional polymer matrix comprising the cellular and sub-cellular components (e.g., nucleic acid molecules) of the biological sample. The three-dimensional polymer matrix can be coupled to the support described herein, covalently or non-covalently. In some embodiments, the three-dimensional polymer matrix is porous and comprises polymerized or cross-linked sub-cellular components, including the target nucleic acid molecules. A polymer matrix may be formed within a biological sample (e.g., a cell or tissue) by flowing one or more polymer precursors (e.g., monomers, such as, for example, ethylene oxide for polyethene glycol) into the biological sample and subjecting the one or more polymer precursors to polymerization or cross-linking. Prior to, during, or subsequent to formation of the polymer matrix, positions of moieties (e.g., DNA, RNA, protein) within the biological sample may be fixed, using for example, a fixation agent (e.g., formaldehyde). A porous matrix may be made according to various methods. For example, a polyacrylamide gel matrix can be polymerized with biotinylated DNA molecules and acrydite-modified streptavidin monomers, using a suitable acrylamide:bis-acrylamide ratio to control the cross-linking density. Additional control over the molecular sieve size and density can be achieved by adding additional cross-linkers such as functionalized polyethylene glycols. Enablement for fixing the biological sample to a surface, as well as generating a polymer matrix within a biological sample, is provided in PCT/US2019/055434 (WO 2020/076976; Three Dimensional Spatial Molecular Indexing), which is hereby incorporated by reference in its entirety.

[00136] In any of the methods described herein, the biological sample may be infused with a swellable polyelectrolyte hydrogel (U.S. patent No. 10,309,879 and Chen 2015 Science 347:543, the contents of these documents are incorporated by reference in their entireties). In some embodiments, a fixed and permeabilized biological sample can be infused with sodium acrylate, acrylamide and a cross-linker N-N'-methylenebisacrylamide. In some embodiments, ammonium persulfate (APS) initiator and tetramethylethylenediamine (TEMED) accelerator were infused to achieve polymerization. In some embodiments, the biological sample can be infused with proteinase K for proteolysis and incubated in a digestion buffer. In some embodiments, the gel inside the biological sample can be swelled by addition of water.

[00137] In any of the methods described herein, the biological sample can be cultured on the support. In some embodiments, the methods comprise culturing the biological sample on the support under a condition suitable for expanding the biological sample for 2-10 generations or more. The cultured biological sample can generate a colony of cells. In some embodiments, the methods comprise culturing the biological sample to confluence or non-

confluence. In some embodiments, the methods comprise culturing the biological sample on the support in a simple or complex cell culture media. For example, the cell culture media comprises D-MEM high glucose (e.g., from Thermo Fisher Scientific, catalog No. 11965118), fetal bovine serum (e.g., 10% FBS; for example from Thermo Fisher Scientific, catalog No. A3160402), MEM non-essential amino acids (e.g., 0.1 mM MEM, for example from Thermo Fisher Scientific, catalog No. 11140050), L-glutamine (e.g., 6 mM L-glutamine, for example from Thermo Fisher Scientific, catalog No. A2916801), MEM sodium pyruvate (e.g., 1 mM sodium pyruvate, for example from Thermo Fisher Scientific, catalog No. 11360070), and an antibiotic (e.g., 1% penicillin-streptomycin-glutamine, for example from Thermo Fisher, catalog No. 10378016). In some embodiments, the methods comprise culturing the biological sample at a humidity and temperature that is suitable for culturing the cell(s) on the support. Example suitable conditions comprise approximately 37 °C with a humidified atmosphere of approximately 5-10% carbon dioxide in air. The biological sample can be cultured with suitable aeration with oxygen and/or nitrogen.

[00138] In any of the methods described herein, the term "simple cell media" or related terms refers to a cell media that typically lacks ingredients to support cell growth and/or proliferation in culture. Simple cell media can be used for example to wash, suspend, or dilute the biological sample. Simple cell media can be mixed with certain ingredients to prepare a cell media that can support cell growth and/or proliferation in culture. A simple cell media comprises any one or any combination of two or more of a buffer, a phosphate compound, a sodium compound, a potassium compound, a calcium compound, a magnesium compound and/or glucose. In some embodiments, the simple cell media comprises PBS (phosphate buffered saline), DPBS (Dulbecco's phosphate-buffered saline), HBSS (Hank's balanced salt solution), DMEM (Dulbecco's Modified Eagle's Medium), EMEM (Eagle's Minimum Essential Medium), and/or EBSS. In some embodiments, the biological sample can be placed in a simple cell media prior to or during the step of conducting any of the nucleic acid methods described herein.

[00139] In any of the methods described herein, the term "complex cell media" or related terms refers to a cell media that can be used to support cell growth and/or proliferation in culture without supplementation or additives. Complex cell media can include any combination of two or more of a buffering system (e.g., HEPES), inorganic salt(s), amino acid(s), protein(s), polypeptide(s), carbohydrate(s), fatty acid(s), lipid(s), purine(s) and their derivatives (e.g., hypoxanthine), pyrimidine(s) and their derivatives, and/or trace element(s). Complex cell media includes fluids obtained from a fluid or tissue extract. Complex cell media includes artificial cell media. In some embodiments, complex cell media can be a serum-containing media, for example complex cell media includes fluids such as fetal bovine serum, blood plasma, blood serum, lymph

fluid, human placental cord serum and amniotic fluid. In some embodiments, complex cell media can be a serum-free media, which are typically (but not necessarily) defined cell culture media. In some embodiments, complex cell media can be a chemically-defined media which typically (but not necessarily) include recombinant polypeptides, and ultra-pure inorganic and/or organic compounds. In some embodiments, complex cell media can be a protein-free media which include for example MEM (minimal essential media) and RPMI-1640 (Roswell Park Memorial Institute). In some embodiments, the complex cell media comprises IMDM (Iscove's Modified Dulbecco's Medium. In some embodiments, the complex cell media comprises DMEM (Dulbecco's Modified Eagle's Medium). In some embodiments, the biological sample can be placed in a complex cell media prior to or during the step of conducting any of the nucleic acid methods described herein.

[00140] In any of the methods described herein, the biological sample may comprise a fixed biological sample. In some embodiments, the biological 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 nucleic acids (e.g. RNA) harbored by the biological sample. In some embodiments, the fixation reagent inhibits loss of nucleic acids from the biological sample.

[00141] In some embodiments, the fixation reagent can cross-link the target to prevent the target from escaping the biological 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).

[00142] 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. 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. In some embodiments, the biological sample can be fixed on the support with 4% paraformaldehyde for about 30-60 minutes and washed with PBS. In some embodiments, the biological sample can be stained, de-stained or un-stained.

ii. Targets

[00143] The methods described herein relate to detecting one or more targets. The one of more targets may comprise a nucleic acid (e.g. an RNA, and/or a DNA) a polypeptide, or a combination thereof. In some embodiments, the one or more targets may comprise one or more naturally-occurring nucleic acids, one or more recombinant nucleic acids, one or more synthesized nucleic acids, or a combination thereof. The one or more targets comprising one or more nucleic acids may comprise linear forms of nucleic acids, circular forms of nucleic acids, branched forms of nucleic acids, or a combination thereof. In some embodiments, the one or more targets comprising one or more nucleic acids may comprise DNA. In some embodiments, the one or more targets may comprise genomic DNA. In some embodiments, the one or more targets may comprise viral DNA. In some embodiments, the one or more targets may comprise cell free DNA (cfDNA). In some embodiments, the DNA may be genomic DNA, methylated DNA, un-methylated DNA, organellar DNA, or a combination thereof. In some cases, the one or more targets may comprise PCR amplified DNA. In some embodiments, the one or more targets may comprise synthetic DNA. The synthetic DNA may comprise synthetic DNA probes. The DNA may be fragmented, unfragmented, or a combination thereof. The DNA may comprise one or more modifications. For example, the DNA may comprise a methyl modification, a phosphorylation modification, a locked nucleic acid modification, or a combination thereof.

[00144] In some embodiments, the one or more targets may comprise RNA. The RNA of the one or more targets may comprise poly-A RNA and/or non-poly-a RNA. The RNA may comprise coding RNA, non-coding RNA, or a combination thereof. The RNA may comprise tRNA, rRNA, small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), microRNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), antisense RNA, non-coding RNA, protein-encoding RNA, or a combination thereof. The one or more targets may comprise RNA that is target RNA, non-target RNA, or a combination thereof. In some cases, the RNA may have been transcribed and have undergone splicing. In some cases, the RNA may not have been spliced. In some cases, the RNA may comprise wild type RNA, mutant RNA or splice variant RNA. In some cases, the RNA may comprise pre-spliced RNA, partially spliced RNA, or fully spliced RNA. In some cases, the RNA may comprise coding RNA, non-coding RNA, housekeeping RNA, cell-specific RNA, tissue-specific RNA, disease-specific RNA, or a combination thereof. In some embodiments, the RNA may comprise 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 RNA may comprise RNA found in healthy cells, diseased cells, or a combination thereof. In some embodiments, the RNA may comprise RNA transcribed from transgenic DNA sequences that are introduced into the biological 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 RNA may comprise RNA that is transcribed from DNA sequences that are not transgenic. The RNA may comprise one or more modifications. For example, the RNA may comprise a phosphorylation (e.g. a 5' phosphorylation), an inter-nucleotide linkage, or a combination thereof.

[00145] A binding complex associated with the RNA or an amplicon generated from binding a nucleic acid molecule to the RNA may be formed without performing reverse transcription of the RNA. For example, the RNA may be bound directly by a probe comprising an RNA-binding sequence. The probe may be ligated to form a circular nucleic acid and amplified using rolling circle amplification to generate one or more amplicons. The one or more amplicons may form a binding complex with a nucleotide conjugated and detected. In this case, the RNA is detected directly without performing reverse transcription.

[00146] The one or more targets may comprise one or more polypeptides. The one of more polypeptides of the one or more targets may comprise one or more proteins, one or more peptides, or a combination thereof. The one or more polypeptides may comprise one or more cell-surface proteins, one or more intracellular proteins, one or more signaling proteins, one or more protein complexes, or a combination thereof. The one or more polypeptides may comprise one or more antibodies or antibody fragments. In some cases, the one or more polypeptides may comprise one or more modifications. The one or more modifications of the one or more polypeptides may comprise one or more phosphoramidate (N-linked) formations, one or more phosphorylations, one or more adenylylations, one or more uridylylations, one or more propionylations, one or more pyroglutamate formations, one or more S-glutathionylations, one or more S-nitrosylations, one or more S-sulfenylations, one or more S-sulfinylations, one or more S-sulfonylations, one or more succinylations, one or more sulfations, one or more glycations, one or more carbamylations, one or more carbonylations, one or more spontaneous isopeptide bond formations, one or more biotinylations, one or more carbamylations, one or more oxidations, one or more myristoylations, one or more palmitoylations, one or more farnesylations, one or more geranylgeranylations, one or more glypiations, one or more glycosylphosphatidylinositols, one or more lipoylations, one or more flavin moiety attachments, one or more heme C attachments, one or more phosphopantetheinylations, one or more retinylidene Schiff base formations, one or more modifications of translation factors, one or more diphthamide formations, one or more ethanolamine phosphoglycerols, one or more hypusine formations, one or more beta-Lysine addition on a lysines, one or more acylations (e.g. O-acylation, N-acylation, and S-acylation), one or more acetylations, one or more formylations, one or more alkylations, one or more amidations, one or more arginylations, one or more polyglutamylations, one or more polyglycylations, one or

more butyrylations, one or more gamma-carboxylations, one or more glycosylation, one or more polysialylations, one or more malonylations, one or more hydroxylations, one or more nucleotide additions, one or more phosphate esters (O-linked), one or more pegylations, one or more ubiquitinations, one or more SUMOylations, one or more neddylations, one or more ISGylations, one or more citrullinations, one or more deamidations, one or more eliminylations, or a combination thereof.

[00147] In some embodiments, the one or more targets may comprise nucleic acids that are clonally-amplified. In some embodiments, the present disclosure provides a method comprising clonal amplification comprising the use of a 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, circle-to-circle amplification, helicase-dependent amplification, recombinase-dependent amplification, single-stranded binding (SSB) protein-dependent amplification, or any combination thereof. In some embodiments, the product of clonal amplification may be detected using the methods described herein.

[00148] The one or more targets of the instant disclosure may have a fixed three-dimensional relationship within the biological sample after the biological sample is coupled to the surface. This fixed three-dimensional relationship, at least partially, enables the identification of spatial and cellular origin within the biological sample following nucleic acid identification using the systems and methods described herein. Nucleic acid analysis inside a biological sample and spatial analysis is described in U.S. Patent Application No. 17/356,929, filed June 24, 2021, entitled "Multivalent Binding Composition for Nucleic Acid Analysis", which is hereby incorporated by reference in its entirety.

[00149] In some embodiments, the one or more targets comprises one or more target nucleic acid sequences. In some embodiments, the one or more targets may be a concatemer comprising multiple repeats of the one or more target nucleic acid sequences. The multiple copies of the target nucleic acid sequence within the concatemer may be covalently linked using phosphodiester bonds. The concatemer may be formed using an amplification reaction, including a rolling circle amplification reaction.

[00150] In some embodiments, two or more targets may be detected. In some embodiments, the two or more targets can each comprise the same targets (e.g. the same nucleic acid sequence). In some embodiments, the two or more targets may comprise different targets (e.g. different nucleic acid sequences). In some embodiments the one or more targets may comprise at least about 1 target, at least about 2 targets, at least about 3 targets, at least about 4

targets, at least about 5 targets, at least about 6 targets, at least about 7 targets, at least about 8 targets, at least about 9 targets, at least about 10 targets, at least about 20 targets, at least about 30 targets, at least about 40 targets, at least about 50 targets, at least about 60 targets, at least about 70 targets, at least about 80 targets, at least about 90 targets, at least about 100 targets, at least about 200 targets, at least about 500 targets, at least about 1000 targets, at least about 2000 targets, at least about 5000 targets, at least about 10000 targets, at least about 12000 targets, at least about 20000 targets, at least about 25000 targets, at least about 50000 targets, at least about 100000 targets, at least about 1000000 targets, or more. In some embodiments the one or more targets may comprise at most about 1 target, at most about 2 targets, at most about 3 targets, at most about 4 targets, at most about 5 targets, at most about 6 targets, at most about 7 targets, at most about 8 targets, at most about 9 targets, at most about 10 targets, at most about 20 targets, at most about 30 targets, at most about 40 targets, at most about 50 targets, at most about 60 targets, at most about 70 targets, at most about 80 targets, at most about 90 targets, at most about 100 targets, at most about 200 targets, at most about 500 targets, at most about 1000 targets, at most about 2000 targets, at most about 5000 targets, at most about 10000 targets, at most about 12000 targets, at most about 20000 targets, at most about 25000 targets, at most about 50000 targets, at most about 100000 targets, or at most about 1000000 targets.

iii. Nucleic acid molecules

[00151] The methods described herein may involve coupling one or more nucleic acid molecules to the one or more targets for detection. The one or more nucleic acid molecules as described herein may comprise one or more probes. The one or more probes may comprise one or more RNA-binding sequences, one or more DNA-binding sequences, one or more oligonucleotide-binding sequences, one or more index sequences, or a combination thereof. In some cases, the one or more nucleic acid molecules may comprise one or more probes configured to bind to an RNA. The one or more probes configured to bind to the RNA may comprise an RNA-binding sequence, an index sequence, or a combination thereof. The one or more probes configured to bind to the RNA my comprise a sequence that is at least partially complementary to a portion of the RNA sequence or reverse complement thereof. The sequence that is at least partially complementary to a portion of the RNA sequence or reverse complement thereof may be an RNA-binding sequence. In some cases, the one or more nucleic acid molecules may comprise one or more probes configured to bind to a DNA. The one or more probes configured to bind to the DNA may comprise an DNA-binding sequence, an index sequence, or a combination thereof. The one or more probes configured to bind to the DNA my comprise a sequence that is at least partially complementary to a portion of the DNA sequence or reverse complement thereof. The

sequence that is at least partially complementary to a portion of the DNA sequence or reverse complement thereof may be an DNA-binding sequence. In some cases, the one or more nucleic acid molecules may comprise one or more probes configured to bind to a polypeptide-binding probe. The one or more probes configured to bind to the polypeptide-binding probe may comprise an oligonucleotide-binding probe that is at least partially complementary to an oligonucleotide of the polypeptide-binding probe. In some cases, the one or more nucleic acid molecules may comprise a target-binding sequence, an index sequence, or a combination thereof. In some cases, the one or more nucleic acid molecules may be configured to bind to at least a portion of an RNA or a reverse complement thereof.

[00152] In some embodiments, the one or more nucleic acid molecules may comprise DNA. In some embodiments, the one or more nucleic acid molecules may comprise RNA. In some embodiments, the one or more nucleic acid molecules may comprise a combination of DNA and RNA. In some embodiments, the one or more nucleic acid molecule may comprise a modification. In some embodiments, the modification may include a phosphorylation, an internucleotide linkage, a fluorescent dye, or a combination thereof. In some embodiments, the internucleotide linkage may comprise a locked nucleic acid residue. In some embodiments, the modification may comprise a 5' modification, a 3' modification, an internal modification, or a combination thereof.

[00153] The one or more nucleic acid molecules may have a variety of lengths. In some embodiments, each of the one or more nucleic acid molecules may have the same length. In some embodiments, each of the one or more nucleic acid molecules may have a different length. In some embodiments, at least one nucleic acid molecule or the one or more nucleic acid molecules has a different length than a second nucleic acid molecule of the one or more nucleic acid molecules. In some embodiments, the length of the one or more nucleic acid molecules may be about 1-100 nucleotides, about 2-98 nucleotides, about 3-96 nucleotides, about 4-94 nucleotides, about 5-92 nucleotides, about 6-90 nucleotides, about 7-88 nucleotides, about 8-86 nucleotides, about 9-84 nucleotides, about 10-82 nucleotides, about 11-80 nucleotides, about 12-78 nucleotides, about 13-76 nucleotides, about 14-74 nucleotides, about 15-72 nucleotides, about 16-70 nucleotides, about 17-68 nucleotides, about 18-66 nucleotides, about 19-64 nucleotides, about 20-62 nucleotides, about 22-60 nucleotides, about 24-58 nucleotides, about 26-56 nucleotides, about 28-54 nucleotides, about 30-52 nucleotides, about 32-50 nucleotides, about 34-48 nucleotides, about 36-46 nucleotides, about 38-44 nucleotides, about 40-42 nucleotides, at least about 1 nucleotide, at least about 2 nucleotides, at least about 3 nucleotides, at least about 4 nucleotides, at least about 5 nucleotides, at least about 6 nucleotides, at least about 7 nucleotides, at least about 8 nucleotides, at least about 9 nucleotides, at least about 10 nucleotides, at least

about 11 nucleotides, at least about 12 nucleotides, at least about 13 nucleotides, at least about 14 nucleotides, at least about 15 nucleotides, at least about 16 nucleotides, at least about 17 nucleotides, at least about 18 nucleotides, at least about 19 nucleotides, at least about 20 nucleotides, at least about 22 nucleotides, at least about 24 nucleotides, at least about 26 nucleotides, at least about 28 nucleotides, at least about 30 nucleotides, at least about 32 nucleotides, at least about 34 nucleotides, at least about 36 nucleotides, at least about 38 nucleotides, at least about 40 nucleotides, at least about 42 nucleotides, at least about 44 nucleotides, at least about 46 nucleotides, at least about 48 nucleotides, at least about 50 nucleotides, at least about 52 nucleotides, at least about 54 nucleotides, at least about 56 nucleotides, at least about 58 nucleotides, at least about 60 nucleotides, at least about 62 nucleotides, at least about 64 nucleotides, at least about 66 nucleotides, at least about 68 nucleotides, at least about 70 nucleotides, at least about 72 nucleotides, at least about 74 nucleotides, at least about 76 nucleotides, at least about 78 nucleotides, at least about 80 nucleotides, at least about 82 nucleotides, at least about 84 nucleotides, at least about 86 nucleotides, at least about 88 nucleotides, at least about 90 nucleotides, at least about 92 nucleotides, at least about 94 nucleotides, at least about 96 nucleotides, at least about 98 nucleotides, or at least about 100 nucleotides. In some embodiments, the nucleic acid molecule is about 10-80 nucleotides in length. In some embodiments, the nucleic acid molecule is about 20-60 nucleotides in length. In some embodiments, the nucleic acid molecule is about 30-50 nucleotides in length.

[00154] In some embodiments, one nucleic acid molecule may couple to a target (e.g. a nucleic acid). In some embodiments, more than one nucleic acid molecule may couple to a target (e.g. a nucleic acid). In some embodiments, one nucleic acid molecule (e.g. a probe comprising an oligonucleotide-binding probe) may couple to a probe that is bound to a target (e.g. a polypeptide-binding probe that is bound to a polypeptide). In some embodiments, more than one nucleic acid molecule (e.g. more than one probe comprising an oligonucleotide-binding probe) may couple to a probe that is bound to a target((e.g. a polypeptide-binding probe that is bound to a polypeptide). In some embodiments, a first set of nucleic acid molecules may couple to a first target and a second set of nucleic acid molecules may couple to a second target. In some embodiments, the number of nucleic acid molecules in the first set and second set are the same. In some embodiments, the number of nucleic acid molecules in the first set and second set are different. In some embodiments, the number of nucleic acid molecules in the first set is greater than the number of nucleic acid molecules in the second set. In some embodiments, the number of nucleic acid molecules in the first set is less than the number of nucleic acid molecules in the second set. In some embodiments, a first set of nucleic acid molecules (e.g. probes comprising

oligonucleotide-binding probes) may couple to a first probe that is bound to a first target (a first polypeptide-binding probe that is bound to a first polypeptide) and a second set of nucleic acid molecules (e.g. probes comprising oligonucleotide-binding probes) may couple to a second target (e.g. a second polypeptide-binding probe that is bound to a second polypeptide). In some embodiments, the number of nucleic acid molecules in the first set and second set are the same. In some embodiments, the number of nucleic acid molecules in the first set and second set are different. In some embodiments, the number of nucleic acid molecules in the first set is greater than the number of nucleic acid molecules in the second set. In some embodiments, the number of nucleic acid molecules in the first set is less than the number of nucleic acid molecules in the second set.

[00155] In some embodiments, a nucleic acid molecule of the one or more nucleic acid molecule comprises at least one binding site (e.g. an RNA-binding sequence, a DNA-binding sequence, an oligonucleotide-binding probe, or a target binding sequence). The binding site may comprise a nucleic acid sequence that hybridizes to the one or more targets for detection. The binding site may have a variety of lengths. In some embodiments, one binding site may have a different length than a second binding site. In some embodiments, the length of the one or more nucleic acid molecules may be about 1-100 nucleotides, about 2-98 nucleotides, about 3-96 nucleotides, about 4-94 nucleotides, about 5-92 nucleotides, about 6-90 nucleotides, about 7-88 nucleotides, about 8-86 nucleotides, about 9-84 nucleotides, about 10-82 nucleotides, about 11-80 nucleotides, about 12-78 nucleotides, about 13-76 nucleotides, about 14-74 nucleotides, about 15-72 nucleotides, about 16-70 nucleotides, about 17-68 nucleotides, about 18-66 nucleotides, about 19-64 nucleotides, about 20-62 nucleotides, about 22-60 nucleotides, about 24-58 nucleotides, about 26-56 nucleotides, about 28-54 nucleotides, about 30-52 nucleotides, about 32-50 nucleotides, about 34-48 nucleotides, about 36-46 nucleotides, about 38-44 nucleotides, about 40-42 nucleotides, at least about 1 nucleotide, at least about 2 nucleotides, at least about 3 nucleotides, at least about 4 nucleotides, at least about 5 nucleotides, at least about 6 nucleotides, at least about 7 nucleotides, at least about 8 nucleotides, at least about 9 nucleotides, at least about 10 nucleotides, at least about 11 nucleotides, at least about 12 nucleotides, at least about 13 nucleotides, at least about 14 nucleotides, at least about 15 nucleotides, at least about 16 nucleotides, at least about 17 nucleotides, at least about 18 nucleotides, at least about 19 nucleotides, at least about 20 nucleotides, at least about 22 nucleotides, at least about 24 nucleotides, at least about 26 nucleotides, at least about 28 nucleotides, at least about 30 nucleotides, at least about 32 nucleotides, at least about 34 nucleotides, at least about 36 nucleotides, at least about 38 nucleotides, at least about 40 nucleotides, at least about 42 nucleotides, at least about 44 nucleotides, at least about 46 nucleotides, at least about 48

nucleotides, at least about 50 nucleotides, at least about 52 nucleotides, at least about 54 nucleotides, at least about 56 nucleotides, at least about 58 nucleotides, at least about 60 nucleotides, at least about 62 nucleotides, at least about 64 nucleotides, at least about 66 nucleotides, at least about 68 nucleotides, at least about 70 nucleotides, at least about 72 nucleotides, at least about 74 nucleotides, at least about 76 nucleotides, at least about 78 nucleotides, at least about 80 nucleotides, at least about 82 nucleotides, at least about 84 nucleotides, at least about 86 nucleotides, at least about 88 nucleotides, at least about 90 nucleotides, at least about 92 nucleotides, at least about 94 nucleotides, at least about 96 nucleotides, at least about 98 nucleotides, or at least about 100 nucleotides in length. In some embodiments, the target-binding sequence is about 5-25 nucleotides in length.

[00156] In some embodiments, the one or more nucleic acid molecules comprise at least one index sequence. The at least one index sequence may denote the identity of the one or more targets. In some embodiments, a single index sequence corresponds one or more targets. In some embodiments, more than one index sequence corresponds to one or more targets. In some embodiments, the one or more nucleic acid molecules comprises at least about 2, 3, 4, 5, 6, 7, 8, 9, or 10 index sequences. The at least one index sequence may be sequence orthogonal to the sequence of the one or more targets. In some embodiments, the at least one index sequence includes at least a portion of a sequence of the one or more targets.

[00157] In some cases, a nucleic acid molecule of the one or more nucleic acid molecules may comprise a binding site (e.g. a target binding site) that is adjacent to an index sequence. For example, the binding site (e.g. a target binding site) may be directly proximal to the index sequence. In another example, the binding site may be separated by at least one nucleotide from the index sequence. In some cases, the nucleic acid molecule may be bound to a target (e.g. an RNA) and the index sequence may be unbound. For example, a nucleic acid molecule may comprise a target-binding sequence that is hybridized to an RNA and the nucleic acid molecule may comprise an index sequence that is not hybridized (e.g. unhybridized) to the RNA. The index sequence may be single-stranded and amenable to hybridization by another nucleic acid sequence (e.g. an index-binding sequence).

[00158] Each of the at least one index sequence may have a variety of lengths. The length of an index sequence may be about 1-25 nucleotides, about 2-24 nucleotides, about 3-23 nucleotides, about 4-22 nucleotides, about 5-21 nucleotides, about 6-20 nucleotides, about 7-19 nucleotides, about 8-18 nucleotides, about 9-17 nucleotides, about 10-16 nucleotides, about 11-15 nucleotides, about 12-14 nucleotides, at least about 1 nucleotide, at least about 2 nucleotides, at least about 3 nucleotides, at least about 4 nucleotides, at least about 5 nucleotides, at least about 6 nucleotides, at least about 7 nucleotides, at least about 8 nucleotides, at least about 9 nucleotides,

at least about 10 nucleotides, at least about 11 nucleotides, at least about 12 nucleotides, at least about 13 nucleotides, at least about 14 nucleotides, at least about 15 nucleotides, at least about 16 nucleotides, at least about 17 nucleotides, at least about 18 nucleotides, at least about 19 nucleotides, at least about 20 nucleotides, at least about 21 nucleotides, at least about 22 nucleotides, at least about 23 nucleotides, at least about 24 nucleotides, or at least about 25 nucleotides. In some embodiments, the index sequence is 4-50 nucleotides in length.

[00159] In some embodiments, the index sequence comprises a primer binding site. In some embodiments, where a nucleic acid molecule comprises more than one index sequence, each index sequence comprises a primer binding site. The primer binding site may be configured to hybridize to a primer nucleic acid sequence as described herein. The primer binding site may hybridize to a primer nucleic acid sequence. The primer binding sites of different index sequences may have a variety of lengths. The length of an primer binding site may be about 1-25 nucleotides, about 2-24 nucleotides, about 3-23 nucleotides, about 4-22 nucleotides, about 5-21 nucleotides, about 6-20 nucleotides, about 7-19 nucleotides, about 8-18 nucleotides, about 9-17 nucleotides, about 10-16 nucleotides, about 11-15 nucleotides, about 12-14 nucleotides, at least about 1 nucleotide, at least about 2 nucleotides, at least about 3 nucleotides, at least about 4 nucleotides, at least about 5 nucleotides, at least about 6 nucleotides, at least about 7 nucleotides, at least about 8 nucleotides, at least about 9 nucleotides, at least about 10 nucleotides, at least about 11 nucleotides, at least about 12 nucleotides, at least about 13 nucleotides, at least about 14 nucleotides, at least about 15 nucleotides, at least about 16 nucleotides, at least about 17 nucleotides, at least about 18 nucleotides, at least about 19 nucleotides, at least about 20 nucleotides, at least about 21 nucleotides, at least about 22 nucleotides, at least about 23 nucleotides, at least about 24 nucleotides, or at least about 25 nucleotides.

[00160] The index sequence may be uniquely associated with the one or more targets for detection. The one or more targets may comprise an RNA, a DNA, a polypeptide, or a combination thereof. In some embodiments, an index sequence may be associated with the one or more targets and at least one other analyte. In some embodiments, the other analyte may be a nucleic acid or a polypeptide. In some embodiment, the other analyte may be an RNA. In some embodiments, the other analyte may comprise naturally-occurring nucleic acid, recombinant nucleic acid, synthesized nucleic acid, or a combination thereof. In some embodiments, the other analyte may be a DNA. In some embodiments, the other analyte may be genomic DNA. In some embodiments, the other analyte may be viral DNA. In some embodiments, the other analyte may be cfDNA. In some embodiments, the other analyte may be poly-A RNA and/or non-poly-a RNA. The RNA may comprise coding RNA, non-coding RNA, or a combination thereof. The RNA may comprise tRNA, rRNA, small nuclear RNA (snRNA), small nucleolar RNA (snoRNA),

microRNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), antisense RNA, non-coding RNA, protein-encoding RNA, or a combination thereof.

[00161] In some embodiments the other analyte may be a polypeptide. In some embodiments, the polypeptide may be a protein, a peptide, or a combination thereof. In some embodiments, the protein may be an antibody, antibody fragment, ribosomal protein, transcription factor, histone, enzyme, structural protein, chemokine, cytokine, or a combination thereof.

[00162] The nucleic acid molecules described herein may comprise one or more padlock probes. The one or more padlock probes may comprise nucleic acid. The nucleic acid of the one or more padlock probes may comprise DNA, RNA, or a combination thereof. The one or more padlock probes may comprise one or more modifications, including but not limited to a phosphorylation (e.g. a 5' phosphorylation), a locked nucleic acid, or a combination thereof. In some embodiments, the padlock probes comprise canonical nucleotides and/or nucleotide analogs. In some embodiments, the padlock probes are modified to confer resistance to nuclease degradation (e.g., ribonuclease degradation). For example, the padlock probes comprise at least one phosphorothioate diester bond at their 5' ends which can render the padlock probes resistant to nuclease degradation. In some embodiments, the padlock probes comprise 2-5 or more consecutive phosphorothioate diester bonds at their 5' ends. In some embodiments, the padlock probes comprise at least one ribonucleotide and/or at least one 2'-O-methyl, 2'-O-methoxyethyl (MOE), 2' fluoro-base nucleotide. In some embodiments, the padlock probes comprise phosphorylated 3' ends. In some embodiments, the padlock probes comprise at least one locked nucleic acid (LNA) base. In some embodiments, the padlock probes comprise a phosphorylated 5' end (e.g., using a polynucleotide kinase).

[00163] A padlock probe of the one or more padlock probes may comprise a first end, a second end, an index sequence, additional sequence, or a combination thereof. The first end of the padlock probe may comprise a sequence that is at least partially complementary to a sequence of a target (e.g. an RNA or a DNA) or reverse complement thereof. The first end of the padlock probe may comprise a sequence that is at least partially complementary to a portion of an oligonucleotide of a polypeptide-binding probe or reverse complement thereof. The second end of the padlock probe may comprise a sequence that is at least partially complementary to a sequence of a target (e.g. an RNA or a DNA) or reverse complement thereof. The second end of the padlock probe may comprise a sequence that is at least partially complementary to a portion of an oligonucleotide of a polypeptide-binding probe or reverse complement thereof. In some cases, the first end of the padlock probe is directly adjacent to the second end of the padlock probe when both the first end and the second end are bound to a target (e.g. an RNA or a DNA) or bound to an oligonucleotide of a polypeptide-binding probe. For example, the first end of the padlock

probe may hybridize to a first sequence of the RNA and the second end of the padlock probe may hybridize to a second sequence of an RNA and the first sequence and the second sequence are contiguous sequences of the RNA such that there are no unhybridized nucleotides on the RNA when bound to the padlock probe at the junction of the first end and the second end of the padlock probe. In another example, the first end of the padlock probe may hybridize to a first portion of an oligonucleotide of a polypeptide-binding probe and the second end of the padlock probe may hybridize to a second portion of the oligonucleotide of the polypeptide-binding probe and the first portion and the second portion are contiguous sequences of the oligonucleotides such that there are no unhybridized nucleotides on the oligonucleotide when bound to the padlock at the junction of the first end and the second end of the padlock probe.

[00164] In some embodiments, the first end of an individual padlock probe has a first random sequence that hybridizes to a first region of a target (e.g. an RNA molecule), and the second end of the individual padlock probe has a second random sequence that hybridizes to a second region of the same target, where a nick or gap is formed between the hybridized first and second binding arms. In some embodiments, individual padlock probes in a set of padlock probes (e.g., a plurality of padlock probes) comprise first and second ends that hybridize to the same target regions of the targets (e.g. RNA molecules) to form a plurality of target-padlock probe complexes having the same target sequence.

[00165] In some cases, the first end of the padlock probe is adjacent to the second end of the padlock probe and separated by a gap (e, g. separated by at least one nucleotide) when both the first end and the second end are bound to a target (e.g. an RNA or a DNA) or bound to an oligonucleotide of a polypeptide-binding probe. For example, the first end of the padlock probe may hybridize to a first sequence of the RNA and the second end of the padlock probe may hybridize to a second sequence of an RNA and the first sequence and the second sequence are noncontiguous sequences of the RNA such that there is at least one unhybridized nucleotide on the RNA when bound to the padlock probe at the junction of the first end and the second end of the padlock probe. In another example, the first end of the padlock probe may hybridize to a first portion of an oligonucleotide of a polypeptide-binding probe and the second end of the padlock probe may hybridize to a second portion of the oligonucleotide of the polypeptide-binding probe and the first portion and the second portion are noncontiguous sequences of the oligonucleotides such that there is at least one unhybridized nucleotide on the oligonucleotide when bound to the padlock at the junction of the first end and the second end of the padlock probe. In cases where the first end of the padlock probe and the second end of the padlock probe are separated by a gap, a gap-filling reaction can be performed to fill in the unhybridized nucleotides of the target or oligonucleotide of the polypeptide-binding probe. The gap-filling reaction can be performed by a

polymerizing enzyme, e.g. a DNA polymerase or an RNA polymerase. In some embodiments, the size of the gap between the hybridized first and second binding arms is 1-30 bases. The 3'OH end of hybridized padlock probe can serve as an initiation site for a reverse transcriptase-catalyzed fill-in reaction (e.g., gap fill-in reaction) using the target RNA molecule as a template. After the fill-in reaction, the remaining nick can be enzymatically ligated to generate covalently closed circular nucleic acid.

[00166] In some embodiments, the gap-filling reaction comprises contacting the circular nucleic acid with a reverse transcriptase enzyme and a plurality of nucleotides. In some embodiments, the reverse transcriptase comprises In some embodiments, the reverse transcriptase enzyme comprises a reverse transcriptase enzyme from AMV (avian myeloblastosis virus), M-MuLV (moloney murine leukemia virus), or HIV (human immunodeficiency virus). In some embodiment, the reverse transcriptase enzyme comprises a recombinant enzyme that exhibits reduced RNase H activity, for example REVERTAID (e.g., from Thermo Fisher Scientific, catalog No. EP0441). In some embodiments, the reverse transcriptase can be a commercially-available enzyme, including MULTISCRIBE (e.g., from Thermo Fisher Scientific, catalog # 4311235), THERMOSCRIPT (e.g., from Thermo Fisher Scientific, catalog # 12236-014), or ARRAYSCRIPT (e.g., from Ambion, catalog No. AM2048). In some embodiments, the reverse transcriptase enzyme comprises SUPERSCRIPT II (e.g., catalog No. 18064014), SUPERSCRIPT III (e.g., catalog No. 18080044), or SUPERSCRIPT IV enzymes (e.g., catalog No. 18090010) (all SUPERSCRIPT enzymes from Invitrogen). In some embodiments, the reverse transcription reaction can include an RNase inhibitor.

[00167] The one or more padlock probes may be circularized to generate one or more circular nucleic acids. For example, a first end of a padlock probe may be ligated to a second end of the padlock probe to form a circular nucleic acid. The padlock probe may be ligated by a ligase, a chemical reaction, or a combination thereof. In some cases, one end of a padlock probe (e.g. a first end or a second end) may comprise a phosphorylation modification (e.g. a 5' phosphorylation). A ligase may ligate the end of the padlock probe comprising the phosphorylation modification to the other end of the padlock probe to form a circular nucleic acid. In some embodiments, the ligase may be a DNA ligase. In some embodiments, the ligase may be an RNA ligase. In some embodiments, the ligase may be a T4 ligase, a SplintR ligase, a T3 ligase, a T4 ligase, a T7 ligase, an TtcB ligase, a Taq ligase or a combination thereof. In some embodiments, the ligase can discriminate between matched and mis-matched hybridized ends to ensure target-specific hybridization. In some cases, one end of a padlock probe may comprise a biorthogonal modification (e.g. an azide or an alkyne). In some cases, a chemical reaction between

the biorthogonal modification of the one end of the padlock probe may be reacted to another end of the padlock probe, which may comprise a second biorthogonal modification to form a covalent bond between the two ends in order to form a circular nucleic acid molecule.

[00168] In some embodiments, the padlock probes comprise at least one universal adaptor sequence including a sample barcode sequence, an amplification primer binding site, a sequencing primer binding site and/or a compaction oligonucleotide binding site. In some embodiments, the padlock probes comprise at least one unique identification sequence (e.g., unique molecular index (UMI). In some embodiments, the padlock probes comprise at least one restriction enzyme recognition sequence.

[00169] In some embodiments, a set of padlock probes (e.g., a plurality of padlock probes) comprise at least two sub-sets (e.g., two sub-populations) of padlock probes (e.g., FIGs. 28, 30 and 32). In some embodiments, individual padlock probes in a first sub-set of padlock probes comprise first and second binding arms that hybridize to the same target regions (e.g., a first target region) of the first target RNA molecules to form a first plurality of RNA-padlock probe complexes having the same RNA sequence. In some embodiments, individual padlock probes in a second sub-set of padlock probes comprise first and second binding arms that hybridize to the same target regions (e.g., a second target region) of the second target RNA molecules to form a second plurality of RNA-padlock probe complexes having the same RNA sequence. In some embodiments, the first and second sub-sets of padlock probes hybridize to different target regions of the same target RNA molecules. In some embodiments, the first and second sub-sets of padlock probes hybridize to different target regions of different target RNA molecules. In some embodiments, the set of padlock probes comprise 2-10 sub-sets of padlock probes, or 10-25 sub-sets of padlock probes, or 25-50 sub-sets of padlock probes, or up to 100 sub-sets of padlock probes. In some embodiments, the set of padlock probes comprise at least 100 sub-sets of padlock probes, at least 500 sub-sets of padlock probes, at least 1000 sub-sets of padlock probes, at least 10,000 sub-sets of padlock probes, or more sub-sets of padlock probes.

iv. Polypeptide-binding probes

[00170] Disclosed herein are polypeptide-binding probes that are useful in the methods described herein for detecting a target (e.g. a polypeptide) *in situ*. The polypeptide-binding probes may be configured to bind one or more polypeptides. The polypeptide-binding probes may comprise a polypeptide binding moiety and an oligonucleotide. The polypeptide binding moiety may be coupled to the oligonucleotide either directly or indirectly using one or more covalent

interactions, one or more non-covalent interactions, or a combination thereof. For example, the polypeptide binding moiety may comprise a polypeptide and the polypeptide of the polypeptide binding moiety may be covalently linked to the oligonucleotide. Additionally, the polypeptide binding moiety may comprise a polypeptide and the polypeptide of the polypeptide binding moiety may be non-covalently linked to the oligonucleotide via a biotin-streptavidin interaction. The polypeptide binding moiety may be linked to the oligonucleotide using one or more linkers. In some embodiments, the one or more linkers may comprise streptavidin (or an avidin -like moiety), biotin, an amine group, a disulfide group, or a combination thereof. In some embodiments, the one or more linkers may not be cleavable or not removable. In some embodiments, the one or more linkers may be cleavable or removable. For example, the one or more linkers can be cleavable with light (e.g., UV light), chemically-cleavable (e.g., dithiothreitol), heat- cleavable, enzymatically cleavable, or a combination thereof.

[00171] The oligonucleotide of the polypeptide-binding probe may comprise DNA, RNA, or a combination thereof. In some cases, the oligonucleotide may comprise one or more non-natural nucleotides. The one or more non-natural nucleotides may comprise one or more of 2-MethoxyEthoxy A, 2-MethoxyEthoxy MeC, 2-MethoxyEthoxy G, 2-MethoxyEthoxy T, 8-oxo dG, deoxyadenosine, N6-Methyl rA, 5-Bromo dU, deoxyUridine, 2,6-Diaminopurine, Dideoxy-C, deoxyInosine, Hydroxymethyl dC, Iso-dG, Iso-dC, 5-Methyl dC, or 5-Nitroindole or a combination thereof. The oligonucleotide may comprise one or more locked nucleic acids (LNA). The oligonucleotide may comprise single-stranded nucleic acids, double-stranded nucleic acids, or a combination thereof. The oligonucleotide may have a length of at least about 2 nucleotides, at least about 3 nucleotides, at least about 4 nucleotides, at least about 5 nucleotides, at least about 6 nucleotides, at least about 7 nucleotides, at least about 8 nucleotides, at least about 9 nucleotides, at least about 10 nucleotides, at least about 11 nucleotides, at least about 12 nucleotides, at least about 13 nucleotides, at least about 14 nucleotides, at least about 15 nucleotides, at least about 16 nucleotides, at least about 17 nucleotides, at least about 18 nucleotides, at least about 19 nucleotides, at least about 20 nucleotides, at least about 25 nucleotides, at least about 30 nucleotides, at least about 40 nucleotides, at least about 45 nucleotides, at least about 50 nucleotides, at least about 55 nucleotides, at least about 60 nucleotides, at least about 70 nucleotides, at least about 80 nucleotides, at least about 90 nucleotides, at least about 100 nucleotides, or more. The oligonucleotide may have a length of at most about 2 nucleotides, at most about 3 nucleotides, at most about 4 nucleotides, at most about 5 nucleotides, at most about 6 nucleotides, at most about 7 nucleotides, at most about 8 nucleotides, at most about 9 nucleotides, at most about 10 nucleotides, at most about 11 nucleotides, at most about 12 nucleotides, at most about 13 nucleotides, at most about 14 nucleotides, at most about 15

nucleotides, at most about 16 nucleotides, at most about 17 nucleotides, at most about 18 nucleotides, at most about 19 nucleotides, at most about 20 nucleotides, at most about 25 nucleotides, at most about 30 nucleotides, at most about 40 nucleotides, at most about 45 nucleotides, at most about 50 nucleotides, at most about 55 nucleotides, at most about 60 nucleotides, at most about 70 nucleotides, at most about 80 nucleotides, at most about 90 nucleotides, at most about 100 nucleotides, or less. The oligonucleotide may have a length of about 2-100 nucleotides, about 3-90 nucleotides, about 4-80 nucleotides, about 5-70 nucleotides, about 6-60 nucleotides, about 7-55 nucleotides, about 8-50 nucleotides, about 9-45 nucleotides, about 10-40 nucleotides, about 11-30 nucleotides, about 12-25 nucleotides, about 13-20 nucleotides, about 14-19 nucleotides, about 15-18 nucleotides, or about 16-17 nucleotides.

[00172] The polypeptide-binding probe may comprise a polypeptide, a nucleic acid, a linker, or a combination thereof. In some cases, the polypeptide-binding probe may comprise a polypeptide. The polypeptide may comprise an antibody, an antibody fragment, a nanobody, an affimer, or a combination thereof. The antibody or antibody fragment may comprise a variety of isotypes including but not limited to IgG, IgM, IgA, IgD, IgE, or a combination thereof. In some cases, the polypeptide-binding probe may comprise an antibody or antibody fragment and the polypeptide binding moiety may comprise an a paratope. In some cases, the polypeptide-binding probe may comprise an antibody or antibody fragment and the polypeptide binding moiety may comprise an antigen binding site. In some cases, the polypeptide-binding probe may comprise a nucleic acid. The nucleic acid of the polypeptide-binding probe may comprise an aptamer. The aptamer may be configured to bind to the target that is detected in the methods described herein.

v. Index-binding sequences

[00173] Disclosed herein are nucleic acid molecules comprising index-binding sequences that are useful in the methods described herein for detecting a target (e.g., an RNA, a DNA, or a polypeptide) in situ. For example, the index-binding sequences of the nucleic acid molecules may enable the nucleic acid molecule to bind the index sequence to initiate an amplification or detection of the index sequence. In some embodiments, the nucleic acid molecule comprising the index-binding sequence is a padlock probe. In some embodiments, the nucleic acid molecule comprising the index-binding sequence is a molecular inversion probe. In some embodiments, the nucleic acid molecule comprising the index-binding sequence may comprise a second index sequence. The index-binding sequence may be configured to couple to an index sequence. In some embodiments, the index-binding sequence may hybridize to an index sequence. In some embodiments, the index-binding sequence may comprise at least one binding region that recognizes and binds at least a portion of an index sequence. In some embodiments, the index-

binding sequence may comprise more than one binding region that recognizes and binds an index sequence. For example, a padlock probe may have two index-binding sequences at a 5' end and 3' end of the padlock probe that, when each hybridized to the 5' end and 3' end of the index sequence, initiates a ligation reaction thereby adjoining the 5' and 3' ends of the padlock probe.

[00174] In some embodiments, the index-binding sequence may bind to the index sequence to form a double-stranded region. In some instances, the double-stranded region is continuous. In some embodiments, the double-stranded region comprises at least one nick. In some embodiments, the double-stranded region may be separated by a single-stranded region. In some cases, the single-stranded region may be one or more nucleotides in length. For example, both ends (the 5' end and the 3' end) of a padlock probe may bind to the index sequence to form a double-stranded region with a nick separating the two ends of the padlock probe. In some embodiments, the index-binding sequence may be ligated after hybridization to the index sequence. In the example of the padlock probe, upon ligation of the 5' end to the 3' end, the padlock probe forms a circular nucleic acid. In some embodiments, a ligation reaction may be performed by a ligase. In some embodiments, the ligase may be a DNA ligase. In some embodiments, the ligase may be an RNA ligase. In some embodiments, the ligase may be a T4 ligase, a SplintR ligase, a T3 ligase, a T7 ligase, an TtcB ligase, or a combination thereof. In some embodiments, the ligation reaction may form a circular nucleic acid.

vi. Amplification reactions

[00175] Disclosed herein are amplification reactions useful for the detection of targets (e.g. RNA, DNA, or polypeptides) in situ. The amplification reactions may be used to increase a signal associated with a target. In some cases, the amplification reactions may generate multiple copies of a target, reverse complement or a target (e.g. a nucleic acid), one or more index sequences associated with a target, or a combination thereof. The amplification reaction may generate multiple copies of the sequence that is amplified. In some instances, the amplification reaction may generate multiple copies of the reverse complement of the sequence being amplified. The amplification reactions may comprise replication of a target or reverse complement thereof or an index sequence associated with the target or may comprise hybridization of additional nucleic acid molecules to generate additional binding sites for a detection moiety as described herein.

[00176] In some cases, one or more amplicons may be generated by the amplification reactions. The one or more amplicons may comprise one or more copies of an index sequence or reverse complement thereof associated with a target. The one or more amplicons may comprise one or more derivatives of the target, the index sequence, or a combination thereof. A

derivative of a nucleic acid for the methods described herein, e.g. the derivative of an index sequence, refers to the reverse complement of the nucleic acid sequence, e.g. the reverse complement of the index sequence. The one or more amplicons may comprise one or more copies of a portion of a target (e.g. an RNA or a DNA) or a reverse complement of a portion of a target (e.g. an RNA or a DNA). The one or more amplicons may comprise one or more copies of a portion of an oligonucleotides of a polypeptide-binding sequence or reverse complement thereof. The one or more amplicons may comprise one or more copies of a portion of an index-binding sequence or reverse complement thereof.

[00177] In some cases, the amplification reactions may comprise a rolling circle amplification reaction, a polymerase chain reaction (PCR), a quantitative polymerase chain reaction (qPCR), loop-mediated isothermal amplification (LAMP), recombinase polymerase amplification (RPA), strand displacement amplification (SDA), helicase-dependent amplification (HAD), multiple displacement amplification (MDA), multiple cross displacement amplification (MCDA), a transcription mediated amplification, a ligase chain reaction, or a combination thereof. The amplification reactions may include one or more temperature incubation steps. In some cases, the amplification reaction may take place at the same temperature (e.g. isothermally). In some cases, the amplification reaction may take place at multiple temperatures. For example, a rolling circle amplification reaction may be performed at 42°C for the duration of the amplification reaction. In some cases, the temperature may vary throughout a range during the amplification reaction. In some cases, an incubation temperature of the amplification reaction may be at least about 15°C, at least about 18°C, at least about 20°C, at least about 25°C, at least about 30°C, at least about 32°C, at least about 34°C, at least about 36°C, at least about 38°C, at least about 40°C, at least about 42°C, at least about 44°C, at least about 46°C, at least about 48°C, at least about 50°C, at least about 52°C, at least about 54°C, at least about 56°C, at least about 58°C, at least about 60°C, at least about 62°C, at least about 64°C, at least about 66°C, at least about 68°C, at least about 70°C, at least about 72°C, at least about 74°C, at least about 76°C, at least about 78°C, at least about 80°C, at least about 82°C, at least about 84°C, at least about 86°C, at least about 88°C, at least about 90°C, at least about 95°C, or more. In some cases, an incubation temperature of the amplification reaction may be at most about 15°C, at most about 18°C, at most about 20°C, at most about 25°C, at most about 30°C, at most about 32°C, at most about 34°C, at most about 36°C, at most about 38°C, at most about 40°C, at most about 42°C, at most about 44°C, at most about 46°C, at most about 48°C, at most about 50°C, at most about 52°C, at most about 54°C, at most about 56°C, at most about 58°C, at most about 60°C, at most about 62°C, at most about 64°C, at most about 66°C, at most about 68°C, at most about 70°C, at most about 72°C, at most about 74°C, at most about 76°C, at most about 78°C, at most about 80°C, at most about 82°C, at

most about 84°C, at most about 86°C, at most about 88°C, at most about 90°C, at most about 95°C, or more. In some cases, an incubation temperature of the amplification reaction may be at most about 18-90°C, about 20-88°C, about 25-86°C, about 30-84°C, about 32-82°C, about 34-80°C, about 36-78°C, about 38-76°C, about 40-74°C, about 42-72°C, about 44-70°C, about 46-68°C, about 48-66°C, about 50-64°C, about 52-62°C, about 54-60°C, or about 56-58 °C. The amplification reactions may be performed for a duration of time. In some cases, the duration of time of the amplification may be correlated to the amount of signal that is generated during the detection step the methods described herein. For example, in some cases, a longer amplification reaction may generate a larger signal associated with a target. The amplification reaction may be at least about 5 minutes, at least about 10 minutes, at least about 15 minutes, at least about 30 minutes, at least about 45 minutes, at least about 1 hour, at least about hours, at least about 3 hours, at least about 4 hours, at least about 5 hours, at least about 6 hours, at least about 7 hours, at least about 8 hours, at least about 9 hours, at least about 10 hours, at least about 12 hours, at least about 15 hours, at least about 20 hours, at least about 24 hours, at least about 36 hours, at least about 48 hours, or more. The amplification reaction may be at most about 5 minutes, at most about 10 minutes, at most about 15 minutes, at most about 30 minutes, at most about 45 minutes, at most about 1 hour, at most about hours, at most about 3 hours, at most about 4 hours, at most about 5 hours, at most about 6 hours, at most about 7 hours, at most about 8 hours, at most about 9 hours, at most about 10 hours, at most about 12 hours, at most about 15 hours, at most about 20 hours, at most about 24 hours, at most about 36 hours, at most about 48 hours, or less. The amplification reaction may be about 5 minutes-48 hours, about 10 minutes-36 hours, about 15 minutes-24 hours, about 30 minutes-20 hours, about 45 minutes-15 hours, about 1 hour -12 hours, about 2 hours-10 hours, about 3 hours-9 hours, about 4 hours-8 hours, or about 5 hours-7 hours.

[00178] In some cases, a rolling circle amplification reaction may be performed. The rolling circle amplification may comprise amplifying a circular nucleic acid. The circular nucleic acid may comprise one or more nucleic acid molecules as described herein. For example, a padlock probe may be ligated to form a circular nucleic acid that is used during a rolling circle amplification reaction. The padlock probe that forms a circular nucleic acid may be a probe comprising a DNA-binding sequence and an index sequence, a probe comprising an RNA-binding sequence and an index sequence, an oligonucleotide-binding probe and an index sequence, or an index-binding sequence. The rolling circle amplification rection may generate one or more amplicons. An amplicon of the one or more amplicons may comprise one or more copies of the component of the circular nucleic acid. For example, the circular nucleic acid may be generated by ligating a padlock probe that comprises a DNA-binding sequence an index sequence. A rolling circle amplification reaction may be performed using this circular nucleic acid to generate one or

more amplicons that comprise one or more copies of the DNA-binding sequence or reverse complement thereof and one or more copies of the index sequence or reverse complement thereof. In another example, the circular nucleic acid may be generated by ligating a padlock probe that comprises an RNA-binding sequence an index sequence. A rolling circle amplification reaction may be performed using this circular nucleic acid to generate one or more amplicons that comprise one or more copies of the RNA-binding sequence or reverse complement thereof and one or more copies of the index sequence or reverse complement thereof. In another example, the circular nucleic acid may be generated by ligating a padlock probe that comprises an oligonucleotide-binding probe an index sequence. A rolling circle amplification reaction may be performed using this circular nucleic acid to generate one or more amplicons that comprise one or more copies of the oligonucleotide-binding probe or reverse complement thereof and one or more copies of the index sequence or reverse complement thereof. In another example, the circular nucleic acid may be generated by ligating a padlock probe that comprises an index-binding sequence an optionally an index sequence. A rolling circle amplification reaction may be performed using this circular nucleic acid to generate one or more amplicons that comprise one or more copies of the index-binding sequence or reverse complement thereof and optionally one or more copies of the index sequence or reverse complement thereof.

[00179] In some embodiments, the amplification reaction is performed by a polymerizing enzyme. In some embodiments, the polymerizing enzyme may be a polymerase. In some embodiments, the polymerase may be a DNA polymerase. In some embodiments, the polymerase may be an RNA polymerase. In some embodiments, the polymerase may have strand-displacement capabilities. In some embodiments, the polymerase may be phi29 polymerase, taq polymerase, T4 polymerase, T7 polymerase, or a combination thereof. For example, the amplification reaction may comprise a rolling circle amplification reaction and a phi29 polymerase may be used to generate one or more amplicons.

[00180] In some embodiments, the rolling circle amplification reaction comprises contacting the covalently circular nucleic acids with a soluble amplification primer (e.g., a universal rolling circle amplification primer), a strand-displacing DNA polymerase, a plurality of nucleotides, or a combination thereof, under a condition suitable for hybridizing individual amplification primers to a circular nucleic acids, and under a condition suitable for conducting primer extension using the circular nucleic acids as a template molecule to generate one or more amplicons. In some embodiments, the plurality of nucleotides in the rolling circle amplification reaction may comprise any mixture of two or more of dATP, dGTP, dCTP, dTTP and/or dUTP. In some embodiments, any of the rolling circle amplification reactions described herein can be conducted in the presence or in the absence of a plurality of compaction oligonucleotides.

[00181] In some embodiments, when the rolling circle amplification reaction includes a plurality of nucleotide which includes dUTP, the resulting amplicon can be cross-linked to a cross-linking reactive group by treating the biological sample with a succinimide ester (NHS), maleimide (Sulfo-SMCC), imidoester (DMP), carbodiimide (DCC, EDC) or phenyl azide. In some embodiments, polymerization of the cross-linking reactive group can be initiated with light or UV light. In some embodiments, the resulting amplicon can be cross-linked to a matrix by treating the biological sample with a cross-linked agarose, cross-linked dextran or cross-linked polyethylene glycol (PEG), polyacrylamide, cellulose alginate or polyamide. In some embodiments, the PEG comprises a sulfo-NHS ester moiety at one or both ends, for example a PEGylated bis(sulfosuccinimidyl)suberate) (e.g., $BS(PEG)_9$ from Thermo Fisher Scientific, catalog No. 21582).

[00182] In some embodiments, the rolling circle amplification reaction can be conducted at a constant temperature (e.g., isothermal) wherein the constant temperature is at room temperature to about 30 °C, or about 30 – 40 °C, or about 40 – 50 °C, or about 50 - 65 °C.

[00183] In some embodiments, the DNA polymerase having a strand displacing activity 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).

[00184] In some embodiments, the rolling circle amplification primers can be modified to increase resistance to nuclease degradation. In some embodiments, the rolling circle amplification primers comprise at least one phosphorothioate diester bond at their 5' ends which can render the amplification primers resistant to exonuclease degradation. In some embodiments, the rolling circle amplification primers comprise 2-5 or more consecutive phosphorothioate diester bonds at their 5' ends. In some embodiments, the rolling circle amplification primers comprise at least one ribonucleotide and/or at least one 2'-O-methyl or 2'-O-methoxyethyl (MOE) nucleotide.

[00185] In some embodiments, the rolling circle amplification reaction can be conducted in the presence of a plurality of compaction oligonucleotides which, when hybridized to an amplicon, compacts the size and/or shape of the amplicon 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 an amplicon and a second region at

the other end that hybridizes to another portion of the same amplicon, where hybridization of the compaction oligonucleotide to a given amplicon compacts the size and/or shape of the amplicon.

[00186] 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 amplicon to pull together distal portions of the amplicon causing compaction of the amplicon to form a DNA nanoball. For example, the 5' region of the compaction oligonucleotide is designed to hybridize to a first portion of the amplicon (e.g., a universal compaction oligonucleotide binding site), and the 3' region of the compaction oligonucleotide is designed to hybridized to a second portion of the amplicon (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 amplicons generated in the absence of the compaction oligonucleotides. The compact and stable characteristics of the DNA nanoballs improves *in situ* sequencing accuracy by increasing signal intensity and the nanoballs retain their shape and size during multiple sequencing cycles.

[00187] 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, or 30-100 nucleotides, or 40-80 nucleotides in length.

[00188] 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. 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.

[00189] The 5' region of the compaction oligonucleotides can be wholly complementary or partially complementary along its length to a first portion of an amplicon. The 3' region of the compaction oligonucleotides can be wholly complementary or partially complementary along its length to a second portion of an amplicon. The 5' region of the compaction oligonucleotides can hybridize to a first universal sequence portion of an amplicon. The 3' region of the compaction oligonucleotides can hybridize to a second universal sequence portion of an amplicon.

[00190] 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.

[00191] 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.

[00192] 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 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 an amplicon. In some embodiments, the compaction oligonucleotides comprise a phosphorylated 5' end (e.g., using a polynucleotide kinase).

[00193] 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.

[00194] 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. 18). 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.

[00195] At least one compaction oligonucleotide can form a guanine tetrad (Figure 12) and hybridize to the universal binding sequences in an amplicon which can cause the amplicon to fold to form an intramolecular G-quadruplex structure (Figure 13). The amplicons 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 biological sample.

[00196] In some embodiments, the plurality of compaction oligonucleotides in the rolling circle amplification reaction have the same sequence. Alternatively, the plurality of compaction oligonucleotides in the rolling circle amplification reaction comprise a mixture of two or more different populations of compaction oligonucleotides having different sequences.

[00197] In some embodiment, the immobilized amplicon 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.

[00198] 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 biological 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 um or smaller.

vii. Reaction conditions

[00199] Disclosed herein, are methods that relate to detection of one or more targets (e.g. nucleic acids or polypeptides) *in situ* under certain reaction conditions. In some cases, one or more nucleic acid molecules (e.g. probes) may be introduced to the biological sample comprising the one or more targets under conditions sufficient to couple a binding sequence of a nucleic molecule (e.g. an RNA-binding sequence, a DNA-binding sequence, or a target-binding sequence) to one or more targets (e.g. RNA or DNA). In some cases, one or more nucleic acid molecules (e.g. a probe or an index-binding molecule) may be introduced to the biological sample comprising one or more targets under sufficient conditions to couple a binding sequence of the nucleic acid molecule (e.g. an oligonucleotide-binding probe or an index-binding sequence) to one or more probes or nucleic acid molecules (e.g. one or more polypeptide-binding probes or one or more

target-binding sequences) that are bound to a target (e.g. a polypeptide or an RNA). In some cases, one or more primer nucleic acid sequences may be introduced to the biological sample comprising one or more targets under conditions sufficient to couple the one or more primer nucleic acid sequences to one or more index sequences or the reverse complement thereof. Conditions sufficient to couple 1) the binding sequences as described herein to the one or more targets 2) nucleic acid molecules to polypeptide-binding probes or target-binding sequences, and/or 3) one or more primer nucleic acid sequences to one or more index sequences or the reverse complement thereof may include a reagent, a temperature, an incubation time, a concentration, or a combination thereof.

[00200] In some embodiments, Conditions sufficient to couple 1) the binding sequences as described herein to the one or more targets 2) nucleic acid molecules to polypeptide-binding probes or target-binding sequences, and/or 3) one or more primer nucleic acid sequences to one or more index sequences or the reverse complement thereof may include a regent. In some embodiments, the reagent may comprise a salt. In some embodiments, the salt may comprise NaCl, $MgCl_2$, $CaCl_2$, KCl, Na_2SO_4, or a combination thereof. In some further embodiments, the coupling conditions comprise the presence of strontium ions, magnesium ions, and/or calcium ions. In some embodiments, the reagent may comprise a buffer. In some embodiments, the buffer may comprise a phosphate buffered saline, tris, ethylenediaminetetraacetic acid, or a combination thereof.

[00201] In some embodiments, the reagent may comprise a solvent. In some embodiments, the solvent may be a polar solvent. In some embodiments, the solvent may be a nonpolar solvent. In some embodiments, the solvent may be an aprotic solvent. In some embodiments, the solvent may be a protic solvent.

[00202] In some embodiments, the reagent may comprise a chaotropic reagent. In some embodiments, the chaotropic reagent may comprise glycine, arginine, histidine hydrochloride, formamide, dimethyl sulfoxide, dimethylformamide, urea, thiourea, guanidine hydrochloride, or a combination thereof.

[00203] In some embodiments Conditions sufficient to couple 1) the binding sequences as described herein to the one or more targets 2) nucleic acid molecules to polypeptide-binding probes or target-binding sequences, and/or 3) one or more primer nucleic acid sequences to one or more index sequences or the reverse complement thereof may include a temperature. The temperature may be an incubation temperature maintained during the time sufficient to couple the target-binding sequence to the target nucleic acid. The temperature may be about 10-74°C, about °C, about 12-72°C, about 14-70°C, about 16-68°C, about 18-66°C, about 20-64°C, about 22-62°C, about 24-60°C, about 26-58°C, about 28-56°C, about 30-54°C, about 32-52°C, about

34-50°C, about 36-48°C, about 38-46°C, about 40-44°C, at least about 10°C, at least about12°C, at least about14°C, at least about16°C, at least about18°C, at least about20°C, at least about22°C, at least about24°C, at least about26°C, at least about28°C, at least about30°C, at least about32°C, at least about34°C, at least about36°C, at least about38°C, at least about40°C, at least about42°C, at least about44°C, at least about46°C, at least about48°C, at least about50°C, at least about52°C, at least about54°C, at least about56°C, at least about58°C, at least about60°C, at least about62°C, at least about64°C, at least about66°C, at least about68°C, at least about70°C, at least about72°C, or at least about74°C.

[00204] In some embodiments Conditions sufficient to couple 1) the binding sequences as described herein to the one or more targets 2) nucleic acid molecules to polypeptide-binding probes or target-binding sequences, and/or 3) one or more primer nucleic acid sequences to one or more index sequences or the reverse complement thereof may include an incubation time. The incubation time may be a time sufficient to couple the target-binding sequence to the target nucleic acid. The incubation time may be at least about 10 minutes, at least about 10 minutes, at least about 20 minutes, at least about 30 minutes, at least about 40 minutes, at least about 50 minutes, at least about 55 minutes, at least about 1 hours, at least about 1.5 hours, at least about 2 hours, at least about 2.5 hours, at least about 3 hours, at least about 3.5 hours, at least about 4 hours, at least about 4.5 hours, at least about 5 hours, at least about 5.5 hours, at least about 6 hours, at least about 6.5 hours, at least about 7 hours, at least about 7.5 hours, at least about 8 hours, at least about 8.5 hours, at least about 9 hours, at least about 9.5 hours, at least about 10 hours, at least about 10.5 hours, at least about 11 hours, at least about 11.5 hours, or at least about 12 hours.

[00205] In some embodiments, Conditions sufficient to couple 1) the binding sequences as described herein to the one or more targets 2) nucleic acid molecules to polypeptide-binding probes or target-binding sequences, and/or 3) one or more primer nucleic acid sequences to one or more index sequences or the reverse complement thereof may include a concentration. The concentration may relate to the concentration of the one or more nucleic acid molecules. In some embodiments the concentration may be about 0.1-10000 nM, about 0.2-5000 nM, about 0.3-1000 nM, about 0.4-900 nM, about 0.5-800 nM, about 0.6-700 nM, about 0.7-600 nM, about 0.8-500 nM, about 0.9-400 nM, about 1-300 nM, about 2-250 nM, about 3-200 nM, about 4-150 nM, about 5-100 nM, about 6-50 nM, about 7-25 nM, about 8-20 nM, about 9-15 nM, at least about 0.1 nM, at least about 0.2 nM, at least about 0.3 nM, at least about 0.4 nM, at least about 0.5 nM, at least about 0.6 nM, at least about 0.7 nM, at least about 0.8 nM, at least about 0.9 nM, at least about 1 nM, at least about 2 nM, at least about 3 nM, at least about 4 nM, at least about 5 nM, at least about 6 nM, at least about 7 nM, at least about 8 nM, at least about 9 nM, at least about 10

nM, at least about 15 nM, at least about 20 nM, at least about 25 nM, at least about 50 nM, at least about 100 nM, at least about 150 nM, at least about 200 nM, at least about 250 nM, at least about 300 nM, at least about 400 nM, at least about 500 nM, at least about 600 nM, at least about 700 nM, at least about 800 nM, at least about 900 nM, at least about 1000 nM, at least about 5000 nM, or at least about 10000 nM. In some embodiments, the concentration of the one or more nucleic acid molecules may be at most about at most about 0.1 nM, at most about 0.2 nM, at most about 0.3 nM, at most about 0.4 nM, at most about 0.5 nM, at most about 0.6 nM, at most about 0.7 nM, at most about 0.8 nM, at most about 0.9 nM, at most about 1 nM, at most about 2 nM, at most about 3 nM, at most about 4 nM, at most about 5 nM, at most about 6 nM, at most about 7 nM, at most about 8 nM, at most about 9 nM, at most about 10 nM, at most about 15 nM, at most about 20 nM, at most about 25 nM, at most about 50 nM, at most about 100 nM, at most about 150 nM, at most about 200 nM, at most about 250 nM, at most about 300 nM, at most about 400 nM, at most about 500 nM, at most about 600 nM, at most about 700 nM, at most about 800 nM, at most about 900 nM, at most about 1000 nM, at most about 5000 nM, or at most about 10000 nM

viii. Primer nucleic acid sequences

[00206] Disclosed herein are primer nucleic acid sequences which are useful in the methods described herein for detecting a target (e.g., RNA) in situ. The methods disclosed herein may include one or more primer nucleic acid sequences. The one or more primer nucleic acid sequence may be configured to bind to an index sequence a reverse complement of an index sequence. The one or more primer nucleic acid sequences may comprise DNA, RNA, or a combination thereof. The primer nucleic acid sequence may be configured to couple to an index sequence. In some embodiments, the primer nucleic acid sequence may hybridize to an index sequence. In some embodiments, the primer nucleic acid may comprise at least one binding region that recognizes and binds at least a portion of an index sequence. In some embodiments, the primer nucleic acid may comprise more than one binding region that recognizes and binds an index sequence. The one or more primer nucleic acid sequences may form double-stranded region when bound to an index sequence or reverse complement thereof, as described herein. The one or more primer nucleic acid sequences may bind directly to a nucleic acid molecule as described herein. For example, a single-stranded oligonucleotide may comprise a sequence that is complementary to a portion of an index sequence and form a double-stranded nucleic acid region upon hybridization of the single-stranded oligonucleotide to the index sequence. In some embodiments, the one or more primer nucleic acid sequence may bind to an index-binding sequence as described herein. In some embodiments, the one or more primer nucleic acid sequences may bind to an amplification product generated from an amplification reaction as described herein. In some

embodiments, the primer nucleic acid sequence binds to at least a portion of an index sequence. In some embodiments, the primer nucleic acid sequence binds to at least a portion of the reverse complement of an index sequence. In some embodiments, the one or more primer nucleic acid sequences binds directly to the target nucleic acid, as described herein. In some embodiments, the primer nucleic acid sequence is a sequencing primer. In some embodiments, the primer nucleic acid sequence is an amplification primer.

[00207] The one or more primer nucleic acid sequences may comprise nucleic acid molecules with a given length. Each of the one or more primer nucleic acid sequences may be the same length or different lengths. In some embodiments, at least two of the one or more primer nucleic acid sequences have the same length. In some embodiments, at least two of the one or more primer nucleic acid sequences have different lengths. The one or more primer nucleic acid sequences may have a variety of lengths. In some embodiments, the one or more primer nucleic acid sequences has a length of 1-100 nucleotides, about 2-98 nucleotides, about 3-96 nucleotides, about 4-94 nucleotides, about 5-92 nucleotides, about 6-90 nucleotides, about 7-88 nucleotides, about 8-86 nucleotides, about 9-84 nucleotides, about 10-82 nucleotides, about 11-80 nucleotides, about 12-78 nucleotides, about 13-76 nucleotides, about 14-74 nucleotides, about 15-72 nucleotides, about 16-70 nucleotides, about 17-68 nucleotides, about 18-66 nucleotides, about 19-64 nucleotides, about 20-62 nucleotides, about 22-60 nucleotides, about 24-58 nucleotides, about 26-56 nucleotides, about 28-54 nucleotides, about 30-52 nucleotides, about 32-50 nucleotides, about 34-48 nucleotides, about 36-46 nucleotides, about 38-44 nucleotides, about 40-42 nucleotides, at least about 1 nucleotide, at least about 2 nucleotides, at least about 3 nucleotides, at least about 4 nucleotides, at least about 5 nucleotides, at least about 6 nucleotides, at least about 7 nucleotides, at least about 8 nucleotides, at least about 9 nucleotides, at least about 10 nucleotides, at least about 11 nucleotides, at least about 12 nucleotides, at least about 13 nucleotides, at least about 14 nucleotides, at least about 15 nucleotides, at least about 16 nucleotides, at least about 17 nucleotides, at least about 18 nucleotides, at least about 19 nucleotides, at least about 20 nucleotides, at least about 22 nucleotides, at least about 24 nucleotides, at least about 26 nucleotides, at least about 28 nucleotides, at least about 30 nucleotides, at least about 32 nucleotides, at least about 34 nucleotides, at least about 36 nucleotides, at least about 38 nucleotides, at least about 40 nucleotides, at least about 42 nucleotides, at least about 44 nucleotides, at least about 46 nucleotides, at least about 48 nucleotides, at least about 50 nucleotides, at least about 52 nucleotides, at least about 54 nucleotides, at least about 56 nucleotides, at least about 58 nucleotides, at least about 60 nucleotides, at least about 62 nucleotides, at least about 64 nucleotides, at least about 66 nucleotides, at least about 68 nucleotides, at least about 70 nucleotides, at least about 72

nucleotides, at least about 74 nucleotides, at least about 76 nucleotides, at least about 78 nucleotides, at least about 80 nucleotides, at least about 82 nucleotides, at least about 84 nucleotides, at least about 86 nucleotides, at least about 88 nucleotides, at least about 90 nucleotides, at least about 92 nucleotides, at least about 94 nucleotides, at least about 96 nucleotides, at least about 98 nucleotides, or at least about 100 nucleotides in length.

[00208] In some embodiments, the one or more primer nucleic acid sequences inhibits incorporation of another nucleotide into the primer nucleic acid sequence. In some embodiments, the one or more primer nucleic acid sequences does not include a 3' hydroxyl group. In some embodiments, the one or more primer nucleic acid sequences may comprise a 3' hydroxyl group substituted with a chain terminating moiety. In some embodiments, the one or more primer nucleic acid sequences comprises a blocking group. In some embodiments, the blocking group may comprise alkyl group, alkenyl group, alkynyl group, allyl group, aryl group, benzyl group, azide group, amine group, amide group, keto group, isocyanate group, phosphate group, thiol group, disulfide group, carbonate group, urea group, silyl group, or a combination thereof.

[00209] In some embodiments, coupling of a primer nucleic acid sequence to an index or a reverse complement of an index generates a primed index sequence. The primed index sequence may comprise a single-stranded nucleic acid, a double-stranded nucleic acid, or a combination thereof. In some embodiments, the primed index sequence is formed between a nucleic acid molecule and a primer nucleic acid sequence, as described herein. In some embodiments, the primed index sequence is formed between a target molecule and a primer nucleic acid sequence, as described herein. In some cases, the primed index sequence is formed between an amplicon and a primer nucleic acid sequence, as described herein.

[00210] In some embodiments, the primer nucleic acid sequence may bind to at least a portion of the one or more amplicons described herein. For example, the primer nucleic acid sequence may bind to a portion of one or more index sequences within the one or more amplicons to generate a primed index sequence. In some cases, the primer nucleic acid sequence may bind to a portion of the one more amplicons that is proximal to one or more index sequences to form a primed index sequences. For example, the portion of the one or more applicants that is proximal to the one or more index sequences may be upstream or downstream from the one or more index sequences and when the primer nucleic acid sequence binds to the portion of the one or more amplicons that is proximal to the one or more index sequences it may form a double-stranded nucleic acid sequence proximal to the one or more index sequences of the one or more amplicons.

[00211] In some embodiments, a primer nucleic acid sequence may bind directly to a target to form a primed index sequence. FIG. 1C illustrates a non-limiting step of annealing a

primer 103 to the target 102 to form a primed index sequence 104. FIG. 1B only shows one primer being used in the annealing step, but more than one primer may be used depending on the types of target nucleic acid. In some instances, an adapter that is used to attach the target to the surface may have the same sequence as the primer used to prepare the primed target nucleic acid. The primer may comprise forward amplification primers, reverse amplification primers, sequencing primers, and/or molecular barcoding sequences, or any combination thereof. In some instances, one primer sequence may be used in the hybridization step. In some instances, a plurality of different primer sequences may be used in the hybridization step. As shown in Fig. 1D, the primed sequence 104 may be combined with a nucleotide conjugate (e.g. detectable nucleotide conjugate)and a polymerase 106 to form a binding complex.

[00212] The methods described herein involve detection of a target using a detectable nucleotide conjugate. In some embodiments, the target is a ribonucleic acid. In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)is introduced to the biological sample comprising the target under conditions sufficient to form binding complex.

[00213] In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises a detectable label. In some embodiments, the detectable label may be conjugated to the nucleotide conjugate (e.g. detectable nucleotide conjugate)via covalent bond. In some embodiments, the detectable label may be conjugated to the nucleotide conjugate (e.g. detectable nucleotide conjugate)via non-covalent bond. In some embodiments, the detectable label may comprise a quantum dot. In some embodiments, the detectable label may comprise a fluorophore. Example fluorescent moieties include, but are not limited to, fluorescein and fluorescein derivatives such as carboxyfluorescein, tetrachlorofluorescein, hexachlorofluorescein, carboxynapthofluorescein, 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.

[00214] In some embodiments, the detection label may be a FRET pair, 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.

[00215] In some embodiments, the one or more nucleotide conjugates (e.g. detectable nucleotide conjugates) are labeled with fluorophores and the detecting step comprises

use of fluorescence imaging; and especially wherein the fluorescence imaging comprises dual wavelength excitation/four wavelength emission fluorescence imaging. In some embodiments, the four different nucleotide conjugates (e.g. detectable nucleotide conjugates), each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein the four different nucleotide conjugates (e.g. detectable nucleotide conjugates) are labeled with separate respective fluorophores, and wherein the detecting step comprises simultaneous 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, four different nucleotide conjugates (e.g. detectable nucleotide conjugates), each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein the four different nucleotide conjugates (e.g. detectable nucleotide conjugates) are labeled with Cy3, Cy3.5, Cy5, and Cy5.5 respectively, and wherein 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; and/or wherein the fluorescence imaging comprises dual wavelength excitation / dual wavelength emission fluorescence imaging. In some embodiments, the four different nucleotide conjugates (e.g. detectable nucleotide conjugates), each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein one, two, three, or four different nucleotide conjugates (e.g. detectable nucleotide conjugates) are respectively labeled, each with a with distinct fluorophore or set of fluorophores, and wherein the detecting step comprises simultaneous excitation at a wavelength sufficient to excite one, two, three, or four fluorophores or sets of fluorophores, and imaging of fluorescence emission at wavelengths sufficient to detect each respective fluorophore. In some embodiments, the three different nucleotide conjugates (e.g. detectable nucleotide conjugates), each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein one, two, or three different nucleotide conjugates (e.g. detectable nucleotide conjugates) are respectively labeled, each with a with distinct fluorophore or set of fluorophores, and wherein the detecting step comprises simultaneous excitation at a wavelength sufficient to excite one, two, or three, fluorophores or sets of fluorophores, and imaging of fluorescence emission at wavelengths sufficient to detect each respective fluorophore, and wherein detection of the fourth nucleotide is determined or determinable with reference to the location of "dark" or unlabeled spots or target nucleotides. In some embodiments, the binding complex consists of three types of nucleotide conjugates (e.g. detectable nucleotide conjugates) and wherein each type of the three types of nucleotide conjugates (e.g. detectable nucleotide conjugates) comprises a

different type of nucleotide. In some embodiments, the detection of the binding complex is performed in the absence of unbound or solution-borne polymer nucleotide conjugates.

[00216] The non-limiting example of the nucleotide conjugate (e.g. detectable nucleotide conjugate)in FIG. 1D comprises four nucleotide conjugates (e.g. detectable nucleotide conjugates) 105a, 105b, 105c, and 105d. Each nucleotide conjugate (e.g. detectable nucleotide conjugate)has multiple copies of a nucleotide attached to the particle, and the four nucleotide conjugates (e.g. detectable nucleotide conjugates) cover four types of nucleotide respectively. The nucleotide conjugate (e.g. detectable nucleotide conjugate)having a nucleotide that is complementary to the next base on the target may form a binding complex with the polymerase and the target nucleic acid. In some instances, the binding complex may include one, two or three nucleotide conjugates (e.g. detectable nucleotide conjugates). In some embodiments, each different type of nucleotide conjugate (e.g. detectable nucleotide conjugate)can be labeled with a separate label. In some embodiments, three of four types of nucleotide conjugates can be labeled, with a fourth either unlabeled or conjugated to an undetectable label. In some embodiments, 1, 2, 3, or 4 nucleotide conjugates (e.g. detectable nucleotide conjugates) can be labeled, either with the same label, or each with a label corresponding to the identity of its conjugated nucleotide, with, respectively, 3, 2, 1, or no nucleotide conjugates (e.g. detectable nucleotide conjugates) that may be either left unlabeled or conjugated to an undetectable label. In some embodiments, detection of a polymerase complex incorporating a nucleotide conjugate (e.g. detectable nucleotide conjugate)may be carried out using four-color detection, such that conjugates corresponding to all four nucleotides are present in a sample, each conjugate having a separate label corresponding to the nucleotide conjugated thereto. In some embodiments, the four nucleotide conjugates (e.g. detectable nucleotide conjugates) may be exposed to or contacted with the target at the same time; in some other embodiments, the four nucleotide conjugates (e.g. detectable nucleotide conjugates) may be exposed to or contacted with the target sequentially, either individually, or in groups of two or three. In some embodiments, detection of a polymerase complex incorporating a nucleotide conjugate (e.g. detectable nucleotide conjugate)may be carried out using three-color detection, such that conjugates corresponding to three of the four nucleotides are present in a sample, with three conjugates having a separate label corresponding to the nucleotide conjugated thereto and one conjugate having no label or being conjugated to an undetectable label. In some embodiments, only three types of conjugates are provided, such that conjugates corresponding to three of the four nucleotides are present in a sample, with three conjugates having a separate label corresponding to the nucleotide conjugated thereto and one conjugate being absent. In some embodiments, the identity of nucleotides corresponding to an unlabeled or absent nucleotide conjugate can be determined with respect to the location and/or

identity of "dark" spots or locations of known targets showing no fluorescence signal. In some embodiments, the detection of the binding complex is performed in the absence of unbound or solution-borne polymer nucleotide conjugates.

[00217] In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises one or more nucleotides or nucleotide analogs. In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise a single type of nucleotide. In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)comprise more than one type of nucleotide. In some embodiments, the nucleotide comprises ATP, ADP, AMP, dATP, dADP, dAMP, TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, dUMP, CTP, CDP, CMP, dCTP, dCDP, dCMP, GTP, GDP, GMP, dGTP, dGDP, or dGMP.

[00218] In some embodiments, the present disclosure provides a reagent comprising one or more nucleotide conjugates (e.g. detectable nucleotide conjugates) as disclosed herein and a buffer. For example, in some embodiments, the present disclosure provides a reagent, wherein said reagent comprises 1, 2, 3, 4, or more nucleotide conjugates (e.g. detectable nucleotide conjugates), wherein each nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises a single type of nucleotide. In some embodiments, a reagent of the present disclosure comprises 1, 2, 3, 4, or more nucleotide conjugates (e.g. detectable nucleotide conjugates), wherein each nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises a single type of nucleotide or nucleotide analog, and wherein said nucleotide or nucleotide moiety may respectively correspond to one or more from the group consisting of ATP, ADP, AMP, dATP, dADP, and dAMP; one or more from the group consisting of TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, and dUMP; one or more from the group consisting of CTP, CDP, CMP, dCTP, dCDP, and dCMP; and one or more from the group consisting of GTP, GDP, GMP, dGTP, dGDP, and dGMP. In some other examples or some further examples, the present disclosure provides a reagent comprising or further comprising 1, 2, 3, 4, or more nucleotide conjugates (e.g. detectable nucleotide conjugates), wherein each nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises a single type of nucleotide or nucleotide moiety, and wherein said nucleotide or nucleotide moiety may respectively correspond to one or more from the group consisting of ATP, ADP, AMP, dATP, dADP, dAMP, TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, dUMP, CTP, CDP, CMP, dCTP, dCDP, dCMP, GTP, GDP, GMP, dGTP, dGDP, and dGMP.

[00219] In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)is included in a mixture of a plurality of nucleotide conjugates (e.g. detectable nucleotide conjugates), wherein each of the plurality of nucleotide conjugates (e.g. detectable

nucleotide conjugates) comprises different types of nucleotide moieties from each other. The type of nucleotide may be ATP, ADP, AMP, dATP, dADP, dAMP, TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, dUMP, CTP, CDP, CMP, dCTP, dCDP, dCMP, GTP, GDP, GMP, dGTP, dGDP, dGMP, or a combination thereof.

[00220] In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises a polymer having a plurality of branches. In some embodiments, the plurality of branches comprises copies of a nucleotide coupled thereto. In some embodiments, the polymer has a star, comb, cross-linked, bottle brush, or dendrimer configuration. In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises one or more binding groups comprising avidin, biotin, affinity tag, or any combination thereof. In some embodiments, a binding group is interposed between a branch of the polymer and a nucleotide.

[00221] In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise a common core. The common core may comprise multiple copies of the same nucleotide covalently bound to or noncovalently bound to the common core. Examples of the common core can include a polymer particle; a metal particle; a quantum dot; a liposome; an emulsion particle, or any other particle (e.g., nanoparticles, microparticles, or the like). In some embodiments, the polymer particle is or comprises a branched polymer; a dendrimer; a cross linked polymer particle such as an agarose, polyacrylamide, acrylate, methacrylate, cyanoacrylate, methyl methacrylate particle; a glass particle; a ceramic particle; or a protein particle. In some embodiment, the common core is a branched polymer. In some embodiments, the common core is or comprises a protein. In some embodiments, the protein is streptavidin, avidin or a combination thereof. In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises multiple polymer branches radiating from the common core, wherein some branches are attached to a nucleotide and/or some branches are attached to an oligonucleotide.

[00222] In some embodiments, the present disclosure provides a nucleotide conjugate (e.g. detectable nucleotide conjugate)comprising a branched polymer having two or more branches and two or more copies of a nucleotide, wherein said nucleotide is attached to a first plurality of said branches or arms, and optionally, wherein one or more interaction moieties are attached to a second plurality of said branches or arms. In some embodiments, said nucleotide conjugate (e.g. detectable nucleotide conjugate)may further comprise one or more labels on the polymer. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)wherein the nucleoside has a surface density of at least 4 nucleotides per polymer. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)comprising or incorporating a nucleotide or nucleotide analog that is modified so as to prevent its incorporation into an extending nucleic acid

chain during a polymerase reaction. In some embodiments, said nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise or incorporate a nucleotide or nucleotide analog that is reversibly modified so as to prevent its incorporation into an extending nucleic acid chain during a polymerase reaction. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)wherein one or more labels comprise a fluorescent label, a FRET donor, and/or a FRET acceptor. In some embodiments, said nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more branches or arms, or 2, 4, 8, 16, 32, 64, or more, branches or arms. In some embodiments, the branches or arms may radiate from a central moiety. In some embodiments, said nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise one or more interaction moieties, which interaction moieties may comprise avidin or streptavidin; a biotin moiety; an affinity tag; an enzyme, antibody, minibody, receptor, or other protein; a non-protein tag; a metal affinity tag, or any combination thereof. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)wherein the polymer comprises polyethylene glycol, polypropylene glycol, polyvinyl acetate, polylactic acid, or polyglycolic acid. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)wherein the nucleotide or nucleotide analog is attached to the branch or arm through a linker; and especially wherein the linker comprises PEG, and wherein the PEG moiety has an average molecular weight of about 1K, about 2K, about 3K, about 4K, about 5K, about 10K, about 15K, about 20K, about 50K, about 100K, about 150K, or about 200K, or greater than about 200K. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)wherein the linker comprises PEG, and wherein the PEG moiety has an average molecular weight of between about 5K and about 20K. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)wherein at least one nucleotide or nucleotide analog comprises a deoxyribonucleotide, a ribonucleotide, a deoxyribonucleoside, or a ribonucleoside; and/or wherein the nucleotide or nucleotide analog is conjugated to the linker through the 5' end of the nucleotide or nucleotide analog. In some embodiments, the present disclosure provides said nucleotide conjugate (e.g. detectable nucleotide conjugate)wherein one of the nucleotides or nucleotide analogs comprises deoxyadenosine, deoxyguanosine, thymidine, deoxyuridine, deoxycytidine, adenosine, guanosine, 5-methyl-uridine, and/or cytidine; and wherein the length of the linker is between 1 and 1,000nm.

[00223] 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 a group consisting of dATP, dGTP, dCTP, dTTP and dUTP. The plurality of nucleotides can comprise at a mixture of any combination of two or more types of nucleotides selected from a 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.

[00224] 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 phosphoramide 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 BH_3. In some embodiments, the chain includes phosphate groups substituted with analogs including phosphoramidate, phosphorothioate, phosphordithioate, and O-methylphosphoroamidite groups.

[00225] 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 moiety 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, 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) ($Pd(PPh_3)_4$) 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 (K_2CO_3) 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.

[00226] 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(hydroxyproyl)phosphine (THPP). In some embodiments, the cleaving agent comprises 4-dimethylaminopyridine (4-DMAP).

[00227] 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'-sulfhydral, 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.

[00228] 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.

[00229] 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, 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) ($Pd(PPh_3)_4$) 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 H2 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.

[00230] 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(hydroxyproyl)phosphine (THPP). In some embodiments, the cleaving agent comprises 4-dimethylaminopyridine (4-DMAP).

[00231] 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.

[00232] The nucleotide can be linked to the common core through a linker, and the nucleotide can be attached to one end or location of a polymer. The nucleotide can be conjugated to the common core through the 5' end of the nucleotide. In some common core-nucleotide conjugates, one nucleotide attached to one end or location of a polymer. In some common core-nucleotide conjugate, multiple nucleotides are attached to one end or location of a polymer. The conjugated nucleotide is sterically accessible to one or more proteins, one or more enzymes, and nucleotide binding moieties. In some embodiments, a nucleotide may be provided separately from a nucleotide binding moiety such as a polymerase. In some embodiments, the linker does not comprise a photo emitting or photo absorbing group.

[00233] The common core can also have a binding moiety. In some embodiments, common cores may self-associate without the use of a separate interaction moiety. In some embodiments, common cores may self-associate due to buffer conditions or salt conditions, e.g., as in the case of calcium-mediated interactions of hydroxyapatite particles, lipid or polymer mediated interactions of micelles or liposomes, or salt-mediated aggregation of metallic (such as iron or gold) nanoparticles.

[00234] The nucleotide conjugate (e.g. detectable nucleotide conjugate)can have one or more labels. Examples of the labels include but are not limited to fluorophores, spin labels, metals or metal ions, colorimetric labels, nanoparticles, PET labels, radioactive labels, or other such label as may render said nucleotide conjugate (e.g. detectable nucleotide conjugate)detectable by such methods as are known in the art of the detection of macromolecules or molecular interactions. The label may be attached to the nucleotide (e.g. by attachment to the 5' phosphate moiety of a nucleotide), to the common core itself (e.g., to the PEG subunits), to an end of the polymer, to a central moiety, or to any other location within said polymer-nucleotide conjugate which would be recognized by one of skill in the art to be sufficient to render said detectable nucleotide conjugate, such as a common core, detectable by such methods as are known in the art or described elsewhere herein. In some embodiments, one or more labels are provided so as to correspond to or differentiate a particular detectable nucleotide conjugate.

[00235] One example of the nucleotide conjugate (e.g. detectable nucleotide conjugate)is a polymer-nucleotide conjugate. Some non-limiting examples of the nucleotide conjugates (e.g. detectable nucleotide conjugates) are shown in FIG. 5A-5C. For example, FIG. 5A shows nucleotide conjugates (e.g. detectable nucleotide conjugates) having various configurations; FIG. 5B shows a nucleotide conjugate (e.g. detectable nucleotide

conjugate)having a polymer branch radiating from the center; and FIG. 5C shows nucleotide conjugates (e.g. detectable nucleotide conjugates) having a binding moiety such as a biotin.

[00236] Examples of the branched polymer include polyethylene glycol (PEG), polypropylene glycol, polyvinyl alcohol, polylactic acid, polyglycolic acid, polyglycine, polyvinyl acetate, a dextran, or other such polymers, or copolymers incorporating any two or more of the foregoing or incorporating other polymers as are known in the art. In some embodiment, the polymer is a PEG. In another embodiment, the polymer may have PEG branches.

[00237] Suitable polymers may be characterized by a repeating unit incorporating a functional group suitable for derivatization such as an amine, a hydroxyl, a carbonyl, or an allyl group. The polymer can also have one or more pre-derivatized substituents such that one or more particular subunits will incorporate a site of derivatization or a branch site, whether or not other subunits incorporate the same site, substituent, or moiety. A pre-derivatized substituent may comprise or may further comprise, for example, a nucleotide, a nucleoside, a nucleotide analog, a label such as a fluorescent label, radioactive label, or spin label, an interaction moiety, an additional polymer moiety, or the like, or any combination of the foregoing.

[00238] In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise a polymer. The polymer can have a plurality of branches. The branched polymer can have various configurations, including but are not limited to stellate ("starburst") forms, aggregated stellate ("helter skelter") forms, bottle brush, or dendrimer. The branched polymer can radiate from a central attachment point or central moiety, or may incorporate multiple branch points, such as, for example, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more branch points. In some embodiments, each subunit of a polymer may optionally constitute a separate branch point.

[00239] The length and size of the branch can differ based on the type of polymer. In some branched polymers, the branch may have a length of between 1 and 1,000 nm, between 1 and 100 nm, between 1 and 200 nm, between 1 and 300 nm, between 1 and 400 nm, between 1 and 500 nm, between 1 and 600 nm, between 1 and 700 nm, between 1 and 800 nm, or between 1 and 900 nm, or more, or having a length falling within or between any of the values disclosed herein. In some branched polymers, the branch may have a size corresponding to an apparent molecular weight of 1K, 2K, 3K, 4K, 5K, 10K, 15K, 20K, 30K, 50K, 80K, 100K, or any value within a range defined by any two of the foregoing. The apparent molecular weight of a polymer may be calculated from the known molecular weight of a representative number of subunits, as determined by size exclusion chromatography, as determined by mass spectrometry, or as determined by any other method as is known in the art. The polymer can have multiple branches. The number of branches in the polymer can be 2, 3, 4, 5, 6, 7, 8, 12, 16, 24, 32, 64, 128 or more, or a number falling within a range defined by any two of these values.

[00240] For the detectable nucleotide conjugate, the branched polymer of 4, 8, 16, 32, or 64 branches can have nucleotides attached to the ends of PEG branches, such that each end has attached thereto 0, 1, 2, 3, 4, 5, 6 or more nucleotides. In one non-limiting example, the branched polymer of between 3 and 128 PEG arms having attached to the polymer branches ends one or more nucleotides, such that each end has attached thereto 0, 1, 2, 3, 4, 5, 6 or more nucleotides or nucleotide analogs. In some embodiments, a branched polymer or dendrimer has an even number of arms. In some embodiments, a branched polymer or dendrimer has an odd number of arms.

[00241] In some embodiments, the nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprisea polymer where each branch or a subset of branches of the polymer may have attached thereto a moiety comprising a nucleotide (e.g., an adenine, a thymine, an uracil, a cytosine, or a guanine residue or a derivative or mimetic thereof), and the moiety is capable of binding to a polymerase, reverse transcriptase, or other nucleotide binding domain. Optionally, the moiety may be capable of being incorporated into an elongating nucleic acid chain during a polymerase reaction. In some instances, said moiety may be blocked such that it is not capable of being incorporated into an elongating nucleic acid chain during a polymerase reaction. In some other instances, said moiety may be reversibly blocked such that it is not capable of being incorporated into an elongating nucleic acid chain during a polymerase reaction until such block is removed, after which said moiety is then capable of being incorporated into an elongating nucleic acid chain during a polymerase reaction.

[00242] The nucleotide can be conjugated to the polymer branch through the 5' end of the nucleotide. In some instances, the nucleotide may be modified so as to inhibit or prevent incorporation of the nucleotideinto an elongating nucleic acid chain during a polymerase reaction. By way of example, the nucleotide may include a 3' deoxyribonucleotide, a 3' azidonucleotide, a 3'-methyl azido nucleotide, or another such nucleotide as is or may be known in the art, so as to not be capable of being incorporated into an elongating nucleic acid chain during a polymerase reaction. In some embodiments, the nucleotide can include a 3'-O-azido group, a 3'-O-azidomethyl group, a 3'-phosphorothioate group, a 3'-O-malonyl group, a 3'-O-alkyl hydroxylamino group, or a 3'-O-benzyl group. In some embodiments, the nucleotide lacks a 3' hydroxyl group.

[00243] The polymer can further have a binding moiety in each branch or a subset of branches. Some examples of the binding moiety include but are not limited to biotin, avidin, streptavidin or the like, polyhistidine domains, complementary paired nucleic acid domains, G-quartet forming nucleic acid domains, calmodulin, maltose-binding protein, cellulase, maltose, sucrose, glutathione-S-transferase, glutathione, O-6-methylguanine-DNA methyltransferase,

benzylguanine and derivatives thereof, benzylcysteine and derivatives thereof, an antibody, an epitope, a protein A, a protein G. The binding moiety can be any interactive molecules or fragment thereof known in the art to bind to or facilitate interactions between proteins, between proteins and ligands, between proteins and nucleic acids, between nucleic acids, or between small molecule interaction domains or moieties.

[00244] In some embodiments, a nucleotide conjugate (e.g. detectable nucleotide conjugate)as provided herein may comprise one or more elements of a complementary interaction moiety. Example complementary interaction moieties include, for example, biotin and avidin; SNAP-benzylguanosine; antibody or FAB and epitope; IgG FC and Protein A, Protein G, Protein A/G, or Protein L; maltose binding protein and maltose; lectin and cognate polysaccharide; ion chelation moieties, complementary nucleic acids, nucleic acids capable of forming triplex or triple helical interactions; nucleic acids capable of forming G-quartets, and the like. One of skill in the art will readily recognize that many pairs of moieties exist and are commonly used for their property of interacting strongly and specifically with one another; and thus any such complementary pair or set is considered to be suitable for this purpose in constructing or envisioning the nucleotide conjugates (e.g. detectable nucleotide conjugates) of the present disclosure. In some embodiments, a nucleotide conjugate (e.g. detectable nucleotide conjugate)as disclosed herein may comprise nucleotide conjugates (e.g. detectable nucleotide conjugates) in which one element of a complementary interaction moiety is attached to one molecule or multivalent ligand, and the other element of the complementary interaction moiety is attached to a separate molecule or multivalent ligand. In some embodiments, a nucleotide conjugate (e.g. detectable nucleotide conjugate)as disclosed herein may comprise nucleotide conjugates (e.g. detectable nucleotide conjugates) in which both or all elements of a complementary interaction moiety are attached to a single molecule or multivalent ligand. In some embodiments, a nucleotide conjugate (e.g. detectable nucleotide conjugate)as disclosed herein may comprise nucleotide conjugates (e.g. detectable nucleotide conjugates) in which both or all elements of a complementary interaction moiety are attached to separate arms of, or locations on, a single molecule or multivalent ligand. In some embodiments, a nucleotide conjugate (e.g. detectable nucleotide conjugate)as disclosed herein may comprise nucleotide conjugates (e.g. detectable nucleotide conjugates) in which both or all elements of a complementary interaction moiety are attached to the same arm of, or locations on, a single molecule or multivalent ligand. In some embodiments, nucleotide conjugates (e.g. detectable nucleotide conjugates) comprising one element of a complementary interaction moiety and nucleotide conjugates (e.g. detectable nucleotide conjugates) comprising another element of a complementary interaction moiety may be simultaneously or sequentially mixed. In some embodiments, interactions between molecules

or particles as disclosed herein allow for the association or aggregation of multiple molecules or particles such that, for example, detectable signals are increased. In some embodiments, fluorescent, colorimetric, or radioactive signals are enhanced. In other embodiments, other interaction moieties as disclosed herein or as are known in the art are contemplated. In some embodiments, a nucleotide conjugate (e.g. detectable nucleotide conjugate)as provided herein may be provided such that one or more molecules comprising a first interaction moiety such as, for example, one or more imidazole or pyridine moieties, and one or more additional molecules comprising a second interaction moiety such as, for example, histidine residues, are simultaneously or sequentially mixed. In some embodiments, said nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises 1, 2, 3, 4, 5, 6, or more imidazole or pyridine moieties. In some embodiments, said nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises 1, 2, 3, 4, 5, 6, or more histidine residues. In such embodiments, interaction between the molecules or particles as provided may be facilitated by the presence of a divalent cation such as nickel, manganese, magnesium, calcium, strontium, or the like. In some embodiments, for example, a $(His)_3$ group may interact with a $(His)_3$ group on another molecule or particle via coordination of a nickel or manganese ion.

ix. Binding complexes

[00245] The methods described herein may involve detecting a target by forming a binding complex between a primed index sequence described herein and a nucleotide conjugate (e.g. detectable nucleotide conjugate)as described herein. The target may be a ribonucleic acid. The binding complex may include one or more non-covalent interactions.

[00246] 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., FIGs. 5A-5C). 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 moiety, 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 moiety (e.g., FIGs. 16-18). In some embodiments, the nucleotide moiety comprises a base, sugar and at least one phosphate group, and the linker is attached to the nucleotide moiety 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 example nucleotide arm is shown in FIG. 19. Example multivalent molecules are shown in FIGs. 5A-C and 16-18. An example

spacer is shown in Figure 6 (top) and example linkers are shown in FIG. 20 (bottom) and FIG. 21. Example nucleotides attached to a linker are shown in FIGs. 22-24. An example biotinylated nucleotide arm is shown in FIG. 25.

[00247] The formation of a binding complex in the context of the methods and systems described herein may have certain advantages. An increase in binding of a nucleotide to an enzyme (e.g., polymerase) or an enzyme complex can be affected by increasing the effective concentration of the nucleotide. The increase can be achieved by increasing the concentration of the nucleotide in free solution, or by increasing the amount of the nucleotide in proximity to the relevant binding site. The increase can also be achieved by physically restricting a number of nucleotides into a limited volume thus resulting in a local increase in concentration, and such as structure may thus bind to the binding site with a higher apparent avidity than would be observed with unconjugated, untethered, or otherwise unrestricted individual nucleotide. One example means of effecting such restriction is by providing a binding complex in which multiple nucleotides are bound to a particle such as a polymer, a branched polymer, a dendrimer, a micelle, a liposome, a microparticle, a nanoparticle, a quantum dot, or other suitable particle known in the art.

[00248] The binding complex can be used to localize detectable signals to active regions of biochemical interactions, such as sites of protein-nucleic acid interactions, nucleic acid hybridization reactions, or enzymatic reactions, such as polymerase reactions. For instance, the binding complex described herein can be utilized to identify sites of base incorporation in elongating nucleic acid chains during polymerase reactions and to provide base discrimination for sequencing and array based applications. The increased binding between the target and the nucleotide in the binding complex, when the nucleotide is complementary to the target nucleic acid, provides enhanced signal that greatly improve base call accuracy and shorten imaging time.

[00249] In addition, the formation of a binding complex as described herein allows sequencing signals from a given sequence to originate within cluster regions containing multiple copies of the target sequence. Sequencing methods incorporating multiple copies of a target sequence have the advantage that signals can be amplified due to the presence of multiple simultaneous sequencing reactions within the defined region, each providing its own signal. The presence of multiple signals within a defined area also reduces the impact of any single skipped cycle, due to the fact that the signal from a large number of correct base calls can overwhelm the signal from a smaller number of skipped or incorrect base calls, therefore providing methods for reducing phasing errors and/or to improve read length in sequencing reactions.

[00250] The formation of a binding complex and the compositions associated with it, as described herein may lead to one or more of: (i) stronger signal for better base-calling

accuracy compared to conventional nucleic acid amplification and sequencing methodologies; ii) allow greater discrimination of sequence-specific signal from background signals; (iii) reduced requirements for the amount of starting material necessary, (iv) increased sequencing rate and shortened sequencing time; (v) reducing phasing errors, and (vi) improving read length in sequencing reactions.

[00251] In some embodiments, the nucleotide moiety 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 phosphoramide linkage. In some embodiments, at least one nucleotide moiety 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 BH_3. In some embodiments, the chain includes phosphate groups substituted with analogs including phosphoramidate, phosphorothioate, phosphordithioate, and O-methylphosphoroamidite groups.

[00252] In some embodiments, the multivalent molecule comprises a core attached to multiple nucleotide arms, and wherein individual nucleotide arms comprise a nucleotide moiety 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 moiety 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 moiety 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, or silyl group. In some embodiments, the chain terminating moiety is cleavable/removable from the nucleotide moiety, 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) ($Pd(PPh_3)_4$) 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 (K_2CO_3) 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.

[00253] In some embodiments, the nucleotide moiety 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-TPP) or Tri(hydroxyproyl)phosphine (THPP). In some embodiments, the cleaving agent comprises 4-dimethylaminopyridine (4-DMAP).

[00254] In some embodiments, the nucleotide moiety comprising 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'-sulfhydral, 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.

[00255] In some embodiments, the multivalent molecule comprises a core attached to multiple nucleotide arms, wherein the nucleotide arms comprise a spacer, linker and nucleotide moiety, and wherein the core, linker and/or nucleotide moiety 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 moiety to permit detection and identification of the nucleotide base.

[00256] In some embodiments, at least one nucleotide arm of a multivalent molecule has a nucleotide moiety 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 moiety to permit detection and identification of the nucleotide base.

[00257] 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 a 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 streptavidin. For example, avidin moiety includes 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.

[00258] In some embodiments, the binding complex may be formed between one or more nucleotides of a detectable nucleotide conjugate, as described herein, and a nucleotide of a primed index sequence, as described herein. In some embodiments, the binding complex may be formed between one or more nucleotides of the detectable nucleotide conjugate, a nucleotide of a primed index sequence and a polymerizing enzyme (e.g., DNA polymerase). In some embodiments, the binding complex comprises one type of detectable nucleotide conjugate. In some embodiments, the binding complex comprises two or more types of nucleotide conjugates (e.g. detectable nucleotide conjugates). In some embodiments, each type of the two or more types of nucleotide conjugates (e.g. detectable nucleotide conjugates) comprises a different type of nucleotide. In some embodiments, the binding complex comprises a blocked nucleotide (e.g., 3'-O-azidomethyl, a 3'-O-alkyl hydroxylamino or 3'-O-methyl nucleotide). In some embodiments, the blocked nucleotide is at the 3' position of the primer nucleic acid sequence of the primed index sequence. In some embodiments, the nucleotide of the nucleotide conjugate (e.g. detectable nucleotide conjugate)is the blocked nucleotide.

[00259] In some embodiments, the binding complex may be multivalent. In some embodiments, the binding complex may be formed between two or more nucleotides of a detectable nucleotide conjugate, as described herein, and a nucleotide of a primed index sequence, as described herein. In some embodiments, the binding complex may be formed between two or more nucleotides of the detectable nucleotide conjugate, a nucleotide of a primed index sequence and a polymerizing enzyme (e.g., DNA polymerase). In some embodiments, the binding complex

comprises one type of detectable nucleotide conjugate. In some embodiments, the binding complex comprises two or more types of nucleotide conjugates (e.g. detectable nucleotide conjugates). In some embodiments, each type of the two or more types of nucleotide conjugates (e.g. detectable nucleotide conjugates) comprises a different type of nucleotide. In some embodiments, the binding complex comprises a blocked nucleotide (e.g., 3'-O-azidomethyl, a 3'-O-alkyl hydroxylamino or 3'-O-methyl nucleotide). In some embodiments, the blocked nucleotide is at the 3' position of the primer nucleic acid sequence of the primed index sequence. In some embodiments, the nucleotide of the nucleotide conjugate (e.g. detectable nucleotide conjugate)is the blocked nucleotide.

[00260] In some embodiments, the nucleotide moieties involved in the formation of the binding complex are not incorporated. In some embodiments, the binding complex further comprises a polymerase. In some embodiments, the present disclosure provides said method wherein the polymerase molecule is catalytically inactive, such as where the polymerase molecule been rendered catalytically inactive by mutation, by chemical modification, or by the absence of a necessary ion or cofactor. In some embodiments, the present disclosure also provides said method wherein the polymerase molecule is catalytically active, and/or wherein the binding complex does not comprise a blocked nucleotide.

[00261] 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 moiety 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 moiety 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 moiety is complementary to a next base of the template nucleic acid, and destabilizes the binding complex when the nucleotide or nucleotide moiety is not complementary to the next base of the template nucleic acid.

[00262] In some embodiments, in any of the sequencing methods that employ multivalent 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 moiety 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 moiety 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. Example multivalent molecules are shown in FIGs. 21-24.

[00263] In some embodiments, in any of the sequencing methods that employ multivalent molecules, 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 moiety 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 moiety 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 moieties 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 moiety in the first binding complex thereby determining the sequence of the first portion of the concatemer template molecule, and identifying the second nucleotide moiety 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. Example multivalent molecules are shown in FIGs. 21-24.

[00264] In some embodiments, 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 composition is deposited under buffer conditions incorporating a polar aprotic solvent. In some embodiments, the contacting is performed under a condition that stabilizes said binding complex when said nucleotide is complementary to a next base of said target and destabilizes said binding complex when said nucleotide is not complementary to said next base of said target nucleic acid.

[00265] In some embodiments, the binding complex may be used in replacement of single unconjugated or untethered nucleotide to form a binding complex with the polymerase and the target and the local concentration of the nucleotide may be increased many fold, which in turn may enhance the signal intensity. The formed binding complex may also have a longer persistence time which in turn helps shorten the imaging step. The high signal intensity resulting from the use of the nucleotide conjugate (e.g. detectable nucleotide conjugate)may remain for the entire binding and imaging step. The strong binding between the polymerase, the primed index sequence, and the nucleotide or nucleotide analog also means that the formed binding complex will remain stable during the washing step and the signal will remain at a high intensity when other reaction mixture and unmatched nucleotide analogs are washed away. After the imaging step, the binding complex may be destabilized, and the primed index sequence may be extended for one base. In some

embodiments, the one base nucleotide used to extend the primed index sequence by one base may be unmodified. In some embodiments, the one base nucleotide used to extend the primed index sequence by one base may not comprise a label.

[00266] In some embodiments, the binding complex described herein may comprise one or more buffers, salts, ions, or additives. In some embodiments, representative additives may include, but are not limited to, betaine, spermidine, detergents such as polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether (Triton X-100), 2-[2-[3,4-bis(2-hydroxyethoxy)oxolan-2-yl]-2-(2-hydroxyethoxy)ethoxy]ethyl dodecanoate (Tween 20), sodium lauryl sulfate (SDS), or nonyl phenoxypolyethoxylethanol (NP-40), ethylene glycol, polyethylene glycol, dextran, polyvinyl alcohol, vinyl alcohol, methylcellulose, heparin, heparan sulfate, glycerol, sucrose, 1,2-propanediol, dimethyl sulfoxide (DMSO), N,N,N-trimethylglycine, ethanol, ethoxyethanol, propylene glycol, polypropylene glycol, block copolymers such as the Pluronic (r) series polymers, arginine, histidine, imidazole, or any combination thereof, or any substance known in the art as a DNA "relaxer" (a compound, with the effect of altering the persistence length of DNA, altering the number of within-polymer junctions or crossings, or altering the conformational dynamics of a DNA molecule such that the accessibility of sites within the strand to DNA binding moieties is increased).

[00267] The binding complex may include zwitterionic compounds as additives. Further representative additives may be found in Lorenz, T.C. J. Vis. Exp. (63), e3998, doi:10.3791/3998 (2012), which is hereby incorporated by reference with respect to its disclosure of additives for the facilitation of nucleic acid binding or dynamics, or the facilitation of processes involving the manipulation, use, or storage of nucleic acids. In some embodiments, representative cations may include, but are not limited to, sodium, magnesium, strontium, potassium, manganese, calcium, lithium, nickel, cobalt, or other such cations as are known in the art to facilitate nucleic acid interactions, such as self-association, secondary or tertiary structure formation, base pairing, surface association, peptide association, protein binding, or the like.

[00268] The present disclosure relates to binding complexes having a plurality of nucleotides conjugated to a particle (e.g., a polymer, branched polymer, dendrimer, or equivalent structure). Contacting the binding complex with a polymerase and a primed target may result in the formation of a binding complex which may be detected and in turn achieve a more accurate determination of the bases of the target nucleic acid.

[00269] In some embodiments, the binding complex may be used in replacement of single unconjugated or untethered nucleotide to form a complex with the polymerase and the target nucleic acid. In some cases, the local concentration of the nucleotide may be increased many fold by the formation of a binding complex, which in turn may enhance the signal intensity, particularly

the correct signal versus mismatch. The binding complex described herein can include at least one nucleotide conjugate (e.g. detectable nucleotide conjugate)for interacting with the target nucleic acid. The multivalent composition can also include two, three, or four different nucleotide conjugates (e.g. detectable nucleotide conjugates), each having a different nucleotide conjugated to the common core.

[00270] The binding complex may comprise 1, 2, 3, 4, or more types of nucleotide conjugates (e.g. detectable nucleotide conjugates), wherein each nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises a different type of nucleotide. A first type of the nucleotide conjugate (e.g. detectable nucleotide conjugate)can comprise a nucleotide selected from the group consisting of ATP, ADP, AMP, dATP, dADP, and dAMP. A second type of the nucleotide conjugate (e.g. detectable nucleotide conjugate)can comprise a nucleotide selected from the group consisting of TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, and dUMP. A third type of the nucleotide conjugate (e.g. detectable nucleotide conjugate)can comprise a nucleotide selected from the group consisting of CTP, CDP, CMP, dCTP, dCDP, and dCMP. A fourth type of the nucleotide conjugate (e.g. detectable nucleotide conjugate)can comprise a nucleotide selected from the group consisting of GTP, GDP, GMP, dGTP, dGDP, and dGMP. In some embodiments, each nucleotide conjugate (e.g. detectable nucleotide conjugate)comprises a single type of nucleotide respectively corresponding to one or more nucleotide selected from the group consisting of ATP, ADP, AMP, dATP, dADP, dAMP TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, dUMP, CTP, CDP, CMP, dCTP, dCDP, dCMP, GTP, GDP, GMP, dGTP, dGDP, and dGMP. Each binding complex may further comprise one or more labels corresponding to the particular nucleotide conjugated to each respective conjugate. Example labels include fluorescent labels, colorimetric labels, electrochemical labels (such as, for example, glucose or other reducing sugars, or thiols or other redox active moieties), luminescent labels, spin labels, radioactive labels, steric labels, affinity tags, or the like.

[00271] 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 moiety, which is selected from a group consisting of dATP, dGTP, dCTP, dTTP and dUTP.

[00272] In some embodiments, a multivalent molecule comprises a core attached to multiple nucleotide arms, where each arm includes a nucleotide moiety. The nucleotide moiety 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 moiety selected from a

group consisting of dATP, dGTP, dCTP, dTTP and dUTP. The plurality of multivalent molecules can comprise at a mixture of any combination of two or more types of multivalent molecules, where individual multivalent molecules in the mixture comprise nucleotide moieties selected from a group consisting of dATP, dGTP, dCTP, dTTP and/or dUTP.

[00273] **FIG. 6** demonstrates how the use of the nucleotide conjugate (e.g. detectable nucleotide conjugate)increases the signal intensity during binding, persistence, and washing/removal steps of the methods described herein. Because of the increased local concentration of the nucleotide on the polymer-nucleotide conjugate, the binding between the polymerase, the primed target strand, and the nucleotide, when the nucleotide is complementary to the next base of the target nucleic acid, becomes more favorable. The formed binding complex has a longer persistence time which in turn helps shorten the imaging step. The high signal intensity results from the use of the nucleotide conjugate (e.g. detectable nucleotide conjugate)may remain for the entire binding and imaging step. The strong binding between the polymerase, the primed target strand, and the nucleotide or nucleotide analog may also result in the formed binding complex remaining stabilized during a washing step and the signal may remain at a high intensity when other reaction mixture and unmatched nucleotide analogs are washed away. In some embodiments, after the imaging step, the binding complex may be destabilized, and the primed index sequence may be extended for one base. After the extension, the binding and imaging steps can be repeated again with the use of the polymer nucleotide conjugate to determine the identity of the next base.

[00274] As an example, a graphical depiction of the increase in signal intensity during binding, persistence, and washing/removal of a multivalent substrate as described herein is provided in **FIG. 10**, which is representative of the changes in signal intensity that have been observed experimentally. Therefore, the compositions and methods of the present disclosure provide a robust and controllable means of establishing and maintaining a ternary enzyme complex, as well as providing vastly improved means by which the presence of said complex may be identified and/or measured, and a means by which the persistence of said complex may be controlled. This provides important solutions to problems such as that of determining the identity of the N+1 base in nucleic acid sequencing applications.

[00275] In some embodiments, the binding complexes disclosed herein associate with polymerase nucleotide complexes in order to form binding complexes with a rate that is time-dependent, though substantially slower than the rate of association known to be obtainable by nucleotides in free solution. In some cases, the on-rate (Kon) is substantially and surprisingly slower than the on rate for single nucleotides or nucleotides not attached to multivalent ligand complexes. In some cases, the off rate (Koff) of the binding complex is substantially slower than

that observed for nucleotides in free solution. In some embodiments, the binding complexes of the present disclosure provide a surprising and beneficial improvement of the persistence of ternary polymerase-polynucleotide-nucleotide complexes (especially over such complexes that are formed with free nucleotides) allowing, for example, significant improvements in imaging quality for nucleic acid sequencing applications, over currently available methods and reagents. Importantly, this property of the multivalent substrates disclosed herein renders the formation of visible ternary complexes controllable, such that subsequent visualization, modification, or processing steps may be undertaken essentially without regard to the dissociation of the complex—that is, the complex can be formed, imaged, modified, or used in other ways as necessary, and will remain stable until a user carries out an affirmative dissociation step, such as exposing the complexes to a dissociation buffer.

[00276] In some embodiments, polymerizing enzymes suitable for the binding interaction describe herein include may include any polymerase as is or may be known in the art. In some cases, the polymerases described herein may be polymerizing enzymes. In some cases, the polymerase may include Klenow DNA polymerase, Thermus aquaticus DNA polymerase I (Taq polymerase), KlenTaq® polymerase, and 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, and 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, or Moloney Murine Leukemia Virus (MMLV) reverse transcriptase, or telomerase. Further non-limiting examples of DNA polymerases can include those from various Archaea genera, such as, Aeropyrum, Archaeglobus, Desulfurococcus, Pyrobaculum, Pyrococcus, Pyrolobus, Pyrodictium, Staphylothermus, Stetteria, Sulfolobus, Thermococcus, and Vulcanisaeta and the like or variants thereof, including such polymerases as are known in the art such as Vent™, Deep Vent™, Pfu, KOD, Pfx, Therminator™, Tgo polymerases, or a combination thereof. In some embodiments, the polymerase is a klenow polymerase.

[00277] The binding complex as described herein may have a longer persistence time when the nucleotide on the nucleotide conjugate (e.g. detectable nucleotide conjugate)is complementary to the target than with a non-complementary nucleotide. In some cases, the binding complex described herein may have a longer persistence time when the nucleotide on the nucleotide conjugate (e.g. detectable nucleotide conjugate)is complementary to the target than a complementary nucleotide that is not conjugated or tethered. In some embodiments, the binding complex may have a persistence time of less than 1s, greater than 1s, greater than 2s, greater than

3s, greater than 5s, greater than 10s, greater than 15s, greater than 20s, greater than 30s, greater than 60s, greater than 120s, greater than 360s, greater than 3600s, or more, or for a time lying within a range defined by any two or more of these values.

[00278] The persistence time can be measured, for example, 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.

[00279] In some cases, persistence time of the binding complex described herein may be altered by the salt environment of the binding complex. In some embodiments, binding complexes formed in the presence of, for example, Mg^{2+} ions may form more quickly than complexes formed with other ions. In some cases, binding complexes formed in the presence of Sr^{2+} ions may form readily and dissociate completely or with substantial completeness upon withdrawal of the ion or upon washing with buffer lacking one or more components of the present compositions, such as, e.g., a polymer and/or one or more nucleotides, and/or one or more interaction moieties, or a buffer containing, for example, a chelating agent which may cause or accelerate the removal of a divalent cation from the multivalent reagent containing complex. In some embodiments, a composition of the present disclosure may comprise Mg^{2+} ions. In some embodiments, a composition of the present disclosure may comprise Ca^{2+} ions. In some embodiments, a composition of the present disclosure comprises Sr^{2+}. In some embodiments, a composition of the present disclosure comprises Cobalt ions. In some embodiments, a composition of the present disclosure may comprise $MgCl_2$. In some embodiments, a composition of the present disclosure may comprise $CaCl_2$. In some embodiments, a composition of the present disclosure may comprise $SrCl_2$. In some embodiments, a composition of the present disclosure may comprise $CoCl_2$. In some embodiments, the composition may comprise no, or substantially no Magnesium. In some embodiments, the composition may comprise no, or substantially no Mg ions. In some embodiments, the composition comprises no, or substantially no Calcium. In some embodiments, the composition comprises no, or substantially no Ca^{2+} ions. In some embodiments, the methods of the present disclosure provide for the contacting of one or more nucleic acids with one or more of the compositions disclosed herein wherein said composition lacks either one of calcium or magnesium, or lacks both calcium and magnesium. In some embodiments, the methods of the present disclosure provide for the contacting of one or more nucleic acids with one or more of the compositions disclosed herein wherein said composition lacks either one of Ca ions or Mg ions, or lacks both Ca ions and Mg ions. In some embodiments, such conditions render the polymerizing enzyme capable of forming the binding complex, but incapable of incorporating the

nucleotide moiety (e.g., nucleotide, nucleotide analog) of the nucleotide conjugate (e.g. detectable nucleotide conjugate)into the primer nucleic acid of the primed template strand.

[00280] In some embodiments, the methods described herein may comprise dissociation of the binding complex. The dissociation of binding complexes may be affected by changing the buffer conditions. In some cases, after the imaging step in the methods described herein, a buffer with increased salt content may be used to cause dissociation of the binding complexes such that nucleotide conjugates (e.g. detectable nucleotide conjugates) can be washed out, providing a means by which signals can be attenuated or terminated, such as in the transition between one sequencing cycle and the next. This dissociation may be affected, in some embodiments, by washing the complexes with a buffer lacking a necessary metal or cofactor. In some embodiments, a wash buffer may comprise one or more compositions for the purpose of maintaining pH control. In some embodiments, a wash buffer may comprise one or more monovalent cations, such as sodium. In some embodiments, a wash buffer lacks or substantially lacks a divalent cation, for example, having no or substantially no strontium, calcium, magnesium, or manganese. In some embodiments, a wash buffer further comprises a chelating agent, such as, for example, EDTA, EGTA, nitrilotriacetic acid, polyhistidine, imidazole, or the like. In some embodiments, a wash buffer may maintain the pH of the environment at the same level as for the bound complex. In some embodiments, a wash buffer may raise or lower the pH of the environment relative to the level seen for the bound complex. In some embodiments, the pH may be within a range from 2-4, 2-7, 5-8, 7-9, 7-10, or lower than 2, or higher than 10, or a range defined by any two of the values provided herein.

[00281] The addition of a particular ion may affect the binding of the polymerase to a primed-index sequence, the formation of a binding complex, the dissociation of a binding complex, or the incorporation of one or more nucleotides into an elongating nucleic acid such as during a polymerase reaction. In some embodiments, relevant anions may comprise chloride, acetate, gluconate, sulfate, phosphate, or the like. In some embodiments, an ion may be incorporated into the compositions of the present disclosure by the addition of one or more acids, bases, or salts, such as $NiCl_2$, $CoCl_2$, $MgCl_2$, $MnCl_2$, $SrCl_2$, $CaCl_2$, $CaSO_4$, $SrCO_3$, $BaCl_2$ or the like. Representative salts, ions, solutions and conditions may be found in Remington: The Science and Practice of Pharmacy, 20th. Edition, Gennaro, A.R., Ed. (2000), which is hereby incorporated by reference in its entirety, and especially with respect to Chapter 17 and related disclosure of salts, ions, salt solutions, and ionic solutions.

[00282] The present disclosure contemplates contacting the binding complex comprising at least one nucleotide conjugate (e.g. detectable nucleotide conjugate)with one or more polymerizing enzymes. In some embodiments, the one or more polymerizing enzymes may

comprise a polymerase. The contacting may be optionally done in the presence of one or more targets. In some embodiments, the targets may be single stranded nucleic acids. In some embodiments, said targets are primed single stranded nucleic acids. In some embodiments, said targets may be double stranded nucleic acids. In some embodiments, said contacting comprises contacting the binding complex with one polymerase. In some embodiments, the contacting comprises contacting of the binding complex multiple polymerases. The polymerase may be bound to a single nucleic acid molecule.

[00283] The binding between a target and a nucleotide conjugate (e.g. detectable nucleotide conjugate)may be provided in the presence of a polymerase that has been rendered catalytically inactive. In one embodiment, the polymerase may have been rendered catalytically inactive by mutation. In one embodiment, the polymerase may have been rendered catalytically inactive by chemical modification. In some embodiments, the polymerase may have been rendered catalytically inactive by the absence of a necessary substrate, ion, or cofactor. In some embodiments, the polymerase enzyme may have been rendered catalytically inactive by the absence of magnesium ions.

[00284] The binding between target and the nucleotide conjugate (e.g. detectable nucleotide conjugate)may occur in the presence of a polymerase wherein the binding solution, reaction solution, or buffer lacks magnesium or manganese. In some embodiments, the binding between the target and nucleotide conjugate (e.g. detectable nucleotide conjugate)may occur in the presence of a polymerase wherein the binding solution, reaction solution, or buffer comprises calcium or strontium.

[00285] In some embodiments, a catalytically inactive polymerase may be used to help a nucleic acid interact with a detectable nucleotide conjugate. In some embodiments, the interaction between the binding complex and a polymerase stabilizes the binding complex so as to render the complex detectable by fluorescence or by other methods as disclosed herein or otherwise known in the art. Unbound nucleotide conjugates (e.g. detectable nucleotide conjugates) may optionally be washed away prior to detection of the binding complex.

[00286] In some embodiments, contacting of one or more targets with the nucleotide conjugates (e.g. detectable nucleotide conjugates) disclosed herein may comprise a solution containing either one of calcium or magnesium, or containing both calcium and magnesium. In some embodiments, contacting one or more targets with the nucleotide conjugates (e.g. detectable nucleotide conjugates) disclosed herein may comprise a solution lacking either one of calcium or magnesium, or lacking both calcium or magnesium, and in a separate step, without regard to the order of the steps, adding to the solution one of calcium or magnesium, or both calcium and magnesium. In some embodiments, the contacting of one or more nucleic acids with the nucleotide

conjugates (e.g. detectable nucleotide conjugates) disclosed herein may comprise a solution lacking strontium, and comprises in a separate step, without regard to the order of the steps, adding to the solution strontium.

x. Substrate

[00287] Disclosed herein, in some embodiments, are substrates configured to immobilize a biological sample (e.g., tissue, cell, *etc.*). In some embodiments, the biological sample immobilized to the substrate comprises a target for *in situ* detection utilizing the methods and systems of the present disclosure. The target may be a ribonucleic acid. The substrate may be a solid support, such as a flow cell. In some embodiments, the biological sample is immobilized to a surface of the substrate. In the example of the flow cell, the surface may be an interior surface of the flow cell. In some cases, the surface of the substrate may improve nucleic acid hybridization performance, amplification performance, or a combination thereof.

[00288] Disclosed herein are solid supports comprising low non-specific binding surface compositions that enable improved nucleic acid hybridization and amplification performance. In general, the disclosed supports may 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 primer sequences that may be used for tethering single-stranded target nucleic acid(s) to the support surface. In some instances, the formulation of the surface, 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 surface 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 support surface is minimized or reduced relative to a comparable monolayer. The formulation of the surface may be varied such that non-specific hybridization on the support surface is minimized or reduced relative to a comparable monolayer. The formulation of the surface may be varied such that non-specific amplification on the support surface is minimized or reduced relative to a comparable monolayer. The formulation of the surface may be varied such that specific amplification rates and/or yields on the support surface 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.

[00289] Examples of materials from which the substrate or 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 substrates are contemplated.

[00290] The substrate or 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, in some instances the substrate or support structure may be locally planar (e.g., comprising a microscope slide or the surface of a microscope slide). Globally, the substrate or 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 instances, the surface of the substrate or support structure used for nucleic acid hybridization and amplification may be a solid, non-porous surface. In some instances, the surface of the substrate or 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.

[00291] The substrate or 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 instances, the substrate or support structure may comprise one or more surfaces within an integrated or assembled microfluidic flow cell. The substrate or support structure may comprise one or more surfaces within a microplate format, e.g., the bottom surface of the wells in a microplate. As noted above, in some preferred embodiments, the substrate or support structure comprises the interior surface (such as the lumen surface) of a capillary. In alternate preferred embodiments the substrate or support structure comprises the interior surface (such as the lumen surface) of a capillary etched into a planar chip.

[00292] As noted, 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, in some instances, 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 instances, 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 instances, 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.

[00293] 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.

[00294] As noted, in some instances, 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 be detection of the amount of label remaining on the surface and comparison of the signal resulting therefrom to an appropriate calibration standard. In some instances, the label may comprise a fluorescent label. In some instances, the label may comprise a radioisotope. In some instances, the label may comprise any other detectable label known to one of skill in the art. In some instances, 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 other molecules) per unit area. In some instances, 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 1molecule 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-Rho11 (Jena Biosciences), 10 μM Aminoallyl-dUTP - ATTO-Rho11 (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 substrates at 37°C for 15 minutes in a 384 well plate format. Each well was rinsed 2-3 x with 50 μl deionized RNase/DNase Free water and 2-3 x 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 (Olympus Corp., Center Valley, PA) with a total internal reflectance fluorescence (TIRF) objective (100X, 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, New York), 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.

[00295] In some instances, 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 instances, 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.

[00296] 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.

[00297] In some instances, 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 instances, a static contact angle may be determined. In some instances, an advancing or receding contact angle may be determined. In some instances, 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 instances, 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.

[00298] In some instances, 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 instances, adequate wash steps may be performed in less than 60, 50, 40, 30, 20, 15, 10, or less than 10 seconds. For example, in some instances adequate wash steps may be performed in less than 30 seconds.

[00299] Some low-binding surfaces of the present disclosure 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, in some instances, 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 instances, 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 instances, 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).

[00300] In some instances, 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.

[00301] In some instances, fluorescence images of the disclosed low background surfaces when used in nucleic acid hybridization or amplification applications to create clusters 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, 20, 210, 220, 230, 240, 250, or greater than 250.

[00302] One or more types of primer may be attached or tethered to the support surface. In some instances, the one or more types of adapters or primers may comprise spacer sequences, adapter sequences for hybridization to adapter-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 instances, 1 primer or adapter sequence may be tethered to at least one layer of the surface. In some instances, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10 different primer or adapter sequences may be tethered to at least one layer of the surface.

[00303] In some instances, the tethered adapter and/or primer sequences may range in length from about 10 nucleotides to about 100 nucleotides. In some instances, the tethered adapter 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 instances, the tethered adapter 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 instances the length of the tethered adapter 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 adapter and/or primer sequences may have any value within this range, e.g., about 24 nucleotides.

[00304] In some instances, the resultant surface density of 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 instances, 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 instances, 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 instances, 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 instances 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 instances, the surface density of target library nucleic acid sequences initially hybridized to adapter 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 instances, the surface density of clonally-amplified target library nucleic acid sequences hybridized to adapter or primer sequences on the support surface may span the same range as that indicated for the surface density of tethered primers.

[00305] 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 / μm^2, while also comprising at least a second region having a substantially different local density.

xi. Detection and identification

[00306] The methods described herein may involve detecting a target in a biological sample in situ. The target may be an RNA, a DNA, or a polypeptide. In some embodiments, a binding complex may be formed between a nucleotide conjugate (e.g. detectable nucleotide conjugate)and a primed index sequence, as described herein. The nucleotide conjugate (e.g. detectable nucleotide conjugate) may be detected using an imaging system. The imaging system may be configured to image the biological sample.

[00307] In some cases, the methods described herein include detection of the target using sequencing-by-binding (SBB). In some embodiments, the sequencing-by-binding (SBB) procedure employs non-labeled chain-terminating nucleotides. In some embodiments, a cycle of sequencing-by-binding (SBB) comprises the steps of **(a)** sequentially contacting a primed concatemer (e.g., a concatemer annealed to a plurality of sequencing primers) 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 concatemer 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 concatemer, 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 concatemer after step (b), thereby producing an extended primer; and **(e)** repeating steps (a) through (d) at least once on the primed concatemer that comprises the extended primer. Example 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).

[00308] In some embodiments, the imaging system described herein may include one or more image sensors. In some embodiments, the one or more image sequences may comprise a photodetector array. Representative imaging systems may be found, for example, in International Application No. PCT/US2024/012802, the contents of which are hereby incorporated by reference in their entirety.

[00309] In some embodiments, the biological sample described herein may be coupled to a substrate. In some embodiments, the biological sample may be immobilized on a surface of the substrate, as described herein. In some embodiments, the surface may be an interior surface of a flow cell. In some embodiments, the surface has a water contact angle of less than or

equal to 45 degrees. In some embodiments, the flow cell may comprise one or more hydrophilic polymer layers. In some embodiments, the one or more hydrophilic polymer layers may comprise a branched polymer. In some embodiments, one or more hydrophilic polymer layers may comprise a polymer comprising polyethylene glycol.

[00310] In some embodiments, the detectable label is a fluorophore. In some embodiments, the fluorophore may be a fluorescent label. The fluorophore or fluorescent label may be any one of the configurations described herein. In some cases, detecting the complex comprises a fluorescence measurement. The fluorescence measurement may be a quantitative measurement, an image, or a combination thereof. In some embodiments, the detecting as described herein may involve imaging the biological sample immobilized within a flow cell using an imaging system as described herein.

[00311] . In some embodiments, the present disclosure provides said method further comprising a dissociation step that destabilizes said binding complex formed between the composition of (a) and the polymer-nucleotide conjugate to remove said polymer-nucleotide conjugate. In some embodiments, the present disclosure provides said method further comprising an extension step to incorporate into said primer nucleic acid a nucleotide that is complementary to said next base of the target nucleic acid, and optionally wherein the extension step occurs currently as or after said dissociation step.

[00312] In some embodiments, multiple rounds of detection are performed. The multiple rounds of detection may comprise the formation of additional binding complexes between a nucleotide conjugate (e.g. detectable nucleotide conjugate) and the primed index sequence as described herein. In some cases, each round of the multiple rounds of detection may reveal a different nucleotide sequence of the index sequence.

[00313] The methods described herein, may comprise sequencing the one or more index sequence using a sequencing method. The sequencing method may comprise formation of binding complexes as described herein. In some embodiments, the sequencing method may comprise incorporation of unlabeled nucleotides. The sequencing method may further comprise incorporating the N+1 or terminal nucleotide into the primed-index sequence as shown in FIG. 1F. In FIG. 1F, the primer strand of the primed-index sequence 104 can be extended for one base to form an extended nucleic acid 108. The extension step can occur after or concurrently with the destabilization of the binding complex. The primed target 104 can be extended using a complementary nucleotide that is attached to the particle in the nucleotide conjugate (e.g. detectable nucleotide conjugate), or using an unconjugated or untethered free nucleotide that is provided after the binding complex has been removed.

[00314] In some embodiments, after the extension step, the contacting step as shown in FIG. 1G can be performed again to form binding complexes and imitate the next sequencing cycle. The contacting, detecting, and extension steps can be repeated for one or more cycles, thereby determining the sequence of the target molecule. For example, FIG. 1H shows the surface images after multiple sequencing cycles, and the images can then be processed to determine the sequences of the target molecules.

[00315] The extension of the primed target may be prevented or inhibited due to a blocked nucleotide on the strand or the use of polymerase that is catalytically inactive. When the nucleotide in the polymer-nucleotide conjugate has a blocking group that prevents the extension of the nucleic acid, incorporation of a nucleotide may be achieved by the removal of a blocking group from said nucleotide (such as by detachment of said nucleotide from its polymer, branched polymer, dendrimer, particle, or the like). When the extension of the primed target is inhibited due to the use of polymerase that is catalytically inactive, incorporation of a nucleotide may be achieved by the provision of a cofactor or activator such as a metal ion.

[00316] In some embodiments, the polymerase(s) is/are capable of binding and incorporating a complementary nucleotide opposite a nucleotide in an amplicon. In some embodiments, the polymerase(s) is/are capable of binding a complementary nucleotide opposite a nucleotide in an amplicon without incorporation. In some embodiments, the sequencing polymerase(s) is/are capable of binding a complementary nucleotide moiety of a multivalent molecule opposite a nucleotide in an amplicon. In some embodiments, the plurality of sequencing polymerases comprise recombinant mutant polymerases.

[00317] 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 altiarchaeales archaeon; Candidatus Hadarchaeum Yellowstonense; Hadesarchaea archaeon; Euryarchaeota archaeon; Thermoplasmata 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, Archaeglobus, Desulfurococcus, Pyrobaculum, Pyrococcus, Pyrolobus, Pyrodictium, 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.

[00318] In some embodiments, the four different nucleic acid binding compositions, or three different nucleic acid binding compositions, each comprising a different nucleotide or nucleotide analog, are used to determine the identity of a terminal nucleotide on a primed index sequence, wherein one of the four or three different nucleic acid binding compositions is labeled with a first fluorophore, one is labeled with a second fluorophore, one is labeled with both the first and second fluorophore, and one is not labeled or is absent, and wherein the detecting step comprises simultaneous excitation at a first excitation wavelength and a second excitation wavelength and images are acquired at a first fluorescence emission wavelength and a second fluorescence emission wavelength. In some embodiments, the first fluorophore is Cy3, the second fluorophore is Cy5, the first excitation wavelength is 532 nm or 568 nm, the second excitation wavelength is 633 nm, the first fluorescence emission wavelength is about 570 nm, and the second fluorescence emission wavelength is about 670 nm. In some embodiments, the detection label can comprise one or more portions of a FRET pair, such that multiple classifications can be performed under a single excitation and imaging step. In some embodiments, the a sequencing reaction cycle comprising the contacting, detecting, and incorporating/extending steps is performed in less than 30 minutes in less than 20 minutes, or in less than 10 minutes. In some embodiments, the an average Q-score for base calling accuracy over a sequencing run is greater than or equal to 30, and/or greater than or equal to 40. In some embodiments, the at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the terminal nucleotides identified have a Q-score of greater than 30 and/or greater than or equal to 40. In some embodiments, the present disclosure provides said method, herein at least 95% of the terminal nucleotides identified have a Q-score of greater than 30.

[00319] The detecting of the methods described herein may enable identification of the target nucleic acid. In some cases, the target may comprise a ribonucleic acid. In some embodiments, detecting at least a portion of an index sequence may allow for identification of the corresponding ribonucleic acid sequence. In some embodiments, detecting may comprise detecting the identity of each nucleotide within the one or more index sequences as described herein. In some embodiments, detecting may comprise detecting the identity of a portion of the nucleotides within the one or more index sequences as described herein. In some cases, identifying the target may comprise detecting the nucleotides within the index sequence, described herein, and associating the index sequence to a target nucleic acid. Associating the index sequence to a target may comprise use of a codebook, a database, a reference table, or a combination thereof.

[00320] In some embodiments, the present disclosure provides a method wherein a sequencing reaction cycle comprising the contacting, detecting, and incorporating steps is performed in less than 30 minutes, less than 20 minutes, or in less than 10 minutes. In some embodiments, the present disclosure provides said method wherein an average Q-score for base calling accuracy over a sequencing run is greater than or equal to 30, or greater than or equal to 40. In some embodiments, the present disclosure provides said method wherein at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the terminal nucleotides identified have a Q-score of greater than 30; or greater than 40. In some embodiments, the present disclosure provides said method wherein at least 95% of the terminal nucleotides identified have a Q-score of greater than 30.

[00321] In some embodiments, the present disclosure provides a composition as disclosed herein for use in establishing or maintaining control over the persistence time of a signal from a labeled nucleic acid complex bound to or associated with a surface.

[00322] In some embodiments, the present disclosure provides a composition as disclosed herein for use in establishing or maintaining control over the persistence time of a fluorescence, luminescence, electrical, electrochemical, colorimetric, radioactive, magnetic, or electromagnetic signal from a labeled nucleic acid complex bound to or associated with a surface.

[00323] In some embodiments, the present disclosure provides a composition as disclosed herein for use in increasing the specificity, accuracy, or read length of a nucleic acid sequencing and/or genotyping application.

[00324] In some embodiments, the present disclosure provides a composition as disclosed herein for use in increasing the specificity, accuracy, or read length in a sequencing by binding, sequencing by synthesis, single molecule sequencing, or ensemble sequencing method.

[00325] In some embodiments, the present disclosure provides a reagent as disclosed herein for use in increasing the contrast to noise ratio (CNR) of a labeled nucleic acid complex bound to or associated with a surface.

[00326] In some embodiments, the present disclosure provides a reagent as disclosed herein for use in establishing or maintaining control over the persistence time of a signal from a labeled nucleic acid complex bound to or associated with a surface.

[00327] In some embodiments, the present disclosure provides a reagent as disclosed herein for use in establishing or maintaining control over the persistence time of a fluorescence, luminescence, electrical, electrochemical, colorimetric, radioactive, magnetic, or electromagnetic signal from a labeled nucleic acid complex bound to or associated with a surface.

[00328] In some embodiments, the present disclosure provides a reagent as disclosed herein for use in increasing the specificity, accuracy, or read length of a nucleic acid sequencing and/or genotyping application.

[00329] In some embodiments, the present disclosure provides a reagent as disclosed herein for use in increasing the specificity, accuracy, or read length in a sequencing by binding, sequencing by synthesis, single molecule sequencing, or ensemble sequencing method.

[00330] In some embodiments, where three of the four nucleotide conjugates (e.g. detectable nucleotide conjugates) are labeled, or where only three of the four nucleotide conjugates (e.g. detectable nucleotide conjugates) are present, the identity of the nucleotide corresponding to the unlabeled or absent conjugate may be established by the absence of a signal or by monitoring of the presence of unlabeled complexes such as by the identification of "dark" spots or unlabeled regions in a sequencing reaction. In some embodiments, detection of a polymerase complex incorporating a nucleotide conjugate (e.g. detectable nucleotide conjugate)may be carried out using two-color detection, such that conjugates corresponding to two of the four nucleotides are present in a sample, with two conjugates having a separate label corresponding to the nucleotide conjugated thereto and two conjugates having no label or being conjugated to an undetectable label. In some embodiments, only two of the four nucleotide conjugates (e.g. detectable nucleotide conjugates) are labeled. In some embodiments where two of the four nucleotide conjugates (e.g. detectable nucleotide conjugates) are labeled, the identity of the nucleotide corresponding to the unlabeled conjugate or conjugates may be established by the absence of a signal or by monitoring of the presence of unlabeled complexes such as by the identification of "dark" spots or unlabeled regions in a sequencing reaction. In some embodiments where two of the four nucleotide conjugates (e.g. detectable nucleotide conjugates) are labeled, the four nucleotide conjugates (e.g. detectable nucleotide conjugates) may be exposed to or contacted with the target sequentially, either individually, or in groups of two or three. In some embodiments two of the four nucleotide conjugates (e.g. detectable nucleotide conjugates) may share a common label, and the four nucleotide conjugates (e.g. detectable nucleotide conjugates) may be exposed to or contacted with the target sequentially, either individually, or in groups of two or three, wherein each contacting step shows the distinction between two or more different bases, such that after two, three, four, or more such contacting steps the identities of all unknown bases have been determined.

[00331] FIG. 1E shows the images captured on the surface after the binding complex is formed between the polymerase, the target nucleic acid, and the nucleotide conjugate (e.g. detectable nucleotide conjugate)having a nucleotide commentary to the next base of the target nucleic acid. The captured image includes four binding complexes 107a, 107b, 107c, and

107d formed on the surface, and each binding complex has a different nucleotide which can be distinguished based on the label (e.g., color) on the detectable nucleotide conjugate. Because of use of the nucleotide conjugate (e.g. detectable nucleotide conjugate)allows binding signals from a given sequence to originate within cluster regions containing multiple copies of the target sequence, the sequencing signals is greatly enhanced. Although FIG. 1E involves four detectable nucleotide conjugate, each having a different type of nucleotide, some methods can use one, two, or three nucleotide conjugates (e.g. detectable nucleotide conjugates), each having a different type of nucleotide and label. In some embodiments, each different type of nucleotide conjugate (e.g. detectable nucleotide conjugate)can be labeled either with the same label, or each with a label corresponding to the identity of its conjugated nucleotide. In some embodiments, three of four types of nucleotide conjugates can be labeled, with a fourth either unlabeled or conjugated to an undetectable label. In some embodiments, 1, 2, 3, or 4 detectable nucleotide conjugates can be labeled with a separate label, with, respectively, 3, 2, 1, or no nucleotide conjugates (e.g. detectable nucleotide conjugates) either unlabeled or conjugated to an undetectable label In some embodiments, a detection step can comprise simultaneous and/or serial excitation of up to 4 different excitation wavelengths, such as wherein the fluorescence imaging is carried out by detecting single and/or multiple fluorescence emission bands that uniquely classify each of the possible base pairing (A,G,C, or T). In some embodiments, four different nucleic acid binding compositions, each comprising a different nucleotide or nucleotide analog, may be used to determine the identity of the terminal nucleotide, wherein one of the four different nucleic acid binding compositions is labeled with a first fluorophore, one is labeled with a second fluorophore, one is labeled with both the first and second fluorophore, and one is not labeled, and wherein the detecting step comprises simultaneous excitation at a first excitation wavelength and a second excitation wavelength and images are acquired at a first fluorescence emission wavelength and a second fluorescence emission wavelength.

[00332] FIG.3 is a non-limiting example of a flowchart outlining the steps in sequencing a target nucleic acid. 301 describes a step of attaching target library sequences to a solid support surface by hybridizing the target molecules to complementary adapters on substrate surface. The target molecules can be single stranded or partially double stranded. Prior to 301, the nucleic acid molecules in the target library may have been prepared to contain fragments complementary to the adaptor sequences through ligation or other methods. 302 describes the step of clonal amplification to generate clusters of target molecules on the surface. 303 describes hybridizing sequencing primers to complementary primer binding sequences on the target to form the primed target nucleic acid. 304 describes combining the polymerase, the binding complex, which contains labeled (e.g., fluorescently-labeled) nucleotide conjugates (e.g. detectable

nucleotide conjugates), and the primed target nucleic acid. 304 may also include a step of washing or removing the unbound reagents including polymerase and detectable nucleotide conjugate.

[00333] In 305, when the nucleotide on the nucleotide conjugate (e.g. detectable nucleotide conjugate)is complementary to the next base of the primed target nucleic acid, the detectable nucleotide conjugate, polymerase, and primed target form a binding complex, which can be detected by detection methods (e.g., florescence imaging) compatible with the label on the detectable nucleotide conjugate. 305 can also include measuring the persistence time of the binding complex. In 306, the binding complex is destabilized to remove the binding of the nucleotide conjugate (e.g. detectable nucleotide conjugate)and polymerase. The dissociation can be achieved by placing the binding complex in a condition (e.g., adding Strontium ions) that will change the conformation of the polymerase and destabilize the binding. 306 may also include a step of washing or removing the dissociated nucleotide conjugate (e.g. detectable nucleotide conjugate)and/or polymerase. 307 describes the step of extending the primed strand of the primed target by a single base addition reaction. After the single base extension, steps 304, 305, 306, and 307 can be repeated in multiple cycles to determine the sequences of the target nucleic acid.

[00334] FIG. 4 is another non-limiting example of a flowchart outlining the steps in sequencing a target nucleic acid, which includes cleaving a nucleotide from the nucleotide conjugate (e.g. detectable nucleotide conjugate)and incorporating the cleaved nucleotide. 401 describes a step of attaching target library sequences to a solid support surface by hybridizing the target molecules to complementary adapters on substrate surface. The target molecules can be single stranded or partially double stranded. Prior to 401, the nucleic acid molecules in the target library may have been prepared to contain fragments complementary to the adaptor sequences through ligation or other methods. 402 describes the step of clonal amplification to generate clusters of target molecules on the surface. 403 describes hybridizing sequencing primers to complementary primer binding sequences on the target to form the primed target nucleic acid. 404 describes combining the polymerase, the binding complex, which contains labeled (e.g., fluorescently-labeled) nucleotide conjugates (e.g. detectable nucleotide conjugates), and the primed target nucleic acid. In the nucleotide conjugates (e.g. detectable nucleotide conjugates), the nucleotides are attached to the common core through chemical bonds or interactions that can be later severed. 404 may also include a step of washing or removing the unbound reagents including polymerase and detectable nucleotide conjugate.

[00335] In 405, when the nucleotide on the nucleotide conjugate (e.g. detectable nucleotide conjugate)is complementary to the next base of the primed target nucleic acid, the detectable nucleotide conjugate, polymerase, and primed target form a binding complex, which can be detected by detection methods (e.g., florescence imaging) compatible with the label on the

detectable nucleotide conjugate. 405 can also include measuring the persistence time of the binding complex. In 406, the polymerase is placed in a condition that would make it catalytically active to incorporate a nucleotide. The condition can include exposing the polymerase to Mg or Mn ions in the reaction solution. The nucleotide that is bound to the polymerase and the primed target is then cleaved from the particle and then incorporated into the primed strand of the primed target nucleic acid. The binding complex is destabilized. 406 may also include a step of washing or removing the dissociated nucleotide conjugate (e.g. detectable nucleotide conjugate)and/or polymerase. After the extension, steps 404, 405, and 406 can be repeated in multiple cycles to determine the sequences of the target nucleic acid.

[00336] FIG. 4 illustrates one exemplified method in which the binding complex is used for detecting a target nuclei acid. As shown in FIG. 4A, the polymer-nucleotide conjugate 201 is placed in contact with polymerase 206, a first nucleic acid molecule 204 and a second nucleic acid molecule 205. The polymer-nucleotide conjugate 201 has multiple polymer branches radiating from the core, and some branches are attached to nucleotide or oligonucleotide 202, and some branches are attached to a label 203. When the nucleotide or oligonucleotide 202 on the polymer-nucleotide conjugate 201 is complementary to at least a fraction of the first nucleic acid 204, a binding complex is formed as shown in FIG. 4B, and the strong binding signal can helps detect target with sequences complementary or partially complementary to the nucleotide or oligonucleotide on the polymer-nucleotide conjugate. In some instances, at least one of the polymerase, nucleic acid molecules, and nucleotide conjugates (e.g. detectable nucleotide conjugates) is attached to a solid support.

xii. Additional embodiments

[00337] Also disclosed herein are additional embodiments for the detection of a ribonucleic acid in situ.

[00338] The disclosed methods of determining the sequence of a target comprise: a) contacting a primed index sequence, wherein the primed index sequence comprises a double-stranded or partially double-stranded target molecule comprising the template strand to be sequenced and a primer strand to be elongated with one or more nucleotide conjugates (e.g. detectable nucleotide conjugates); and b) detecting the binding of the one or more nucleotide conjugates (e.g. detectable nucleotide conjugates) to the primed index sequence, thereby determining the presence of one of said one or more nucleotide conjugates (e.g. detectable nucleotide conjugates) on the primed index sequence and the identity of the next nucleotide (i.e., the N+1 or terminal nucleotide) to be incorporated into the complementary strand.

[00339] The sequencing method may further comprise incorporating the N+1 or terminal nucleotide into the primer strand, and then repeating the contacting, detecting, and incorporating steps for one or more additional iterations, thereby determining the sequence of the template strand of the nucleic acid molecule. After the step of detecting the binding complex, the primed strand of the primed target is extended for one base before another round of analysis is performed. The primed target can be extended using a conjugated nucleotide of the nucleotide conjugate (e.g. detectable nucleotide conjugate)that is attached to the polymer in the binding complex, or using an unconjugated or untethered free nucleotide that is provided after the binding complex has been removed.

[00340] The extension of the primed target may be prevented or inhibited due to a blocked nucleotide on the strand or the use of polymerase that is sufficiently catalytically inactive rendering the polymerase unable to incorporate a nucleotide into the primer of the primed nucleic acid. When the nucleotide in the nucleotide conjugate (e.g. detectable nucleotide conjugate)has a blocking group that prevents the extension of the nucleic acid, incorporation of a nucleotide may be achieved by the removal of a blocking group from said nucleotide (such as by detachment of said nucleotide from its polymer, branched polymer, dendrimer, common core, or the like). When the extension of the primed target is inhibited due to the use of polymerase that is catalytically inactive, incorporation of a nucleotide may be achieved by the provision of a cofactor or activator such as a metal ion.

[00341] Detection of the ternary complex is achieved prior to, concurrently with, or following the incorporation of the nucleotide residue. In some embodiments, a primed target may comprise a target with multiple primed locations for the attachment of polymerases and/or nucleic acid binding moieties. In some embodiments, multiple polymerases may be attached to a single target molecule, such as at multiple sites within a target molecule. In some embodiments, multiple polymerases may be bound to a binding complex disclosed herein comprising multiple nucleotides. In some embodiments, a target molecule may be a product of a strand displacement synthesis, a rolling circle amplification, a concatenation or fusion of multiple copies of a query sequence, or other such methods as are known in the art or as are disclosed elsewhere herein to produce nucleic acid molecules comprising multiple copies of an identical sequence. Therefore, in some embodiments, multiple polymerases may be attached at multiple identical or substantially identical locations within a target which comprises multiple identical or substantially identical copies of a query sequence. In some embodiments, said multiple polymerases may then be involved in interactions with one or more multivalent binding complexes; however, in preferred embodiments, the number of binding sites within a target is at least two, and the number of nucleotides or substrate moieties present on a nucleotide conjugate (e.g. detectable nucleotide

conjugate)such as a nucleotide conjugate (e.g. detectable nucleotide conjugate)is also greater than or equal to two.

[00342] It may be advantageous to provide the binding complexes in combination with other elements such as to provide optimized signals, for example to provide identification of a nucleotide at a particular position in a nucleic acid sequence. In some embodiments, the compositions disclosed herein are provided in combination with a surface providing low background binding or low levels of protein binding, especially a hydrophilic or polymer coated surface. Representative surfaces may be found, for example, in U.S. Patent Application No. 16/363,842, the contents of which are hereby incorporated by reference in their entirety.

[00343] In some instances, the nucleic acid molecule is tethered to the surface of a solid support, e.g., through hybridization of the template strand to an adapter nucleic acid sequence or primer nucleic acid sequence that is tethered to the solid support. In some instances, the solid support comprises a glass, fused-silica, silicon, or polymer substrate. In some instances, the solid support comprises a low non-specific binding coating comprising one or more hydrophilic polymer layers (e.g. PEG layers) where at least one of the hydrophilic polymer layers comprises a branched polymer molecule (e.g., a branched PEG molecule comprising 4, 8, 16, or 32 branches).

[00344] The solid support comprises oligonucleotide adapters or primers tethered to at least one hydrophilic polymer layer at a surface density ranging from about 1,000 primer molecules per μm^2 to about 1,000,000 primer molecules per μm^2. In some instances, the surface density of oligonucleotide 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 instances, the surface density of oligonucleotide 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 instances 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.

[00345] One of ordinary skill would recognize that in a series of iterative sequencing reactions, occasionally one or more sites will fail to incorporate a nucleotide during a given cycle, thus leading one or more sites to be unsynchronized with the bulk of the elongating nucleic acid chains. Under conditions in which sequencing signals are derived from reactions occurring on single copies of a target nucleic acid, these failures to incorporate will yield discrete errors in the output sequence. It is an object of the present disclosure to describe methods for reducing this type of error in sequencing reactions. For example, the use of nucleotide conjugates

(e.g. detectable nucleotide conjugates) that are capable of incorporation into the elongating strand, by providing increased probabilities of rebinding upon premature dissociation of a ternary polymerase complex, can reduce the frequency of "skipped" cycles in which a base is not incorporated. Thus, in some embodiments, the present disclosure contemplates the use of nucleotide conjugates (e.g. detectable nucleotide conjugates) as disclosed herein in which the nucleoside moiety is comprised within a nucleotide having a free, or reversibly modified, 5' phosphate, diphosphate, or triphosphate moiety, and wherein the nucleotide is connected to the common core or polymer as disclosed herein, through a labile or cleavable linkage. In some embodiments, the present disclosure contemplates a reduction in the intrinsic error rate due to skipped incorporations as a result of the use of the multivalent substrates disclosed herein.

[00346] The present disclosure also contemplates sequencing reactions in which sequencing signals from or relating to a given sequence are derived from or originate within definable regions containing multiple copies of the target sequence. Sequencing methods incorporating multiple copies of a target sequence have the advantage that signals can be amplified due to the presence of multiple simultaneous sequencing reactions within the defined region, each providing its own signal. The presence of multiple signals within a defined area also reduces the impact of any single skipped cycle, due to the fact that the signal from a large number of correct base calls can overwhelm the signal from a smaller number of skipped or incorrect base calls. The present disclosure further contemplates the inclusion of free, unlabeled nucleotides during elongation reactions, or during a separate part of the elongation cycle, in order to provide incorporation at sites that may have been skipped in previous cycles. For example, during or following an incorporation cycle, unlabeled blocked nucleotides may be added such that they may be incorporated at skipped sites. The unlabeled blocked nucleotides may be of the same type or types as the nucleotide attached to the multivalent binding substrate or substrates that are or were present during a particular cycle, or a mixture of 1, 2, 3, 4 or more types of unlabeled blocked nucleotides may be included.

[00347] When each sequencing cycle proceeds perfectly, each reaction within the defined region will provide an identical signal. However, as noted elsewhere herein, in a series of iterative sequencing reactions, occasionally one or more sites will fail to incorporate a nucleotide during a given cycle, thus leading one or more sites to be unsynchronized with the bulk of the elongating nucleic acid chains. This issue, referred to as "phasing," leads to degradation of the sequencing signal as the signal is contaminated with spurious signals from sites having skipped one or more cycles. This, in turn, creates the potential for errors in base identification. The progressive accumulation of skipped cycles through multiple cycles also reduces the effective read length, due to progressive degradation of the sequencing signal with each cycle. It is a further

object of this disclosure to provide methods for reducing phasing errors and/or to improve read length in sequencing reactions.

[00348] The sequencing method can include contacting a target or multiple targets, comprising multiple linked or unlinked copies of a target sequence, with the binding complexes described herein. Contacting said target nucleic acid, or multiple targets comprising multiple linked or unlinked copies of a target sequence, with one or more nucleotide conjugates (e.g. detectable nucleotide conjugates) may provide a substantially increased local concentration of the correct nucleotide being interrogated in a given sequencing cycle, thus suppressing signals from improper incorporations or phased nucleic acid chains (e.g., those elongating nucleic acid chains which have had one or more skipped cycles).

[00349] Methods of obtaining nucleic acid sequence information can include contacting a target nucleic acid, or multiple targets, wherein said target or multiple targets comprise multiple linked or unlinked copies of a target sequence, with one or more nucleotide conjugates (e.g. detectable nucleotide conjugates). This method results in a reduction in the error rate of sequencing as indicated by reduction in the misidentification of bases, the reporting of nonexistent bases, or the failure to report correct bases. In some embodiments, said reduction in the error orate of sequencing may comprise a reduction of 5%, 10%, 15%, 20% 25%, 50%, 75%, 100%, 150%, 200%, or more compared to the error rate observed using monovalent ligands, including free nucleotides, labeled free nucleotides, protein or peptide bound nucleotides, or labeled protein or peptide bound nucleotides.

[00350] The method of obtaining nucleic acid sequence information can include contacting a target nucleic acid, or multiple targets, wherein said templet nucleic acid or multiple targets comprise multiple linked or unlinked copies of a target sequence, with one or more nucleotide conjugates (e.g. detectable nucleotide conjugates). This method results in an increase in average read length of 5%, 10%, 15%, 20% 25%, 50%, 75%, 100%, 150%, 200%, 300%, or more compared to the average read length observed using monovalent ligands, including free nucleotides, labeled free nucleotides, protein or peptide bound nucleotides, or labeled protein or peptide bound nucleotides.

[00351] Methods of obtaining nucleic acid sequence information, said methods comprising contacting a target nucleic acid, or multiple targets, wherein said target or multiple targets comprise multiple linked or unlinked copies of a target sequence, with one or more nucleotide conjugates (e.g. detectable nucleotide conjugates). This method results in an increase in average read length of 10 nucleotides (NT), 20NT, 25NT, 30NT, 50NT, 75NT, 100NT, 125NT, 150NT, 200NT, 250NT, 300NT, 350NT, 400NT, 500NT , or more compared to the average read

length observed using monovalent ligands, including free nucleotides, labeled free nucleotides, protein or peptide bound nucleotides, or labeled protein or peptide bound nucleotides.

[00352] The use of binding complex for sequencing effectively shortens the sequencing time. The sequencing reaction cycle comprising the contacting, detecting, and incorporating steps is performed in a total time ranging from about 5 minutes to about 60 minutes. In some instances, the sequencing reaction cycle is performed in at least 5 minutes, at least 10 minutes, at least 20 minutes, at least 30 minutes, at least 40 minutes, at least 50 minutes, or at least 60 minutes. In some instances, the sequencing reaction cycle is performed in at most 60 minutes, at most 50 minutes, at most 40 minutes, at most 30 minutes, at most 20 minutes, at most 10 minutes, or at most 5 minutes. 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 instances the sequencing reaction cycle may be performed in a total time ranging from about 10 minutes to about 30 minutes. Those of skill in the art will recognize that the sequencing cycle time may have any value within this range, e.g., about 16 minutes.

[00353] The use of binding complex for sequencing provides an more accuracy base readout. The disclosed compositions and methods for nucleic acid sequencing will provide an average Q-score for base-calling accuracy over a sequencing run that ranges from about 20 to about 50. In some instances, the average Q-score is at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, or at least 50. Those of skill in the art will recognize that the average Q-score may have any value within this range, e.g., about 32.

[00354] In some instances, the disclosed compositions and methods for nucleic acid sequencing will provide a Q-score of greater than 30 for at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% of the terminal (or N+1) nucleotides identified. In some instances, the disclosed compositions and methods for nucleic acid sequencing will provide a Q-score of greater than 35 for at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% of the terminal (or N+1) nucleotides identified. In some instances, the disclosed compositions and methods for nucleic acid sequencing will provide a Q-score of greater than 40 for at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% of the terminal (or N+1) nucleotides identified. In some instances, the disclosed compositions and methods for nucleic acid sequencing will provide a Q-score of greater than 45 for at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% of the terminal (or N+1) nucleotides identified. In some instances, the disclosed compositions and methods for nucleic acid sequencing will provide a Q-score of

greater than 50 for at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% of the terminal (or N+1) nucleotides identified.

[00355] The disclosed low non-specific binding supports and associated nucleic acid hybridization and amplification methods may be used for the analysis of nucleic acid molecules derived from any of a variety of different cell, tissue, or sample types known to those of skill in the art. For example, nucleic acids may be extracted from cells, or tissue samples comprising one or more types of cells, derived from eukaryotes (such as animals, plants, fungi, protista), archaebacteria, or eubacteria. In some cases, nucleic acids may be extracted from prokaryotic or eukaryotic cells, such as adherent or non-adherent eukaryotic cells. Nucleic acids are variously extracted from, for example, primary or immortalized rodent, porcine, feline, canine, bovine, equine, primate, or human cell lines. Nucleic acids may be extracted from any of a variety of different cell, organ, or tissue types (e.g., white blood cells, red blood cells, platelets, epithelial cells, endothelial cells, neurons, glial cells, astrocytes, fibroblasts, skeletal muscle cells, smooth muscle cells, gametes, or cells from the heart, lungs, brain, liver, kidney, spleen, pancreas, thymus, bladder, stomach, colon, or small intestine). Nucleic acids may be extracted from normal or healthy cells. Alternately or in combination, acids are extracted from diseased cells, such as cancerous cells, or from pathogenic cells that are infecting a host. Some nucleic acids may be extracted from a distinct subset of cell types, e.g., immune cells (such as T cells, cytotoxic (killer) T cells, helper T cells, alpha beta T cells, gamma delta T cells, T cell progenitors, B cells, B-cell progenitors, lymphoid stem cells, myeloid progenitor cells, lymphocytes, granulocytes, Natural Killer cells, plasma cells, memory cells, neutrophils, eosinophils, basophils, mast cells, monocytes, dendritic cells, and/or macrophages, or any combination thereof), undifferentiated human stem cells, human stem cells that have been induced to differentiate, rare cells (e.g., circulating tumor cells (CTCs), circulating epithelial cells, circulating endothelial cells, circulating endometrial cells, bone marrow cells, progenitor cells, foam cells, mesenchymal cells, or trophoblasts). Other cells are contemplated and consistent with the disclosure herein.

[00356] Nucleic acid extraction from cells or other biological samples may be performed using any of a number of techniques known to those of skill in the art. For example, a typical DNA extraction procedure comprises (i) collection of the cell sample or tissue sample from which DNA is to be extracted, (ii) disruption of cell membranes (i.e., cell lysis) to release DNA and other cytoplasmic components, (iii) treatment of the lysed sample with a concentrated salt solution to precipitate proteins, lipids, and RNA, followed by centrifugation to separate out the precipitated proteins, lipids, and RNA, and (iv) purification of DNA from the supernatant to remove detergents, proteins, salts, or other reagents used during the cell membrane lysis step.

[00357] A variety of suitable commercial nucleic acid extraction and purification kits are consistent with the disclosure herein. Examples include, but are not limited to, the QIAamp® kits (for isolation of genomic DNA from human samples) and DNAeasy kits (for isolation of genomic DNA from animal or plant samples) from Qiagen (Germantown, MD), or the Maxwell® and ReliaPrep™ series of kits from Promega (Madison, WI).

[00358] The methods described herein employ *in situ* sequencing including for example one-pass sequencing, reiterative short read sequencing and/or batch-sequencing. In some embodiments the sequencing can be conducted within a biological sample including a single cell, multiple cells, a tissue or a tumor.

[00359] In some embodiments, the methods comprise repeatedly conducting a short number of sequencing cycles of the same region of the template molecules. By conducting reiterative short sequencing cycles, the RNA content of the biological sample can be discovered. Compared to long read sequencing workflows, the reiterative short sequencing cycles described herein use a reduced amount of sequencing reagents which reduces cost and saves time. Methods for conducting reiterative short sequencing cycles has many uses including but not limited to detecting specific RNAs of interest, mutant RNA sequences, splice variants, and their abundance levels thereof.

[00360] In some embodiments, the methods comprise conducting batch sequencing of different sub-populations of template molecules wherein each sub-population of template molecules carries a batch-specific sequencing primer binding site.

[00361] One of the purposes of the methods described herein is to detect and image the spatial localization of RNAs within a biological sample using massively parallel sequencing.

[00362] The workflow comprises the following general steps. A biological sample is placed on a solid support which is positioned on a fluorescent microscope configured to detect fluorescent signals. The biological sample is treated with a chemical fixation reagent to retain the target RNAs inside the cells. The biological sample is also permeabilized to permit manipulation of the target RNAs inside the cells. The target RNAs are selectively hybridized to target-specific padlock probes for generating circularized padlock probes that correspond to particular target RNAs. The padlock probes carry at least a universal adaptor for a sequencing primer binding site. The padlock probes can optionally carry a target barcode sequence wherein the target barcode sequence corresponds to a given target RNA. In some embodiments, the target barcode sequence uniquely identifies a target RNA. The circularized padlock probes are subjected to a ligation and/or fill-in reaction to form covalently closed circular padlock probes which are amplified inside the biological sample via rolling circle amplification in the presence of optional compaction oligonucleotides to generate single-stranded concatemers. The concatemers carry tandem repeat

units of an RNA-of-interest, the universal sequencing primer binding site, and optionally the target barcode sequence. The concatemers are sequenced inside the biological sample. The full length of the RNA-of-interest in the concatemer can be sequenced. The full length of the target barcode can be sequenced.

[00363] Additionally, a short portion of the RNA or target barcode region in the concatemer can be re-sequenced at least once (e.g., reiterative sequencing) from the same start position to generate overlapping sequencing reads that can be aligned to a reference sequence. For example, the same portion of the concatemer molecule can be sequenced at least two, three, four, five, or up to 50 times. The start sequencing site can be any location of the concatemer and is dictated by the sequencing primers which are designed to anneal to a selected position within the concatemer. The reiterative short sequencing reads increase the redundancy of sequencing information for individual bases in the RNA region. Reiteratively sequencing one strand of the concatemer template molecule provides enough base coverage to reveal the presence of target RNAs in the biological sample so that pairwise sequencing of the complementary strand is not necessary.

[00364] A concatemer template molecule includes multiple sequencing primer binding sites along the same concatemer molecule which can be used to generate multiple usable sequencing reads for increased sequencing depth. Together, reiteratively sequencing one strand of the concatemer templates increases sequencing base coverage and sequencing depth compared to sequencing a one-copy template molecule.

[00365] The methods described herein can be conducted in uni-plex or multi-plex modes. Two or more different target RNAs can be detected and imaged simultaneously inside a biological sample using different target-specific padlock probes, and universal sequencing primers. For example, the presence of a housekeeping RNA and at least one target RNA in a biological sample can be simultaneously detected and imaged using any of the sequencing methods described herein.

[00366] In the methods described herein, the RNA is not extracted from the biological sample and sequencing information does not need to be tracked and mapped back to an image of the biological sample. Rather, RNA is retained inside the biological sample to permit direct imaging of the spatial location of target RNAs within the cells. Additionally, RNA within the biological sample is not fragmented and enrichment of target RNA is not necessary. Use of target-specific and/or random-sequence padlock probes enables detection of both poly-A and non-poly-A RNAs in either uni-plex or multi-plex modes.

[00367] The methods described herein offer several advantages over other *in situ* transcriptomics workflows, including a simpler workflow, fewer reagents, lower cost, less time, gentler conditions on the biological sample, and no requirement for specialized equipment.

Methods for Conducting *in situ* RNA Sequencing

[00368] The present disclosure provides methods for *in situ* sequencing comprising **step (a):** providing a biological sample deposited on a support, wherein the biological sample harbors a plurality of RNA including at least a first target RNA. In some embodiments, the biological sample harbors a plurality of RNA including at least a first sub-population of target RNA and a second sub-population of target RNA. In some embodiments, the biological sample is fixed and permeabilized.

[00369] In some embodiments, the biological sample of step (a) harbors 2-25 different target RNA molecules, or harbors 25-50 different target RNA molecules, or harbors 50-75 different target RNA molecules, or harbors 75-100 different target RNA molecules. In some embodiments, the biological sample harbors more than 100 different target RNA molecules, or more than 250 different target RNA molecules, or more than 500 different target molecules, or more than 1000 different target RNA molecules, or more. In some embodiments, the biological sample harbors more than 10,000 different target RNA molecules.

[00370] In some embodiments, the biological sample of step (a) comprises a whole cell, a plurality of whole cells, an intact tissue or an intact tumor. In some embodiments, the biological sample comprises a fresh biological sample, a freshly-frozen biological sample, a sectioned biological sample, or an FFPE biological sample. In some embodiments, the biological sample is deposited onto a solid support. In some embodiments, the biological sample is deposited onto a solid support which is passivated with a coating that promotes cell adhesion. In some embodiments, the biological sample is deposited on a support that lacks immobilized capture oligonucleotides. In some embodiments, the biological sample comprises an expanded biological sample that has been cultured in a simple or complex cell culture media.

[00371] In some embodiments, the biological sample of step (a) can be deposited (*e.g.*, seeded) onto a support comprising a planar or non-planar support. In some embodiments, the support comprises a solid or semi-solid support. In some embodiments, the support comprises a porous, semi-porous or non-porous support. The support can be made of any material such as glass, plastic or a polymer material. In some embodiments, the surface of the support can be coated with one or more compounds to produce a passivated layer on the support. In some embodiments, the passivated layer forms a porous or semi-porous layer.

[00372] In some embodiments, the biological sample of step (a) can be deposited (*e.g.*, seeded) onto a support which is passivated with a coating that promotes cell adhesion. In

some embodiments, the biological sample can be deposited on a support that lacks immobilized capture primers which can bind target polynucleotides from the biological sample. In some embodiments, the support can be coated with one or more compounds that generate a charged coated surface. In some embodiments, the support is coated with a lysine compound, poly-lysine compound, arginine compound, poly-arginine compound, poly(L)-ornithine compound, or an amino-terminated compound (*e.g.*, including amino-terminated PEG). The support can be coated with an unbranched compound, a branched compound, or a mixture of unbranched and branched compounds. In some embodiments, the support can be coated with modified peptides, including, for example and without limitation, cationic anti-microbial peptides or dual surface anti-microbial peptides. In some embodiments, the support can be coated with polycyclic peptide antibiotics comprising thioether amino acids lanthionine or methyllanthionine and/or unsaturated amino acids dehydroalaine and 2-aminoisobutryic acid. In some embodiments, the support can be coated with at least one small peptide such as melittin. In some embodiments, the support can be coated with a compound that promotes integrin-mediated cell adhesion. For example, and without limitation, the support can be coated with tripeptide arginyl-glycyl-aspartic acid (Arg-Gly-Asp; also known as RGD). In some embodiments, the support can be coated with amines or polymers having -NH_2 groups which promote cell adhesion, including for example polyethyleneimine (PEI) or polydopamine (PDA). In some embodiments, the support can be coated with proteoglycans, collagen, elastin, fibronectin, laminin and/or vitronectin. In some embodiments, the support can be coated with collagen type I, collagen type II, collagen type IV.

[00373] In some embodiments, the methods for *in situ* sequencing comprise **step (b):** providing a plurality of padlock probes wherein individual padlock probes in the plurality comprise a linear oligonucleotide comprising a first binding arm at one end and a second binding arm at the other end (e.g., first and second padlock binding arms) and a universal sequence region between the first and second binding arms, wherein the first and second binding arms of a padlock probe can hybridize to portions of an RNA molecule to form an RNA-padlock probe complex (e.g., FIGs. 14, 16 and 18). In some embodiments, the first and second binding arms of individual padlock probes can hybridize to proximal regions of an RNA molecule to form a nick or gap between the first and second binding arms. In some embodiments, the first binding arm of an individual padlock probe comprises a first target-specific sequence that selectively hybridizes to a first region of a target RNA molecule, and the second binding arm of the same individual padlock probe has a second target-specific sequence that selectively hybridizes to a second region of the same target RNA molecule, wherein a nick or gap is formed between the hybridized first and second binding arms, thereby circularizing the padlock probe (e.g., FIGs. 14, 16 and 18).

[00374] In some embodiments, in step (b), the universal sequence region of individual padlock probes comprises any one or any combination of two or more universal sequences including: (i) a universal binding site for a sequencing primer; (ii) a universal binding site for an amplification primer; and/or (iii) a universal binding site for a compaction oligonucleotide (e.g., FIGs. 14, 16 and 18).

[00375] In some embodiments, in step (b), individual padlock probes comprise an optional target barcode sequence between the first and second binding arms. In some embodiments, the target barcode sequence, and the first and second binding arms of a given padlock probe, correspond to a target RNA molecule.

[00376] In some embodiments, the plurality of padlock probes of step (b) comprise at least a first and second sub-population of padlock probes wherein the padlock probes in each sub-population can bind to their cognate RNA molecules (e.g., FIGs. 14, 16 and 18). In some embodiments, the padlock probes of the first sub-population can bind a first sub-population of target RNA molecules. In some embodiments, the padlock probes of the second sub-population can bind a second sub-population of target RNA molecules, wherein the first sub-population of RNA molecules differs from the second sub-population of RNA molecules.

[00377] In some embodiments, the methods for *in situ* sequencing comprise **step (c):** contacting the biological sample with the plurality of padlock probes wherein the contacting is conducted under a condition suitable for moving the plurality of the padlock probes into the biological sample and suitable for hybridizing individual padlock probes to at least a portion of their cognate target RNA inside the biological sample thereby forming a plurality of RNA-padlock probe complexes. In some embodiments, the contacting of step (c) comprises: hybridizing the first and second binding arms of individual padlock probes to proximal positions on a target RNA molecule to form a circularized padlock probe having a nick or gap between the hybridized first and second binding arms (e.g., FIGs. 14, 16 and 18).

[00378] In some embodiments, the contacting of step (c) comprises: hybridizing a first sub-population of padlock probes to a first sub-population of target RNA molecules inside the biological sample, wherein the first and second binding arms of individual padlock probes of the first sub-population hybridize to proximal positions on a target RNA molecule of the first sub-population to form a first RNA-padlock probe complex comprising a circularized padlock probe having a nick or gap between the hybridized first and second binding arms (e.g., FIG. 28 left, FIG. 30 left, and FIG. 32 left). In FIG. 28, the first padlock probe comprises (i) a first target barcode sequence (target BC-1) that uniquely identifies the first target RNA, (ii) a first sequencing primer binding site (Seq-1) (or a complementary sequence thereof), (iii) a universal binding site for an amplification primer (universal RCA) (or a complementary sequence thereof), and (iv) a universal

binding site for a compaction oligonucleotide (or a complementary sequence thereof). The second padlock probe comprises (i) a second target barcode sequence (target BC-2) that uniquely identifies the second target RNA, (ii) a first sequencing primer binding site (Seq-1) (or a complementary sequence thereof), (iii) a universal binding site for an amplification primer (universal RCA) (or a complementary sequence thereof), and (iv) a universal binding site for a compaction oligonucleotide (or a complementary sequence thereof).

[00379] FIG. 29 shows a rolling circle and sequencing workflow comprising generating first and second concatemers by conducting rolling circle amplification using first and second circular nucleic acids (respectively). The first and second concatemers are subjected to a sequencing workflow essentially simultaneously using a plurality of first sequencing primers, a plurality of sequencing polymerases, and a plurality of nucleotide reagents. The first and second concatemers undergo sequencing essentially simultaneously.

[00380] In FIG. 30, the first padlock probe comprises (i) a first batch-specific sequencing primer binding site (Batch Seq-1) (or a complementary sequence thereof), (ii) a universal binding site for an amplification primer (universal RCA) (or a complementary sequence thereof), and (iii) a universal binding site for a compaction oligonucleotide (or a complementary sequence thereof). The second padlock probe comprises (i) a second batch-specific sequencing primer binding site (Batch Seq-2) (or a complementary sequence thereof), (ii) a universal binding site for an amplification primer (universal RCA) (or a complementary sequence thereof), and (iii) a universal binding site for a compaction oligonucleotide (or a complementary sequence thereof).

[00381] FIG. 31 shows a rolling circle and sequencing workflow comprising generating first and second concatemers by conducting rolling circle amplification using first and second circular nucleic acids (respectively). The first and second concatemers are subjected to a first sequencing workflow using first batch-specific sequencing primers, sequencing polymerases, and a plurality of nucleotide reagents. The first concatemers undergo first batch sequencing but the second concatemers do not. The first and second concatemers are subjected to a second sequencing workflow using second batch-specific sequencing primers, sequencing polymerases, and a plurality of nucleotide reagents. The second concatemers undergo second batch sequencing but the first concatemers do not.

[00382] In. FIG. 32, the first padlock probe comprises (i) a first target barcode sequence (target BC-1) that uniquely identifies the first target RNA, (ii) a first batch-specific sequencing primer binding site (Batch Seq-1) (or a complementary sequence thereof), (iii) a universal binding site for an amplification primer (universal RCA) (or a complementary sequence thereof), and (iv) a universal binding site for a compaction oligonucleotide (or a complementary sequence thereof). The second padlock probe comprises (i) a second target barcode sequence

(target BC-2) that uniquely identifies the second target RNA, (ii) a second batch-specific sequencing primer binding site (Batch Seq-2) (or a complementary sequence thereof), (iii) a universal binding site for an amplification primer (universal RCA) (or a complementary sequence thereof), and (iv) a universal binding site for a compaction oligonucleotide (or a complementary sequence thereof).

[00383] FIG. 33 shows This figure shows a rolling circle and sequencing workflow comprising generating first and second concatemers by conducting rolling circle amplification using first and second circular nucleic acids (respectively). The first and second concatemers are subjected to a first sequencing workflow using first batch-specific sequencing primers, sequencing polymerases, and a plurality of nucleotide reagents. The first concatemers undergo first batch sequencing but the second concatemers do not. The first and second concatemers are subjected to a second sequencing workflow using second batch-specific sequencing primers, sequencing polymerases, and a plurality of nucleotide reagents. The second concatemers undergo second batch sequencing but the first concatemers do not.

[00384] In FIG. 35, the concatemer includes tandem repeat units where each unit comprises: (i) a universal sequencing primer binding site (Seq), (ii) universal compaction oligonucleotide binding site (CO), (iii) a sequence that corresponds to a target RNA, and (iv) a target barcode sequence that corresponds to the given target RNA (BC). In some embodiments, universal sequencing primers (solid arrows) hybridize to the universal sequencing primer binding sites and no more than 50 sequencing cycles are conducted to generate a plurality of first sequencing read products (dashed arrows), where the first sequencing read products include only the target barcode sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include only the target barcode sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is once again repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include only the target barcode sequence. In some embodiments, the reiterative sequencing can be conducted up to 50 times. The sequences of all of the first sequencing read products can be determined and aligned with a first reference sequence (e.g., reference barcode sequence) to confirm the presence of the first target RNA molecules inside the biological sample.

[00385] In FIG. 36, the concatemer includes tandem repeat units where each unit comprises: (i) a universal sequencing primer binding site (Seq), (ii) universal compaction oligonucleotide binding site (CO), (iii) a sequence that corresponds to a target RNA, and (iv) a

target barcode sequence that corresponds to the given target RNA (BC). In some embodiments, universal sequencing primers (solid arrows) hybridize to the universal sequencing primer binding sites and no more than 50 sequencing cycles are conducted to generate a plurality of first sequencing read products (dashed arrows), where the first sequencing read products include the target barcode sequence and a portion of the insert sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include the target barcode sequence and a portion of the insert sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is once again repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include the target barcode sequence and a portion of the insert sequence. In some embodiments, the reiterative sequencing can be conducted up to 50 times. The sequences of all of the first sequencing read products can be determined and aligned with a first reference sequence (e.g., reference barcode sequence and the insert sequence that corresponds to the target RNA) to confirm the presence of the first target RNA molecules inside the biological sample.

[00386] In FIG. 37, the concatemer includes tandem repeat units where each unit comprises: (i) a universal sequencing primer binding site (Seq), (ii) universal compaction oligonucleotide binding site (CO), and (iii) a sequence that corresponds to a target RNA. In some embodiments, universal sequencing primers (solid arrows) hybridize to the universal sequencing primer binding sites and no more than 50 sequencing cycles are conducted to generate a plurality of first sequencing read products (dashed arrows), where the first sequencing read products include a portion of the insert sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include a portion of the insert sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is once again repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include a portion of the insert sequence. In some embodiments, the reiterative sequencing can be conducted up to 50 times. The sequences of all of the first sequencing read products can be determined and aligned with a first reference sequence (e.g., the insert sequence that corresponds to the target RNA) to confirm the presence of the first target RNA molecules inside the biological sample.

[00387] In FIG. 38, the concatemer includes tandem repeat units where each unit comprises: (i) a universal sequencing primer binding site (Seq) and (ii) a sequence that corresponds to a target RNA. In some embodiments, universal sequencing primers (solid arrows) hybridize to the universal sequencing primer binding sites and no more than 50 sequencing cycles are conducted to generate a plurality of first sequencing read products (dashed arrows), where the first sequencing read products include a portion of the insert sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include a portion of the insert sequence. The plurality of first sequencing read products are removed from the concatemer, and the sequencing is once again repeated where no more than 50 sequencing cycles are conducted to generate another plurality of first sequencing read products (dashed arrows), where the first sequencing read products include a portion of the insert sequence. In some embodiments, the reiterative sequencing can be conducted up to 50 times. The sequences of all of the first sequencing read products can be determined and aligned with a first reference sequence (e.g., the insert sequence that corresponds to the target RNA) to confirm the presence of the first target RNA molecules inside the biological sample.

[00388] In some embodiments, the contacting of step (c) comprises: hybridizing a second sub-population of padlock probes to a second sub-population of target RNA molecules inside the biological sample, wherein the first and second binding arms of individual padlock probes of the second sub-population hybridize to proximal positions on a target RNA molecule of the second sub-population to form a second RNA-padlock probe complex comprising a circularized padlock probe having a nick or gap between the hybridized first and second binding arms (e.g., FIG. 28 right, FIG. 30 right and FIG. 32 right).

[00389] In some embodiments, the size of the gap between the hybridized first and second binding arms is 1-25 bases. The 3'OH end of hybridized padlock probe can serve as an initiation site for a polymerase-catalyzed fill-in reaction (e.g., a reverse transcription gap fill-in reaction) using the target RNA molecule as a template. After the fill-in reaction, the remaining nick can be enzymatically ligated to generate covalently closed circular padlock probes.

[00390] In some embodiments, the methods for *in situ* sequencing comprise **step (d):** contacting the biological sample with a padlock probe closing reagent, wherein the contacting is conducted under a condition suitable for moving the padlock probe closing reagent into the biological sample, and the condition is suitable for generating a plurality of covalently closed padlock probes (e.g., FIGs. 28, 30 and 32).

[00391] In some embodiments, the padlock probe closing reagent of step (d) comprises a ligase enzyme when the circularize padlock probes from step (c) comprise a nick between the hybridized first and second binding arms. In some embodiments, the ligase enzyme can discriminate between matched and mis-matched hybridized ends to ensure target-specific hybridization. In some embodiments, the ligation reaction comprises use of a ligase enzyme, including a T3, T4, T7 or Taq DNA ligase enzyme (e.g., FIGs. 28, 30 and 32).

[00392] In some embodiments, the padlock probe closing reagent of step (d) comprises a reverse transcription gap fill-in reagent and a ligase enzyme when the circularize padlock probes from step (c) comprise a gap between the hybridized first and second binding arms. In some embodiments, the 3' extendible end of one of the binding arms of a padlock probe can be used as an initiation site for the reverse transcription gap fill-in reaction and uses a portion of the RNA molecule as a template molecule thereby forming a circularized padlock probe having a nick. The nick can be closed by conducting an enzymatic ligation reaction to form a covalently closed padlock probe (e.g., FIGs. 28, 30 and 32).

[00393] In some embodiments, in step (d), the reverse transcription gap fill-in reagent comprises a plurality of reverse transcriptase enzymes and a plurality of nucleotides. In some embodiments, the reverse transcriptase enzyme comprises a reverse transcriptase enzyme from AMV (avian myeloblastosis virus), M-MuLV (moloney murine leukemia virus), or HIV (human immunodeficiency virus). In some embodiment, the reverse transcriptase enzyme comprises a recombinant enzyme that exhibits reduced RNase H activity, for example REVERTAID (e.g., from Thermo Fisher Scientific, catalog No. EP0441). In some embodiments, the reverse transcriptase can be a commercially-available enzyme, including MULTISCRIBE (e.g., from Thermo Fisher Scientific, catalog # 4311235), THERMOSCRIPT (e.g., from Thermo Fisher Scientific, catalog # 12236-014), or ARRAYSCRIPT (e.g., from Ambion, catalog No. AM2048). In some embodiments, the reverse transcriptase enzyme comprises SUPERSCRIPT II (e.g., catalog No. 18064014), SUPERSCRIPT III (e.g., catalog No. 18080044), or SUPERSCRIPT IV enzymes (e.g., catalog No. 18090010) (all SUPERSCRIPT enzymes from Invitrogen). In some embodiments, the reverse transcription reaction can include an RNase inhibitor. In some embodiments, the plurality of nucleotides comprises any combination of dATP, dCTP, dGTP, dTTP and/or dUTP. In some embodiments, the ligase enzyme can discriminate between matched and mis-matched hybridized ends to ensure target-specific hybridization. In some embodiments, the ligation reaction comprises use of a ligase enzyme, including a T3, T4, T7 or Taq DNA ligase enzyme.

[00394] In some embodiments, the methods for *in situ* sequencing comprise step (e): contacting the biological sample with a rolling circle amplification reagent, wherein the

contacting is conducted under a condition suitable for moving the rolling circle amplification reagent into the biological sample. In some embodiments, the rolling circle amplification reagent comprises a plurality of strand-displacing DNA polymerases and a plurality of nucleotides. In some embodiments, the rolling circle amplification reagent further comprises a plurality of soluble amplification primers (e.g., FIGs. 14-15,FIGs. 16-17 and FIGs. 18-19).

[00395] In some embodiments, the plurality of strand-displacing DNA polymerases of step (e) comprise polymerases that exhibit the ability to displace a complementary strand from a template strand and catalyze new strand synthesis. Strand displacing polymerases include, for example and without limitation, mesophilic and thermophilic polymerases. Strand displacing polymerases include, for example and without limitation, wild type enzymes, and variants including exonuclease minus mutants, mutant versions, chimeric enzymes and truncated enzymes. Examples of strand displacing polymerases include, for example and 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™).

[00396] In some embodiments, the plurality of nucleotides of step (e) comprise any combination of dATP, dGTP, dCTP, dTTP and/or dUTP. In some embodiments, at least one of the nucleotides comprises a scissile moiety. For example, a nucleotide having a scissile moiety can be incorporated in a primer extension reaction and the incorporated nucleotide can be enzymatically converted into an abasic site. In some embodiments, nucleotides comprising a scissile moiety include uridine, 8-oxo-7,8-dihydrogunine and deoxyinosine. In some embodiments, uridine can be converted to an abasic site using uracil DNA glycosylase (UDG). In some embodiments, 8oxoG can be converted to an abasic site using formamidopyrimidine [fapy]-DNA glycosylase (FPG glycosylase). In some embodiments, deoxyinosine can be converted to an abasic site using 3-Methyladenine DNA glycosylase II (AlkA glycosylase). In some embodiments, the plurality of nucleotides comprises at least one nucleotide can be labeled with a detectable reporter moiety. For example, a detectable reporter moiety can comprise a fluorophore.

[00397] In some embodiments, individual soluble amplification primers of step (e) can hybridize to a universal binding sequence on a padlock probe, wherein the padlock probe that is hybridized to its cognate RNA molecule. In some embodiments, 3′ end of the soluble

amplification primers are extendible. In some embodiments, 3′ end of the soluble amplification primers provides an initiation site for rolling circle amplification.

[00398] In some embodiments, the soluble amplification primers of step (e) can be any length, for example 5-25 bases, or 25-50 bases, or 50-75 bases, or 75-100 bases in length or longer. The soluble amplification primers each comprise a 5' end and 3' end. In some embodiments, the 3' end of the soluble amplification primers can include a 3' OH moiety which serves as a nucleotide polymerization initiation site in a polymerase-catalyzed primer extension reaction. In some embodiments, the 3' end of the soluble amplification primers have a chain terminating moiety which blocks a polymerase-catalyzed primer extension reaction. The chain terminating moiety can be removed and converted to an extendible 3'OH moiety.

[00399] In some embodiments, the soluble amplification primers of step (e) can be modified to confer resistance to nuclease degradation (e.g., ribonuclease degradation). For example, the soluble amplification primers comprise at least one phosphorothioate diester bond at their 5' ends which can render the soluble amplification primers resistant to nuclease degradation. In some embodiments, the soluble amplification primers comprise 2-5 or more consecutive phosphorothioate diester bonds at their 5' ends. In some embodiments, the plurality of soluble amplification primers comprise at least one ribonucleotide and/or at least one 2'-O-methyl, 2'-O-methoxyethyl (MOE), 2' fluoro-base nucleotide. In some embodiments, the soluble amplification primers comprise phosphorylated 3' ends. In some embodiments, the soluble amplification primers comprise locked nucleic acid (LNA) bases. In some embodiments, the soluble amplification primers comprise a phosphorylated 5' end (e.g., using a polynucleotide kinase).

[00400] In some embodiments, the methods for *in situ* sequencing comprise **step (f):** conducting rolling circle amplification inside the biological sample thereby generating a plurality of concatemers inside the biological sample. In some embodiments, the concatemers comprise tandem repeat units, wherein a unit comprises a complementary sequence of a portion of the target RNA molecule and any universal sequence(s) present in the padlock probe of step (b). In some embodiments, the tandem repeat units include a complementary sequence of (i) a universal binding site for a sequencing primer; (ii) a universal binding site for an amplification primer; and/or (iii) a universal binding site for a compaction oligonucleotide (e.g., FIGs. 15, 17 and 19).

[00401] In some embodiments, the rolling circle amplification reaction of step (f) can be conducted using a suitable temperature and/or pH that promotes activity of the strand-displacing polymerase of step (e).

[00402] In some embodiments, the rolling circle amplification of step (f) generates a first sub-population of concatemer molecules wherein individual concatemer molecules in the

first sub-population comprise tandem repeat units, wherein a unit comprises a complementary sequence of a portion of the first target RNA molecule and any universal sequence(s) present in the first padlock probe of step (b). In some embodiments, the tandem repeat units include a complementary sequence of (i) a universal binding site for a first sequencing primer; (ii) a universal binding site for a first amplification primer; and/or (iii) a universal binding site for a first compaction oligonucleotide (e.g., FIG. 29 left, FIG. 31 left and FIG. 33 left).

[00403] In some embodiments, the rolling circle amplification of step (f) generates a second sub-population of concatemer molecules wherein individual concatemer molecules in the second sub-population comprise tandem repeat units, wherein a unit comprises a complementary sequence of a portion of the second target RNA molecule and any universal sequence(s) present in the second padlock probe of step (b). In some embodiments, the tandem repeat units include a complementary sequence of (i) a universal binding site for a second sequencing primer; (ii) a universal binding site for a second amplification primer; and/or (iii) a universal binding site for a second compaction oligonucleotide (e.g., FIG. 29 right, FIG. 31 right and FIG. 33 right).

[00404] In some embodiments, the first and second universal sequencing primer binding sites have the same sequence or different sequences.

[00405] In some embodiments, the first and second universal amplification primer binding sites have the same sequence or different sequences.

[00406] In some embodiments, the first and second universal compaction oligonucleotide binding sites have the same sequence or different sequences.

[00407] In some embodiments, the rolling circle amplification reaction of step (f) can be conducted in the presence or absence of a plurality of compaction oligonucleotides. In some embodiments, the compaction oligonucleotides bind portions of concatemer molecules generated by conducting a rolling circle amplification reaction. In some embodiments, individual compaction oligonucleotides comprise a 5′ region that binds a first portion of the concatemer molecule and a 3′ region that binds a second portion of the same concatemer molecule. In some embodiments, the 5′ and 3′ regions of a compaction oligonucleotide can hybridize to binding sites in a concatemer molecule to pull together distal portions of the concatemer molecule causing compaction of the concatemer molecule, *e.g.*, to form a DNA nanoball. In some embodiments, a compaction oligonucleotide can bind a portion of the concatemer molecule having a universal binding sequence for a compaction oligonucleotide.

[00408] In some embodiments, in the rolling circle amplification reaction of step (f), the plurality of nucleotides can include dATP, dGTP, dCTP, dTTP and/or dUTP, and the resulting concatemer molecule can be cross-linked to a cross-linking reactive group by treating

the biological sample with a succinimide ester (NHS), maleimide (Sulfo-SMCC), imidoester (DMP), carbodiimide (DCC, EDC) or phenyl azide. In some embodiments, polymerization of the cross-linking reactive group can be initiated with light or ultraviolet (UV) light. In some embodiments, the resulting concatemer molecule can be cross-linked to an intracellular matrix, *e.g.*, by treating the biological sample with a cross-linked agarose, cross-linked dextran or cross-linked polyethylene glycol (PEG), polyacrylamide, cellulose alginate or polyamide. In some embodiments, the PEG comprises a sulfo-NHS ester moiety at one or both ends, for example a PEGylated bis(sulfosuccinimidyl)suberate) (*e.g.*, $BS(PEG)_9$ from Thermo Fisher Scientific™, catalog No. 21582).

[00409] In some embodiments, the methods for *in situ* sequencing comprise step (g): sequencing the plurality of concatemer molecules inside the biological sample. In some embodiments, the sequencing of step (g) comprises contacting the plurality of concatemers inside the biological sample with (i) a plurality of universal sequencing primers, (ii) a plurality of sequencing polymerases, and (iii) a plurality of nucleotide reagents, under a condition suitable for hybridizing the plurality of universal sequencing primers to their respective universal sequencing primer binding sites on the concatemers and conducting a plurality of polymerase-catalyzed sequencing cycles thereby generating a plurality of sequencing read products inside the biological sample (e.g., FIGs. 15, 17 and 19). In some embodiments, the plurality of sequencing primers comprise the same sequence or different sequences. In some embodiments, the plurality of nucleotide reagents comprises nucleotides, nucleotide analogs and/or multivalent molecules. In some embodiments, the plurality of nucleotide reagents comprises non-labeled or detectably labeled nucleotide reagents. In some embodiments, the plurality of sequencing read products comprise detectably labeled read products that are detectable by imaging. In some embodiments, the plurality of sequencing read products comprise non-labeled read products.

[00410] In some embodiments, the plurality of concatemer molecules inside the biological sample can be sequenced essentially simultaneously according to step (g1) (e.g., FIG. 29). In some embodiments, the sequencing of step (g1) comprises contacting the plurality of concatemers inside the biological sample with (i) a plurality of universal sequencing primers, (ii) a plurality of sequencing polymerases, and (iii) a plurality of nucleotide reagents, under a condition suitable for hybridizing the plurality of universal sequencing primers to their respective universal sequencing primer binding sites on the concatemers and conducting a plurality of polymerase-catalyzed sequencing cycles thereby generating a plurality of sequencing read products inside the biological sample. In some embodiments, the plurality of sequencing primers comprise the same sequence (e.g., FIG. 29). In some embodiments, the plurality of nucleotide reagents comprises nucleotides, nucleotide analogs and/or multivalent molecules. In some

embodiments, the plurality of nucleotide reagents comprises non-labeled or detectably labeled nucleotide reagents. In some embodiments, the plurality of sequencing read products comprise detectably labeled read products that are detectable by imaging. In some embodiments, the plurality of sequencing read products comprise non-labeled read products.

[00411] In some embodiments, the sequencing of step (g1) comprises conducting 2-50 polymerase-catalyzed sequencing cycles inside the biological sample, or conducting 50-100 polymerase-catalyzed sequencing cycles inside the biological sample, or conducting 100-150 polymerase-catalyzed sequencing cycles inside the biological sample, or conducting 150-200 polymerase-catalyzed sequencing cycles inside the biological sample, or conducting 200-250 polymerase-catalyzed sequencing cycles inside the biological sample, or conducting 250-300 polymerase-catalyzed sequencing cycles inside the biological sample, or conducting more than 300 polymerase-catalyzed sequencing cycles inside the biological sample. In some embodiments, the sequencing of step (g1) comprises conducting 300-500 polymerase-catalyzed sequencing cycles inside the biological sample.

[00412] In some embodiments, in the sequencing of step (g1), at least a portion of the target RNA region of the concatemer molecules can be sequenced. In some embodiments, the full length of the target RNA region of the concatemer molecules can be sequenced. In some embodiments, at least a portion of the target barcode sequence of the concatemer molecules can be sequenced. In some embodiments, the full length of the target barcode sequence of the concatemer molecules can be sequenced. In some embodiments, the RNA region and the target barcode sequence of the concatemer molecules can be sequenced.

[00413] In some embodiments, the sequencing of step (g1) comprises sequencing at least a portion of the plurality of concatemer molecules using an optical imaging system comprising a field-of-view (FOV) greater than 1.0 mm^2.

[00414] In some embodiments, in the sequencing of step (g1), the plurality of sequencing read products are detectable by imaging.

[00415] In some embodiments, in the sequencing of step (g1), the plurality of the sequencing read products are detectable by imaging, and wherein the sequencing comprises simultaneously imaging the plurality of detectable sequencing read products in the biological sample (e.g., co-localization of the plurality of sequencing read products).

[00416] In some embodiments, the plurality of concatemer molecules inside the biological sample can be sequenced in separate batches according to step (g2) (e.g., FIGs. 17 and 19). In some embodiments, in the sequencing of step (g2), the plurality of concatemer molecules comprises a first and second sub-population of concatemer molecules, wherein the concatemer

molecules of the first sub-population are sequenced in a first batch and the concatemer molecules of the second sub-population are sequenced in a subsequent second batch.

[00417] In some embodiments, in step (g2), the concatemers in the first sub-population comprise tandem repeat units wherein each unit comprises a first universal sequencing primer binding site, a first RNA region (or a complementary sequence thereof) and an optional first target barcode sequence (e.g., FIG. 31 left and FIG. 33 left).

[00418] In some embodiments, in step (g2), the concatemers in the second sub-population comprise tandem repeat units wherein each unit comprises a second universal sequencing primer binding site, a second RNA region (or a complementary sequence thereof) and an optional second target barcode sequence (e.g., FIG. 31 right and FIG. 33 right). In some embodiments, the first and second universal sequencing primer binding sites comprises different sequences and bind different sequencing primers (e.g., FIGs. 31 and 33). In some embodiments, the first and second RNA regions comprise different sequences. In some embodiments, the first and second target barcode sequences comprise different sequences.

[00419] In some embodiments, the sequencing of step (g2) comprises contacting the plurality of concatemers of the first sub-population inside the biological sample with (i) a first plurality of universal sequencing primers, (ii) a first plurality of sequencing polymerases, and (iii) a first plurality of nucleotide reagents, under a condition suitable for hybridizing the first plurality of universal sequencing primers to their respective universal sequencing primer binding sites on the concatemers of the first sub-population and conducting a plurality of polymerase-catalyzed sequencing cycles thereby generating a first plurality of sequencing read products inside the biological sample (e.g., FIG. 31 left and FIG. 33 left). In some embodiments, the first plurality of sequencing primers comprise the same sequence. In some embodiments, the first plurality of nucleotide reagents comprises nucleotides, nucleotide analogs and/or multivalent molecules. In some embodiments, the first plurality of nucleotide reagents comprises non-labeled or detectably labeled nucleotide reagents. In some embodiments, the first plurality of sequencing read products comprise detectably labeled read products that are detectable by imaging. In some embodiments, the first plurality of sequencing read products comprise non-labeled read products.

[00420] In some embodiments, the sequencing of step (g2) comprises conducting 2-50 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer molecules, or conducting 50-100 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer molecules, or conducting 100-150 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer molecules, or conducting 150-200 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer molecules, or conducting 200-250 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer

molecules, or conducting 250-300 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer molecules, or conducting more than 300 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer molecules. In some embodiments, the sequencing of step (g2) comprises conducting 300-500 polymerase-catalyzed sequencing cycles on the first sub-population of concatemer molecules.

[00421] In some embodiments, in the sequencing of step (g2), at least a portion of the target RNA region of the concatemer molecules of the first sub-population are sequenced. In some embodiments, the full length of the target RNA region of the concatemer molecules of the first sub-population are sequenced. In some embodiments, at least a portion of the target barcode sequence of the concatemer molecules of the first sub-population are sequenced. In some embodiments, the full length of the target barcode sequence of the concatemer molecules of the first sub-population are sequenced. In some embodiments, the RNA region and the target barcode sequence of the concatemer molecules of the first sub-population are sequenced.

[00422] In some embodiments, the sequencing of step (g2) comprises sequencing at least a portion of the plurality of concatemer molecules of the first sub-population using an optical imaging system comprising a field-of-view (FOV) greater than 1.0 mm^2.

[00423] In some embodiments, in the sequencing of step (g2), the plurality of sequencing read products of the first sub-population are detectable by imaging, wherein the sequencing comprises simultaneously imaging the plurality of detectable sequencing read products of the first sub-population in the biological sample (e.g., co-localization of the first sub-population of sequencing read products).

[00424] In some embodiments, the sequencing of step (g2) comprises, after sequencing the concatemers of the first sub-population, contacting the plurality of concatemers of the second sub-population inside the biological sample with (i) a second plurality of universal sequencing primers, (ii) a second plurality of sequencing polymerases, and (iii) a second plurality of nucleotide reagents, under a condition suitable for hybridizing the second plurality of universal sequencing primers to their respective universal sequencing primer binding sites on the concatemers of the second sub-population and conducting a plurality of polymerase-catalyzed sequencing cycles thereby generating a second plurality of sequencing read products inside the biological sample (e.g., FIG. 31 right and FIG. 33 right). In some embodiments, the second plurality of sequencing primers comprise the same sequence. In some embodiments, the second plurality of nucleotide reagents comprises nucleotides, nucleotide analogs and/or multivalent molecules. In some embodiments, the second plurality of nucleotide reagents comprises non-labeled or detectably labeled nucleotide reagents. In some embodiments, the second plurality of sequencing read products comprise detectably labeled read products that are detectable by

imaging. In some embodiments, the second plurality of sequencing read products comprise non-labeled read products.

[00425] In some embodiments, the sequencing of step (g2) comprises conducting 2-50 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules, or conducting 50-100 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules, or conducting 100-150 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules, or conducting 150-200 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules, or conducting 200-250 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules, or conducting 250-300 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules, or conducting more than 300 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules. In some embodiments, the sequencing of step (g2) comprises conducting 300-500 polymerase-catalyzed sequencing cycles on the second sub-population of concatemer molecules.

[00426] In some embodiments, in the sequencing of step (g2), at least a portion of the target RNA region of the concatemer molecules of the second sub-population are sequenced. In some embodiments, the full length of the target RNA region of the concatemer molecules of the second sub-population are sequenced. In some embodiments, at least a portion of the target barcode sequence of the concatemer molecules of the second sub-population are sequenced. In some embodiments, the full length of the target barcode sequence of the concatemer molecules of the second sub-population are sequenced. In some embodiments, the RNA region and the target barcode sequence of the concatemer molecules of the second sub-population are sequenced.

[00427] In some embodiments, the sequencing of step (g2) comprises sequencing at least a portion of the plurality of concatemer molecules of the second sub-population using an optical imaging system comprising a field-of-view (FOV) greater than 1.0 mm^2.

[00428] In some embodiments, in the sequencing of step (g2), the plurality of sequencing read products of the second sub-population are detectable by imaging, wherein the sequencing comprises simultaneously imaging the plurality of detectable sequencing read products of the second sub-population in the biological sample (e.g., co-localization of the second sub-population of sequencing read products).

[00429] In some embodiments, the plurality of concatemer molecules inside the biological sample can be reiteratively sequenced according to step (g3) (e.g., FIGs. 21-24). In some embodiments, the sequencing of step (g3) comprises step (1) contacting the plurality of concatemer molecules inside the biological sample with (i) a plurality of universal sequencing primers, (ii) a plurality of sequencing polymerases, and (iii) a plurality of nucleotide reagents,

under a condition suitable for hybridizing the plurality of universal sequencing primers to their respective universal sequencing primer binding sites on the concatemer molecules. In some embodiments, the sequencing further comprises step (2) conducting no more than 2-50 sequencing cycles to generate a first plurality of sequencing read products. In some embodiments, the sequencing read products generated in step (2) are no longer than 50 bases in length. In some embodiments, the sequencing further comprises step (3) removing the first plurality of sequencing read products from the concatemer molecules and retaining the plurality of concatemer molecules inside the biological sample. In some embodiments, the sequencing further comprises step (4) repeating steps (1) – (3) at least once. In some embodiments, step (4) comprises repeating steps (1) - (3) at least 2 times, at least 3 times, at least 4 times, at least 5 times, at least 6 times, at least 7 times, at least 8 times, at least 9 times, or at least 10 times. In some embodiments, step (4) comprises repeating steps (1) - (3) up to 10 times, up to 20 times, up to 30 time, up to 40 times, or up to 50 times (e.g., FIGs. 21-24).

[00430] In some embodiments, the reiterative sequencing can be conducting using a sequencing-by-binding procedure, labeled and/or non-labeled chain-terminating nucleotides, or multivalent molecules. Descriptions of these three sequencing methods is described below.

[00431] In some embodiments, in the sequencing of step (g3), the plurality of universal sequencing primers can be hybridized to concatemer molecules in step (1) with a hybridization reagent comprising an SSC buffer (e.g., 2X saline-sodium citrate) buffer with formamide (e.g., 10-20% formamide). The hybridization conditions comprise a temperature of about 20-30 °C, for about 10-60 minutes.

[00432] In some embodiments, in the sequencing of step (g3), the plurality of sequencing read products can be removed from the concatemers in step (3) and the plurality of concatemers can be retained inside the biological sample using a de-hybridization reagent comprising an SSC buffer (e.g., saline-sodium citrate) buffer, with or without formamide, at a temperature that promotes nucleic acid denaturation such as for example 30 – 90 °C.

[00433] In some embodiments, in the sequencing of step (g3), at least a portion of the target RNA region of the concatemer molecules can be sequenced. In some embodiments, the full length of the target RNA region of the concatemer molecules can be sequenced. In some embodiments, at least a portion of the target barcode sequence of the concatemer molecules can be sequenced. In some embodiments, the full length of the target barcode sequence of the concatemer molecules can be sequenced. In some embodiments, the RNA region and the target barcode sequence of the concatemer molecules can be sequenced.

[00434] In some embodiments, the sequencing of step (g3) comprises sequencing at least a portion of the plurality of concatemer molecules using an optical imaging system comprising a field-of-view (FOV) greater than 1.0 mm^2.

[00435] In some embodiments, in the sequencing of step (g3), the plurality of sequencing read products are detectable by imaging.

[00436] In some embodiments, in the sequencing of step (g3), the plurality of the sequencing read products are detectable by imaging, and wherein the sequencing comprises simultaneously imaging the plurality of detectable sequencing read products in the biological sample (e.g., co-localization of the plurality of sequencing read products).

[00437] In some embodiments, the plurality of nucleotide reagents of steps (g), (g1), (g2) and (g3) comprise a plurality of nucleotides that are detectably labeled or non-labeled. In some embodiments, individual nucleotides are linked to a detectable reporter moiety. In some embodiments, the detectable reporter moiety comprises a fluorophore. In some embodiments, the plurality of detectably labeled nucleotide analogs comprise a plurality of chain terminating nucleotides, where the chain terminating moiety is linked to the 3' nucleotide sugar position to form a 3' blocked nucleotide analog. In some embodiments, the chain terminating moiety can be removed to convert the 3' blocked nucleotide analog to an extendible nucleotide having a 3' OH group on the sugar. In some embodiments, the labeled nucleotide analogs are linked to a different fluorophore that corresponds to the nucleo-bases adenine, cytosine, guanine, thymine or uracil, where the different fluorophores emit a fluorescent signal during the sequencing of steps (g), (g1) and (g2). In some embodiments, a sequencing cycle comprises (1) contacting the concatemer/sequencing primer duplex with a sequencing polymerase and a detectably labeled chain terminating nucleotide under a condition suitable for polymerase-catalyzed incorporation of the detectably labeled chain terminating nucleotide into the terminal end of the sequencing primer, (2) detecting and imaging the fluorescent signal and color emitted by the incorporated chain terminating nucleotide, and (3) removing the chain terminating moiety (e.g., unblocking) and retaining the concatemer/sequencing primer duplex. In some embodiments, 2-500 sequencing cycles can be conducted on the plurality of concatemers inside the biological sample to generate a plurality of sequencing read products. In some embodiments, the sequencing read products can be 2-500 nucleotides in length. In some embodiments, the sequence of the first sequencing read product can be determined and aligned with a first reference sequence to confirm the presence of the first target RNA molecules inside the biological sample. In some embodiments, the sequence of the second sequencing read product can be determined and aligned with a second reference sequence to confirm the presence of the second target RNA molecules inside the biological sample. In some embodiments, the sequencing reactions are conducted on a sequencing apparatus

having a detector that captures fluorescent signals from the sequencing reactions inside the biological sample. The sequencing apparatus can be configured to relay the fluorescent signal data captured by the detector to a computer system that is programmed to display images of different fluorescent spots which are co-located in the biological sample, where individual fluorescent spots correspond to different target RNA molecules. In some embodiments, when the sequencing is conducted using different fluorescently-labeled nucleotide reagents that correspond to different nucleo-bases (e.g., A, G, C, T/U), then the images can have different color fluorescent spots co-located in the same biological sample at different sequencing cycles.

[00438] In some embodiments, out-of-sync phasing and/or pre-phasing events can occur during synchronized sequencing reactions on concatemer molecules (e.g., sequencing of steps (g), (g1), (g2) and (g3), where the sequencing reactions comprise polymerase-catalyzed sequencing reactions employing detectably labeled chain terminator nucleotides. In some embodiments, a sequencing reaction on one concatemer molecule moves ahead (e.g., pre-phasing) or falls behind (e.g., phasing) of the sequencing of the other concatemer molecules. During sequencing, a fluorescent signal is typically detected which corresponds to incorporation of a labeled chain terminator nucleotide. Thus, phasing and pre-phasing events can be detected and monitored using incorporation of a labeled chain terminator nucleotide.

[00439] In some embodiments, the plurality of nucleotide reagents of steps (g), (g1), (g2) and (g3) comprise a plurality of multivalent molecules each comprising a core attached to a plurality of nucleotide-arms, wherein the nucleotide-arms are attached to a nucleotide moiety (e.g., FIGs. 1-5). In some embodiments, individual multivalent molecules are labeled with a detectably reporter moiety. In some embodiments, the detectable reporter moiety comprises a fluorophore. In some embodiments, the core of the multivalent molecule is labeled with a fluorophore, wherein the fluorophore which is attached to a given core of the multivalent molecule corresponds to the nucleotide base (e.g., adenine, guanine, cytosine, thymine or uracil) of the nucleotide arm. In some embodiments, at least one of the nucleotide arms of the multivalent molecule comprises a linker and/or nucleotide base that is attached to a fluorophore wherein the fluorophore which is attached to a given nucleotide base corresponds to the nucleotide base (e.g., adenine, guanine, cytosine, thymine or uracil) of the nucleotide arm. In some embodiments, a sequencing cycle comprises (1) contacting the concatemer/sequencing primer duplex with a first sequencing polymerase to form a polymerase complex, (2) contacting the polymerase complex with a detectably labeled multivalent molecule under a condition suitable for binding a complementary nucleotide moiety of the multivalent molecule to the polymerase complex thereby forming a multivalent-binding complex, and the condition is suitable for inhibiting incorporation of the complementary nucleotide moiety into the terminal end of the sequencing primer, (3) detecting and imaging the

fluorescent signal and color emitted by the bound detectably labeled multivalent molecule, (4) removing the first sequencing polymerase and the bound detectably labeled multivalent molecule, and retaining the concatemer/sequencing primer duplex, (5) contacting the retained concatemer/sequencing primer duplex with a second sequencing polymerase and a non-labeled chain terminating nucleotide under a condition suitable for polymerase-catalyzed incorporation of the non-labeled chain terminating nucleotide into the terminal end of the sequencing primer, and (6) removing the chain terminating moiety (e.g., unblocking) and retaining the concatemer/sequencing primer duplex. 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 mm^2 and the cycle time for scanning large area (> 10 mm^2) can be less than 5 minutes. In some embodiments, 2-500 sequencing cycles can be conducted on the plurality of concatemers inside the biological sample to generate a plurality of sequencing read products. In some embodiments, the sequencing read products can be 2-500 nucleotides in length. In some embodiments, the sequence of the first sequencing read product can be determined and aligned with a first reference sequence to confirm the presence of the first target RNA molecules inside the biological sample. In some embodiments, the sequence of the second sequencing read product can be determined and aligned with a second reference sequence to confirm the presence of the second target RNA molecules inside the biological sample. In some embodiments, the sequencing reactions are conducted on a sequencing apparatus having a detector that captures fluorescent signals from the sequencing reactions inside the biological sample. The sequencing apparatus can be configured to relay the fluorescent signal data captured by the detector to a computer system that is programmed to display images of different fluorescent spots which are co-located in the biological sample, where individual fluorescent spots correspond to different target RNA molecules.

[00440] In some embodiments, when sequencing with detectably labeled multivalent molecules, step (2) in which multivalent-binding complexes are formed and step (3) in which the bound detectably labeled multivalent molecules are imaged and detected, the conditions are gentle compared to sequencing workflows that employ detectable labeled chain terminating nucleotides. For example, steps (2) and (3) can be conducted at a gentle temperature of about 35 – 45 °C, or about 39 – 42 °C. Steps (2) and (3) can be conducted at a gentle temperature which can help retain the compact size and shape of a DNA nanoball during multiple sequencing cycles (e.g., 2-500 cycles) which can improve FWHM (full width half maximum) of a spot image of the DNA nanoball inside a biological 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 um or smaller.

[00441] In some embodiments, out-of-sync phasing and/or pre-phasing events can occur during synchronized polymerase-catalyzed sequencing reactions employing detectably labeled multivalent molecules. During sequencing, a fluorescent signal can be detected which corresponds to binding of complementary nucleotide moiety of a multivalent molecule to the polymerase complex thereby forming a multivalent-binding complex. Thus, phasing and pre-phasing events can be detected and monitored using binding of labeled multivalent molecules. In some embodiments, when conducting 2-500 sequencing cycles with detectably labeled multivalent molecules, the phasing and/or pre-phasing rate can be less than about 5%, or less than about 1%, or less than about 0.01%, or less than about 0.001%. By contrast, the phasing and/or pre-phasing rates for conducting 2-500 sequencing cycles using labeled chain terminator nucleotides can be about 5%.

SYSTEMS

[00442] Disclosed herein, in some embodiments, are systems configured to perform any of the methods described herein. The systems described herein may comprise a computing device and instructions for performing any of the methods described herein. The computing device may include one or more computer processors, an operating system configured to perform executable instructions, a memory, a computer program including instructions for performing any of the methods described herein, or a combination thereof. The system may comprise a fluid flow controller and/or fluid dispensing system configured to sequentially and iteratively contact the primed index sequences with the disclosed polymerase and nucleotide conjugates (e.g. detectable nucleotide conjugates) and/or reagents. The contacting may be performed within one or more flow cells. In some instances, the flow cells may comprise a component of the system. In some instances, the flow cells may be removable and/or disposable components of the system.

[00443] An aspect of the present disclosure provides a system for identifying an RNA in situ. The system may comprise a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of any of the methods described herein related to an RNA. The system may also comprise the RNA.

[00444] Another aspect of the present disclosure provides a system for identifying a polypeptide in situ. The system may comprise a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a

memory, and a computer program including instructions comprising the method of any of the methods described herein related to a polypeptide. The system may also comprise the polypeptide. The system may also comprise a polypeptide-binding probe.

[00445] An aspect of the present disclosure provides a system for identifying an DNA in situ. The system may comprise a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of any of the methods described herein related to an DNA. The system may also comprise the DNA.

[00446] The systems described herein may also include a cartridge comprising a sample collection unit and an assay assembly, wherein the sample collection unit is configured to collect a sample, and wherein the assay assembly comprises at least one reaction site containing a binding complex adapted to interact with said analyte, allowing the predetermined portion of sample to react with assay reagents contained within the assay assembly to yield a signal indicative of the presence of the analyte in the sample, and detecting the signal generated from the analyte.

[00447] In some embodiments, the system may include a biological sample, as described herein. The biological sample may be a cell, a tissue, a body fluid, or a combination thereof. In some cases, the cell may be a cancer cell. In some cases, the tissue may be formalin-fixed, paraffin embedded. In some other cases, the tissue may be fresh-frozen. The sample may be extracted from a subject, wherein the subject may be human. The biological sample may be provided in a tube, on a slide, in a container, in a box, in a vial, or a combination thereof. In some cases, the biological sample may be frozen, cryopreserved, stored at room temperature, stored at a refrigerated temperature, or a combination thereof. In some cases, the biological sample may comprise one or more targets, as described herein. The one or more targets may comprise a ribonucleic acid.

[00448] The systems described herein may include one or more nucleic acid molecules, as described herein. The one or more nucleic acid molecules may recognize the one or more target molecule. The one of more nucleic acid molecules may be stored in a variety of formats including but not limited to in solution, lyophilized, in solid form, flash-frozen, frozen, cryopreserved, or a combination thereof. The one or more nucleic acid molecules may be stored in a tube, a vial, a reservoir, a container, or a combination thereof. In some cases, the one or more nucleic acid molecules may be stored at about at least -80°C, at about at least -70°C, at about at least -60°C, at about at least -50°C, at about at least -40°C, at about at least -30°C, at about at least -20°C, at about at least -10°C, at about at least 0°C, at about at least 10°C, at about at least 20°C, or at about at least 30°C. In some embodiments, the one or more nucleic acid molecules may be stored at about at most -80°C, at about at most -70°C, at about at most -60°C, at about at most -

50°C, at about at most -40°C, at about at most -30°C, at about at most -20°C, at about at most -10°C, at about at most 0°C, at about at most 10°C, at about at most 20°C, or at about at most 30°C. The systems described herein may comprise more than one nucleic acid molecules, wherein each nucleic acid molecule binds to a different sequence. In some cases, the different sequence may be a different sequence within the same target molecules. In some cases, the different sequence may be a different sequence on a different target molecule. The systems described herein may comprise at least about 1, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, at least about 100, at least about 200, at least about 500, at least about 1000, at least about 2000, at least about 5000, at least about 10000, at least about 12000, at least about 20000, at least about 25000, at least about 50000, at least about 100000, or at least about 1000000 different nucleic acid molecules.

[00449] The systems described herein may include one or more primer nucleic acid sequences, as described herein. The one or primer nucleic acid sequences may recognize the index sequences, as described herein. The one of more primer nucleic acid sequences may be stored in a variety of formats including but not limited to in solution, lyophilized, in solid form, flash-frozen, frozen, cryopreserved, or a combination thereof. The one or more primer nucleic acid sequences may be stored in a tube, a vial, a reservoir, a container, or a combination thereof. In some cases, the one or more primer nucleic acid sequences may be stored at about at least -80°C, at about at least -70°C, at about at least -60°C, at about at least -50°C, at about at least -40°C, at about at least -30°C, at about at least -20°C, at about at least -10°C, at about at least 0°C, at about at least 10°C, at about at least 20°C, or at about at least 30°C. In some embodiments, the one or more primer nucleic acid sequences may be stored at about at most -80°C, at about at most -70°C, at about at most -60°C, at about at most -50°C, at about at most -40°C, at about at most -30°C, at about at most -20°C, at about at most -10°C, at about at most 0°C, at about at most 10°C, at about at most 20°C, or at about at most 30°C. The systems described herein may comprise more than one primer nucleic acid sequences, wherein each primer nucleic acid sequence binds to a different sequence. The systems described herein may comprise at least about 1, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, at least about 100, at least about 200, at least about 500, at least about 1000, at least about 2000, at least about 5000, at least about 10000, at least about 12000, at least about 20000, at least about 25000, at least about 50000, at least about 100000, or at least about 1000000 different primer nucleic acid sequences.

[00450] The systems described herein may include one or more index-binding sequences, as described herein. The one or more index-binding sequences may recognize the one or more index sequences, as described herein. The one of more index-binding sequences may be stored in a variety of formats including but not limited to in solution, lyophilized, in solid form, flash-frozen, frozen, cryopreserved, or a combination thereof. The one or more nucleic acid molecule may be stored in a tube, a vial, a reservoir, a container, or a combination thereof. In some cases, the one or more index-binding sequences may be stored at least about -80°C, at least about -70°C, at least about -60°C, at least about -50°C, at least about -40°C, at least about -30°C, at least about -20°C, at least about -10°C, at least about 0°C, at least about 10°C, at least about 20°C, or at least about 30°C. In some embodiments, the one or more index-binding sequences may be stored at most about -80°C, at most about -70°C, at most about -60°C, at most about -50°C, at most about -40°C, at most about -30°C, at most about -20°C, at most about -10°C, at most about 0°C, at most about 10°C, at most about 20°C, or at most about 30°C. The systems described herein may comprise more than one index-binding sequences, wherein each nucleic acid molecule binds to a different index sequence. The systems described herein may comprise at least about 1, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, at least about 100, at least about 200, at least about 500, at least about 1000, at least about 2000, at least about 5000, at least about 10000, at least about 12000, at least about 20000, at least about 25000, at least about 50000, at least about 100000, or at least about 1000000 different index-binding sequences.

[00451] The systems described herein may comprises a ligase, as described herein. The ligase may be a T4 ligase, a SplintR ligase, a T7 ligase, or a combination thereof. The ligase may be stored at a variety of temperatures including at least about -80°C, at least about -70°C, at least about -60°C, at least about -50°C, at least about -40°C, at least about -30°C, at least about -20°C, at least about -10°C, at least about 0°C, at least about 10°C, at least about 20°C, or at least about 30°C. In some embodiments, the ligase may be stored at most about -80°C, at most about -70°C, at most about -60°C, at most about -50°C, at most about -40°C, at most about -30°C, at most about -20°C, at most about -10°C, at most about 0°C, at most about 10°C, at most about 20°C, or at most about 30°C. The ligase may be stored in solution, lyophilized, frozen, flash-frozen, cryopreserved, or a combination thereof. In some cases, the ligase may be stored in combination with a glycerol solution.

[00452] The systems described herein may include a polymerizing enzyme, as described herein. The polymerizing enzyme may be a polymerase. The polymerase may be a phi29 polymerase. The polymerizing enzyme may be stored in a variety of formats including in solution,

lyophilized, frozen, flash-frozen, cryopreserved, or a combination thereof. In some cases, the polymerizing enzyme may be stored in combination with a glycerol solution. The polymerizing enzyme may be stored at a variety of temperatures including at least about -80°C, at least about -70°C, at least about -60°C, at least about -50°C, at least about -40°C, at least about -30°C, at least about -20°C, at least about -10°C, at least about 0°C, at least about 10°C, at least about 20°C, or at least about 30°C. In some embodiments, the polymerizing enzyme may be stored at most about -80°C, at most about -70°C, at most about -60°C, at most about -50°C, at most about -40°C, at most about -30°C, at most about -20°C, at most about -10°C, at most about 0°C, at most about 10°C, at most about 20°C, or at most about 30°C.

[00453] In some embodiments, the system may comprise a buffer, reagent, or combination thereof. The buffer and/or reagent may comprise a salt, a solvent, a chaotropic reagent, or a combination thereof. The salt may comprise any combination of one or more salts described herein, including but not limited to NaCl, $MgCl_2$, $CaCl_2$, or a combination thereof. The solvent may comprise any combination of solvents described herein, including a polar solvent. The chaotropic reagent may comprise any chaotropic reagent previously disclosed herein. The buffer and/or reagent may be formulated with a variety of components. The buffer and/or reagent may be used in one or more steps of any of the methods described herein.

[00454] In some embodiments, the system may include a substrate, as described herein. The substrate may be an interior surface of a flow cell. In some embodiments, the flow cell may comprise a heater, wherein the heater may be integrated with the flow cell. The flow cell shape may be similar to a microscope slide. In some embodiments, the flow cell may comprise a capillary flow cell. Representative flow cells may be found, for example, in International Application No. PCT/US2022/037831, the contents of which are hereby incorporated by reference in their entirety.

[00455] The systems described herein may comprise one or more computer processors. The one or more computer processors may comprise a computing processing unit, one or more graphic processing unit, one or more general-purpose processing unit, one or more field-programmable gate array, or a combination thereof. In some embodiments, the one or more computer processors may comprise one or more In some embodiments, the one or more computer processors may include a central processing unit. In some embodiments, the one or more computer processors may include a graphic processing unit. In some embodiments, the one or more computer processors may include a general-purpose processing unit. In some embodiments, the one or more computer processors may include a field-programmable gate array. In some instances, the one or more processors may be a single core or multi core processor, or a plurality of processors may be configured for parallel processing. Although the disclosure is described with

reference to a processor, other types of integrated circuits and logic devices are also applicable. The processor may have any suitable data operation capability. For example, the processor may perform 512 bit, 256 bit, 128 bit, 64 bit, 32 bit, or 16 bit data operations.

[00456] The systems described herein may comprise a memory or memory storage unit. The memory or memory storage unit may be a data storage unit (or data repository) for storing data. The memory or storage units may store files, such as drivers, libraries and saved programs. The memory or storage units may store user data, e.g. ., user preferences and user programs. The systems described herein may comprise an operating system configured to perform executable instructions.

[00457] The systems described herein may comprise a computer program. The computer program may include instructions for performing one or more of the methods described herein. The instructions may be stored in a memory location. The instructions may be directed to the one or more processors, for example a computer processing unity, which can program or otherwise configure the computer processing unit to implement the methods described herein.

i. Sequencing system

[00458] Disclosed herein, in some embodiments, are systems sequencing systems for performing nucleic acid sequencing *in situ* to identify a target molecule in a biological sample. The sequencing system may comprise a fluidic system, an imaging system, a flow cell, reagents, buffers, reservoirs or a combination thereof. The fluidic system, the imaging system, or a combination thereof may be part of the computer-implemented systems described herein.

[00459] In some embodiments, the systems described herein may include a fluidic system. In some embodiments, the fluidic system may be configured to provide fluid flow control capability for delivering reagents to the biological sample in a flow cell, as described herein. The fluidic system may comprise fluid flow controller. The fluid flow controller may be configured to control sequential and iterative delivery of a reagent to a flow cell, as described herein. The fluidic system may comprise a fluid dispensing system for dispensing reagents and/or other solutions into microplates, capillary flow cell devices and cartridges, microfluidic devices and cartridges. The fluid dispensing system may be programable. In some embodiments, the fluid dispensing system may be automated. Suitable automated, programmable fluid-dispensing systems are commercially available from a number of vendors, e.g. Beckman Coulter, Perkin Elmer, Tecan, Velocity 11, and many others. In a preferred aspect of the disclosed systems, the fluid-dispensing system further comprises a multichannel dispense head, e.g. a 4 channel, 8 channel, 16 channel, 96 channel, or 384 channel dispense head, for simultaneous delivery of programmable volumes of liquid (e.g. ranging from about 1 microliter to several milliliters) to multiple wells or locations on

a flow cell cartridge or microfluidic cartridge. In some embodiments, the fluidic system may comprise a flow cell, a pump, a valve, or a combination thereof. Representative flow cells may be found, for example, in PCT Patent Application No. PCT/US2022/037831, the contents of which are hereby incorporated by reference in their entirety.

[00460] In any of the methods described herein, the biological sample can be deposited onto a solid support (e.g., a flowcell). In some embodiments, the biological sample is deposited onto a flowcell having walls (e.g., top or first wall, and bottom or second wall) and a gap in-between, where the gap can be filled with a fluid, where the flowcell is positioned in a fluorescence optical imaging system. The biological sample has a thickness that may require using the imaging system to focus separately on the first and second surfaces of the flowcell, when using a traditional imaging system. For improved imaging of the sequencing reaction of the concatemers in the biological sample, the flowcell can be positioned in a high performance fluorescence imaging system, which comprises two or more tube lenses which are designed to provide optimal imaging performance for the first and second surfaces of the flowcell at two or more fluorescence wavelengths. In some embodiments, the high-performance imaging system further comprises a focusing mechanism configured to refocus the optical system between acquiring images of the first and second surfaces of the flowcell. In some embodiments, the high performance imaging system is configured to image two or more fields-of-view on at least one of the first flowcell surface or the second flowcell surface.

[00461] In any of the methods described herein, the solid support comprises a flowcell having a coating that promotes cell adhesion. In some embodiments, the flowcell comprises a support which can be a planar or non-planar support. The support can be solid or semi-solid. In some embodiments, the support can be porous, semi-porous or non-porous. The support can be made of any material such as glass, plastic or a polymer material. In some embodiments, the surface of the support can be coated with one or more compounds to produce a passivated layer on the support. In some embodiments, the passivated layer forms a porous or semi-porous layer. In some embodiments, the support is coated with a lysine compound, poly-lysine compound, arginine compound or an amino-terminated compound. The support can be coated with an unbranched compound, a branched compound, or a mixture of unbranched and branched compounds. In some embodiments, the support is coated with surface primers for capturing nucleic acids from the biological sample. Alternatively, the support lacks surface primers.

[00462] The sequencing system may include an imaging system. The imaging system may comprise an imaging module comprising one or more light sources, one or more optical components, one or more image sensors, or a combination thereof. The imaging system

may be configured to image and detect the binding of the disclosed nucleic acid binding compositions to target molecules tethered to a solid support or the interior of a flow cell. The disclosed compositions, reagents, and methods may be used for any of a variety of nucleic acid sequencing and analysis applications. Examples include, but are not limited to, DNA sequencing, RNA sequencing, whole genome sequencing, targeted sequencing, exome sequencing, genotyping, and the like.

[00463] The sequencing system may also include computer control systems that are programmed to implement methods of the disclosure. The computer system is programmed or otherwise configured to implement methods of the disclosure including nucleic acid sequencing methods, interpretation of nucleic acid sequencing data and analysis of cellular nucleic acids, such as RNA (e.g., mRNA), and characterization of cells from sequencing data. The computer system can be an electronic device of a user or a computer system that is remotely located with respect to the electronic device. The electronic device can be a mobile electronic device.

[00464] In any of the methods described herein, any combination of the steps for conducting *in situ* reiterative short read sequencing can be performed in an automated mode using the fluid dispensing system, including cell seeding, cell fixation, cell permeabilization, reverse transcription reactions, padlock probe hybridization, padlock probe ligation reaction, rolling circle amplification, and sequencing.

[00465] The present disclosure provides apparatus and methods for growing/culturing a biological sample on a flowcell and conducting nucleic acid workflows of the cultured biological sample on the flowcell.

[00466] In some embodiments, the biological sample is deposited on the flowcell. The flowcell can be coated with a reagent that promotes cell adhesion 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.

[00467] In some embodiments, the automated fluid dispensing system can be used to deliver simple or complex cell culture media to the biological sample on the flowcell. In some embodiments, the biological sample can be cultured/expanded on the flowcell for 2-10 generations or more. In some embodiments, the biological sample can be expanded to confluence or non-confluence.

[00468] In some embodiments, the automated fluid dispensing system can be used to deliver fixation reagents to the expanded biological sample on the flowcell, and the biological sample can be incubated under conditions suitable for cell fixation.

[00469] In some embodiments, the automated fluid dispensing system can be used to deliver permeabilization reagents to the fixed biological sample on the flowcell, and the biological sample can be incubated under conditions suitable for cell permeabilization.

[00470] In some embodiments, the automated fluid dispensing system can be used to deliver reagents for conducting padlock probe hybridization, circularization and ligation under a condition suitable for generating a plurality of covalently closed circular padlock probes inside the biological sample.

[00471] 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 inside the biological sample.

[00472] 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 inside the biological sample. 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 mm^2 and the cycle time for scanning large area (> 10 mm^2) can be less than 5 minutes.

[00473] 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 inside the biological sample.

[00474] 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 another plurality of sequencing read products inside the biological sample. 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 mm^2 and the cycle time for scanning large area (> 10 mm^2) can be less than 5 minutes.

ii. Computing systems

[00475] Referring to **FIG. 13**, a block diagram is shown depicting an exemplary machine that includes a computer system **1300** (e.g., a processing or computing system) within which a set of instructions can execute for causing a device to perform or execute any one or more of the aspects and/or methodologies for static code scheduling of the present disclosure. The components in **FIG. 13** are examples only and do not limit the scope of use or functionality of

any hardware, software, embedded logic component, or a combination of two or more such components implementing particular embodiments.

[00476] Computer system **1300** may include one or more processors **1301**, a memory **1303**, and a storage **1308** that communicate with each other, and with other components, via a bus **140**. The bus **140** may also link a display **1302**, one or more input devices **1303** (which may, for example, include a keypad, a keyboard, a mouse, a stylus, etc.), one or more output devices **1304**, one or more storage devices **1305**, and various tangible storage media **1306**. All of these elements may interface directly or via one or more interfaces or adaptors to the bus **140**. For instance, the various tangible storage media **1306** can interface with the bus **140** via storage medium interface **1326**. Computer system **1300** may have any suitable physical form, including but not limited to one or more integrated circuits (ICs), printed circuit boards (PCBs), mobile handheld devices (such as mobile telephones or PDAs), laptop or notebook computers, distributed computer systems, computing grids, or servers.

[00477] Computer system **1300** includes one or more processor(s) **1301** (e.g., central processing units (CPUs), general purpose graphics processing units (GPGPUs), or quantum processing units (QPUs)) that carry out functions. Processor(s) **1301** optionally contains a cache memory unit **1302** for temporary local storage of instructions, data, or computer addresses. Processor(s) **1301** are configured to assist in execution of computer readable instructions. Computer system **1300** may provide functionality for the components depicted in **FIG. 13** as a result of the processor(s) **1301** executing non-transitory, processor-executable instructions embodied in one or more tangible computer-readable storage media, such as memory **1303**, storage **1308**, storage devices **1305**, and/or storage medium **1306**. The computer-readable media may store software that implements particular embodiments, and processor(s) **1301** may execute the software. Memory **1303** may read the software from one or more other computer-readable media (such as mass storage device(s) **1305**, **1306**) or from one or more other sources through a suitable interface, such as network interface **1320**. The software may cause processor(s) **1301** to carry out one or more processes or one or more steps of one or more processes described or illustrated herein. Carrying out such processes or steps may include defining data structures stored in memory **1303** and modifying the data structures as directed by the software.

[00478] The memory **1303** may include various components (e.g., machine readable media) including, but not limited to, a random access memory component (e.g., RAM **1304**) (e.g., static RAM (SRAM), dynamic RAM (DRAM), ferroelectric random access memory (FRAM), phase-change random access memory (PRAM), etc.), a read-only memory component (e.g., ROM **1305**), and any combinations thereof. ROM **1305** may act to communicate data and instructions unidirectionally to processor(s) **1301**, and RAM **1304** may act to communicate data

and instructions bidirectionally with processor(s) **1301**. ROM **1305** and RAM **1304** may include any suitable tangible computer-readable media described below. In one example, a basic input/output system **1306** (BIOS), including basic routines that help to transfer information between elements within computer system **1300**, such as during start-up, may be stored in the memory **1303**.

[00479] Fixed storage **1308** is connected bidirectionally to processor(s) **1301**, optionally through storage control unit **1307**. Fixed storage **1308** provides additional data storage capacity and may also include any suitable tangible computer-readable media described herein. Storage **1308** may be used to store operating system **1309**, executable(s) **13130**, data **131311**, applications **1321312** (application programs), and the like. Storage **1308** can also include an optical disk drive, a solid-state memory device (e.g., flash-based systems), or a combination of any of the above. Information in storage **1308** may, in appropriate cases, be incorporated as virtual memory in memory **1303**.

[00480] In one example, storage device(s) **1305** may be removably interfaced with computer system **1300** (e.g., via an external port connector (not shown)) via a storage device interface **1325**. Particularly, storage device(s) **1305** and an associated machine-readable medium may provide non-volatile and/or volatile storage of machine-readable instructions, data structures, program modules, and/or other data for the computer system **1300**. In one example, software may reside, completely or partially, within a machine-readable medium on storage device(s) **1305**. In another example, software may reside, completely or partially, within processor(s) **1301**.

[00481] Bus **140** connects a wide variety of subsystems. Herein, reference to a bus may encompass one or more digital signal lines serving a common function, where appropriate. Bus **140** may be any of several types of bus structures including, but not limited to, a memory bus, a memory controller, a peripheral bus, a local bus, and any combinations thereof, using any of a variety of bus architectures. As an example, and not by way of limitation, such architectures include an Industry Standard Architecture (ISA) bus, an Enhanced ISA (EISA) bus, a Micro Channel Architecture (MCA) bus, a Video Electronics Standards Association local bus (VLB), a Peripheral Component Interconnect (PCI) bus, a PCI-Express (PCI-X) bus, an Accelerated Graphics Port (AGP) bus, HyperTransport (HTX) bus, serial advanced technology attachment (SATA) bus, and any combinations thereof.

[00482] Computer system **1300** may also include an input device **1303**. In one example, a user of computer system **1300** may enter commands and/or other information into computer system **1300** via input device(s) **1303**. Examples of an input device(s) **1303** include, but are not limited to, an alpha-numeric input device (e.g., a keyboard), a pointing device (e.g., a mouse or touchpad), a touchpad, a touch screen, a multi-touch screen, a joystick, a stylus, a

gamepad, an audio input device (e.g., a microphone, a voice response system, etc.), an optical scanner, a video or still image capture device (e.g., a camera), and any combinations thereof. In some embodiments, the input device is a Kinect, Leap Motion, or the like. Input device(s) **1303** may be interfaced to bus **140** via any of a variety of input interfaces **1323** (e.g., input interface **1323**) including, but not limited to, serial, parallel, game port, USB, FIREWIRE, THUNDERBOLT, or any combination of the above.

[00483] In particular embodiments, when computer system **1300** is connected to network **130**, computer system **1300** may communicate with other devices, specifically mobile devices and enterprise systems, distributed computing systems, cloud storage systems, cloud computing systems, and the like, connected to network **130**. Communications to and from computer system **1300** may be sent through network interface **1320**. For example, network interface **1320** may receive incoming communications (such as requests or responses from other devices) in the form of one or more packets (such as Internet Protocol (IP) packets) from network **130**, and computer system **1300** may store the incoming communications in memory **1303** for processing. Computer system **1300** may similarly store outgoing communications (such as requests or responses to other devices) in the form of one or more packets in memory **1303** and communicated to network **130** from network interface **1320**. Processor(s) **1301** may access these communication packets stored in memory **1303** for processing.

[00484] Examples of the network interface **1320** include, but are not limited to, a network interface card, a modem, and any combination thereof. Examples of a network **130** or network segment **130** include, but are not limited to, a distributed computing system, a cloud computing system, a wide area network (WAN) (e.g., the Internet, an enterprise network), a local area network (LAN) (e.g., a network associated with an office, a building, a campus or other relatively small geographic space), a telephone network, a direct connection between two computing devices, a peer-to-peer network, and any combinations thereof. A network, such as network **130**, may employ a wired and/or a wireless mode of communication. In general, any network topology may be used.

[00485] Information and data can be displayed through a display **1302**. Examples of a display **1302** include, but are not limited to, a cathode ray tube (CRT), a liquid crystal display (LCD), a thin film transistor liquid crystal display (TFT-LCD), an organic liquid crystal display (OLED) such as a passive-matrix OLED (PMOLED) or active-matrix OLED (AMOLED) display, a plasma display, and any combinations thereof. The display **1302** can interface to the processor(s) **1301**, memory **1303**, and fixed storage **1308**, as well as other devices, such as input device(s) **1303**, via the bus **140**. The display **1302** is linked to the bus **140** via a video interface **1322**, and transport of data between the display **1302** and the bus **140** can be controlled via the graphics

control **1321**. In some embodiments, the display is a video projector. In some embodiments, the display is a head-mounted display (HMD) such as a VR headset. In further embodiments, suitable VR headsets include, by way of non-limiting examples, HTC Vive, Oculus Rift, Samsung Gear VR, Microsoft HoloLens, Razer OSVR, FOVE VR, Zeiss VR One, Avegant Glyph, Freefly VR headset, and the like. In still further embodiments, the display is a combination of devices such as those disclosed herein.

[00486] In addition to a display **1302**, computer system **1300** may include one or more other peripheral output devices **1304** including, but not limited to, an audio speaker, a printer, a storage device, and any combinations thereof. Such peripheral output devices may be connected to the bus **140** via an output interface **1324**. Examples of an output interface **1324** include, but are not limited to, a serial port, a parallel connection, a USB port, a FIREWIRE port, a THUNDERBOLT port, and any combinations thereof.

[00487] In addition, or as an alternative, computer system **1300** may provide functionality as a result of logic hardwired or otherwise embodied in a circuit, which may operate in place of or together with software to execute one or more processes or one or more steps of one or more processes described or illustrated herein. Reference to software in this disclosure may encompass logic, and reference to logic may encompass software. Moreover, reference to a computer-readable medium may encompass a circuit (such as an IC) storing software for execution, a circuit embodying logic for execution, or both, where appropriate. The present disclosure encompasses any suitable combination of hardware, software, or both.

[00488] Those of skill in the art will appreciate that the various illustrative logical blocks, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented as electronic hardware, computer software, or combinations of both. To clearly illustrate this interchangeability of hardware and software, various illustrative components, blocks, modules, circuits, and steps have been described above generally in terms of their functionality.

[00489] The various illustrative logical blocks, modules, and circuits described in connection with the embodiments disclosed herein may be implemented or performed with a general purpose processor, a digital signal processor (DSP), an application specific integrated circuit (ASIC), a field programmable gate array (FPGA) or other programmable logic device, discrete gate or transistor logic, discrete hardware components, or any combination thereof designed to perform the functions described herein. A general purpose processor may be a microprocessor, but in the alternative, the processor may be any conventional processor, controller, microcontroller, or state machine. A processor may also be implemented as a combination of computing devices, e.g., a combination of a DSP and a microprocessor, a plurality

of microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration.

[00490] The steps of a method or algorithm described in connection with the embodiments disclosed herein may be embodied directly in hardware, in a software module executed by one or more processor(s), or in a combination of the two. A software module may reside in RAM memory, flash memory, ROM memory, EPROM memory, EEPROM memory, registers, hard disk, a removable disk, a CD-ROM, or any other form of storage medium known in the art. An exemplary storage medium is coupled to the processor such the processor can read information from, and write information to, the storage medium. In the alternative, the storage medium may be integral to the processor. The processor and the storage medium may reside in an ASIC. The ASIC may reside in a user terminal. In the alternative, the processor and the storage medium may reside as discrete components in a user terminal.

[00491] In accordance with the description herein, suitable computing devices include, by way of non-limiting examples, server computers, desktop computers, laptop computers, notebook computers, sub-notebook computers, netbook computers, notepad computers, set-top computers, media streaming devices, handheld computers, Internet appliances, mobile smartphones, tablet computers, personal digital assistants, video game consoles, and vehicles. Those of skill in the art will also recognize that select televisions, video players, and digital music players with optional computer network connectivity are suitable for use in the system described herein. Suitable tablet computers, in various embodiments, include those with booklet, slate, and convertible configurations, known to those of skill in the art.

[00492] In some embodiments, the computing device includes an operating system configured to perform executable instructions. The operating system is, for example, software, including programs and data, which manages the device's hardware and provides services for execution of applications. Those of skill in the art will recognize that suitable server operating systems include, by way of non-limiting examples, FreeBSD, OpenBSD, NetBSD®, Linux, Apple® Mac OS X Server®, Oracle® Solaris®, Windows Server®, and Novell® NetWare®. Those of skill in the art will recognize that suitable personal computer operating systems include, by way of non-limiting examples, Microsoft® Windows®, Apple® Mac OS X®, UNIX®, and UNIX-like operating systems such as GNU/Linux®. In some embodiments, the operating system is provided by cloud computing. Those of skill in the art will also recognize that suitable mobile smartphone operating systems include, by way of non-limiting examples, Nokia® Symbian® OS, Apple® iOS®, Research In Motion® BlackBerry OS®, Google® Android®, Microsoft® Windows Phone® OS, Microsoft® Windows Mobile® OS, Linux®, and Palm® WebOS®. Those of skill in the art will also recognize that suitable media streaming device operating systems include, by way of non-limiting

examples, Apple TV®, Roku®, Boxee®, Google TV®, Google Chromecast®, Amazon Fire®, and Samsung® HomeSync®. Those of skill in the art will also recognize that suitable video game console operating systems include, by way of non-limiting examples, Sony® PS3®, Sony® PS4®, Microsoft® Xbox 360®, Microsoft Xbox One, Nintendo® Wii®, Nintendo® Wii U®, and Ouya®.

Non-transitory computer readable storage medium

[00493] In some embodiments, the platforms, systems, media, and methods disclosed herein include one or more non-transitory computer readable storage media encoded with a program including instructions executable by the operating system of an optionally networked computing device. In further embodiments, a computer readable storage medium is a tangible component of a computing device. In still further embodiments, a computer readable storage medium is optionally removable from a computing device. In some embodiments, a computer readable storage medium includes, by way of non-limiting examples, CD-ROMs, DVDs, flash memory devices, solid state memory, magnetic disk drives, magnetic tape drives, optical disk drives, distributed computing systems including cloud computing systems and services, and the like. In some cases, the program and instructions are permanently, substantially permanently, semi-permanently, or non-transitorily encoded on the media.

Computer program

[00494] In some embodiments, the platforms, systems, media, and methods disclosed herein include at least one computer program, or use of the same. A computer program includes a sequence of instructions, executable by one or more processor(s) of the computing device's CPU, written to perform a specified task. Computer readable instructions may be implemented as program modules, such as functions, objects, Application Programming Interfaces (APIs), computing data structures, and the like, which perform particular tasks or implement particular abstract data types. In light of the disclosure provided herein, those of skill in the art will recognize that a computer program may be written in various versions of various languages.

[00495] The functionality of the computer readable instructions may be combined or distributed as desired in various environments. In some embodiments, a computer program comprises one sequence of instructions. In some embodiments, a computer program comprises a plurality of sequences of instructions. In some embodiments, a computer program is provided from one location. In other embodiments, a computer program is provided from a plurality of locations. In various embodiments, a computer program includes one or more software modules. In various embodiments, a computer program includes, in part or in whole, one or more web

applications, one or more mobile applications, one or more standalone applications, one or more web browser plug-ins, extensions, add-ins, or add-ons, or combinations thereof.

[00496] In some embodiments, the computer programs described herein may be used to perform at least one function. The computer programs described herein may perform functions related to storing data, receiving data, analyzing data, exporting data, or a combination thereof. In some embodiments, the computer programs described herein may perform functions related to applying selection criteria, including in silico selection criteria, functional selection criteria, or a combination thereof. In some embodiments, the computer programs may receive sequence information, including sequence information for nucleic acid segments. The sequence information may be configured as an array, a table, a list, or combination thereof. The sequence information may be formatted in a variety of ways, including, but not limited to a .txt file, a FASTA file, an .xls file, or a combination thereof. The computer programs described herein may apply selection criterion or selection criteria to a set of nucleic acid segments. The computer programs may sort the nucleic acid segments, determine or compute characteristics of the nucleic acid segments, perform calculations, reorder the nucleic acid segments, or a combination thereof. In some embodiments, the computer programs described herein may store information related to the nucleic acid segments. In some embodiments, the computer program may use information stored related to the nucleic acid segments to apply selection criteria to the nucleic acid segments. In certain embodiments, the computer program may receive information and/or data related to nucleic acid segments, selection criteria, or a combination thereof. In some embodiments, the computer programs may perform functions related to analyzing data from functional assays, including, but not limited to functional assays described herein. In some embodiments, analyzing data from functional assays may comprise image analysis, image quantification, intensity quantification, feature identification, or a combination thereof. The computer programs described herein may also export information. In some embodiments, the exported information may comprise images, files, data tables, documents, folders, or a combination thereof.

Web application

[00497] In some embodiments, a computer program includes a web application. In light of the disclosure provided herein, those of skill in the art will recognize that a web application, in various embodiments, utilizes one or more software frameworks and one or more database systems. In some embodiments, a web application is created upon a software framework such as Microsoft® .NET or Ruby on Rails (RoR). In some embodiments, a web application utilizes one or more database systems including, by way of non-limiting examples, relational, non-relational, object oriented, associative, XML, and document oriented database systems. In further embodiments, suitable relational database systems include, by way of non-limiting examples,

Microsoft® SQL Server, mySQL™, and Oracle®. Those of skill in the art will also recognize that a web application, in various embodiments, is written in one or more versions of one or more languages. A web application may be written in one or more markup languages, presentation definition languages, client-side scripting languages, server-side coding languages, database query languages, or combinations thereof. In some embodiments, a web application is written to some extent in a markup language such as Hypertext Markup Language (HTML), Extensible Hypertext Markup Language (XHTML), or eXtensible Markup Language (XML). In some embodiments, a web application is written to some extent in a presentation definition language such as Cascading Style Sheets (CSS). In some embodiments, a web application is written to some extent in a client-side scripting language such as Asynchronous JavaScript and XML (AJAX), Flash® ActionScript, JavaScript, or Silverlight®. In some embodiments, a web application is written to some extent in a server-side coding language such as Active Server Pages (ASP), ColdFusion®, Perl, Java™, JavaServer Pages (JSP), Hypertext Preprocessor (PHP), Python™, Ruby, Tcl, Smalltalk, WebDNA®, or Groovy. In some embodiments, a web application is written to some extent in a database query language such as Structured Query Language (SQL). In some embodiments, a web application integrates enterprise server products such as IBM® Lotus Domino®. In some embodiments, a web application includes a media player element. In various further embodiments, a media player element utilizes one or more of many suitable multimedia technologies including, by way of non-limiting examples, Adobe® Flash®, HTML 5, Apple® QuickTime®, Microsoft® Silverlight®, Java™, and Unity®.

[00498] Referring to **FIG. 14**, in a particular embodiment, an application provision system comprises one or more databases **200** accessed by a relational database management system (RDBMS) **1410**. Suitable RDBMSs include Firebird, MySQL, PostgreSQL, SQLite, Oracle Database, Microsoft SQL Server, IBM DB2, IBM Informix, SAP Sybase, Teradata, and the like. In this embodiment, the application provision system further comprises one or more application severs **1420** (such as Java servers, .NET servers, PHP servers, and the like) and one or more web servers **1430** (such as Apache, IIS, GWS and the like). The web server(s) optionally expose one or more web services via app application programming interfaces (APIs) **1440**. Via a network, such as the Internet, the system provides browser-based and/or mobile native user interfaces.

[00499] Referring to **FIG. 15**, in a particular embodiment, an application provision system alternatively has a distributed, cloud-based architecture **1500** and comprises elastically load balanced, auto-scaling web server resources **1510** and application server resources **1520** as well synchronously replicated databases **1530**.

Mobile application

[00500] In some embodiments, a computer program includes a mobile application provided to a mobile computing device. In some embodiments, the mobile application is provided to a mobile computing device at the time it is manufactured. In other embodiments, the mobile application is provided to a mobile computing device via the computer network described herein.

[00501] In view of the disclosure provided herein, a mobile application is created by techniques known to those of skill in the art using hardware, languages, and development environments known to the art. Those of skill in the art will recognize that mobile applications are written in several languages. Suitable programming languages include, by way of non-limiting examples, C, C++, C#, Objective-C, Java™, JavaScript, Pascal, Object Pascal, Python™, Ruby, VB.NET, WML, and XHTML/HTML with or without CSS, or combinations thereof.

[00502] Suitable mobile application development environments are available from several sources. Commercially available development environments include, by way of non-limiting examples, AirplaySDK, alcheMo, Appcelerator®, Celsius, Bedrock, Flash Lite, .NET Compact Framework, Rhomobile, and WorkLight Mobile Platform. Other development environments are available without cost including, by way of non-limiting examples, Lazarus, MobiFlex, MoSync, and Phonegap. Also, mobile device manufacturers distribute software developer kits including, by way of non-limiting examples, iPhone and iPad (iOS) SDK, Android™ SDK, BlackBerry® SDK, BREW SDK, Palm® OS SDK, Symbian SDK, webOS SDK, and Windows® Mobile SDK.

[00503] Those of skill in the art will recognize that several commercial forums are available for distribution of mobile applications including, by way of non-limiting examples, Apple® App Store, Google® Play, Chrome WebStore, BlackBerry® App World, App Store for Palm devices, App Catalog for webOS, Windows® Marketplace for Mobile, Ovi Store for Nokia® devices, Samsung® Apps, and Nintendo® DSi Shop.

Standalone application

[00504] In some embodiments, a computer program includes a standalone application, which is a program that is run as an independent computer process, not an add-on to an existing process, e.g., not a plug-in. Those of skill in the art will recognize that standalone applications are often compiled. A compiler is a computer program(s) that transforms source code written in a programming language into binary object code such as assembly language or machine code. Suitable compiled programming languages include, by way of non-limiting examples, C, C++, Objective-C, COBOL, Delphi, Eiffel, Java™, Lisp, Python™, Visual Basic, and VB .NET, or combinations thereof. Compilation is often performed, at least in part, to create an executable

program. In some embodiments, a computer program includes one or more executable complied applications.

Web browser plug-in

[00505] In some embodiments, the computer program includes a web browser plug-in (e.g., extension, etc.). In computing, a plug-in is one or more software components that add specific functionality to a larger software application. Makers of software applications support plug-ins to enable third-party developers to create abilities which extend an application, to support easily adding new features, and to reduce the size of an application. When supported, plug-ins enable customizing the functionality of a software application. For example, plug-ins are commonly used in web browsers to play video, generate interactivity, scan for viruses, and display particular file types. Those of skill in the art will be familiar with several web browser plug-ins including, Adobe® Flash® Player, Microsoft® Silverlight®, and Apple® QuickTime®. In some embodiments, the toolbar comprises one or more web browser extensions, add-ins, or add-ons. In some embodiments, the toolbar comprises one or more explorer bars, tool bands, or desk bands.

[00506] In view of the disclosure provided herein, those of skill in the art will recognize that several plug-in frameworks are available that enable development of plug-ins in various programming languages, including, by way of non-limiting examples, C++, Delphi, Java™, PHP, Python™, and VB .NET, or combinations thereof.

[00507] Web browsers (also called Internet browsers) are software applications, designed for use with network-connected computing devices, for retrieving, presenting, and traversing information resources on the World Wide Web. Suitable web browsers include, by way of non-limiting examples, Microsoft® Internet Explorer®, Mozilla® Firefox®, Google® Chrome, Apple® Safari®, Opera Software® Opera®, and KDE Konqueror. In some embodiments, the web browser is a mobile web browser. Mobile web browsers (also called microbrowsers, mini-browsers, and wireless browsers) are designed for use on mobile computing devices including, by way of non-limiting examples, handheld computers, tablet computers, netbook computers, subnotebook computers, smartphones, music players, personal digital assistants (PDAs), and handheld video game systems. Suitable mobile web browsers include, by way of non-limiting examples, Google® Android® browser, RIM BlackBerry® Browser, Apple® Safari®, Palm® Blazer, Palm® WebOS® Browser, Mozilla® Firefox® for mobile, Microsoft® Internet Explorer® Mobile, Amazon® Kindle® Basic Web, Nokia® Browser, Opera Software® Opera® Mobile, and Sony® PSP™ browser.

Software modules

[00508] In some embodiments, the platforms, systems, media, and methods disclosed herein include software, server, and/or database modules, or use of the same. In view of the disclosure provided herein, software modules are created by techniques known to those of skill in the art using machines, software, and languages known to the art. The software modules disclosed herein are implemented in a multitude of ways. In various embodiments, a software module comprises a file, a section of code, a programming object, a programming structure, a distributed computing resource, a cloud computing resource, or combinations thereof. In further various embodiments, a software module comprises a plurality of files, a plurality of sections of code, a plurality of programming objects, a plurality of programming structures, a plurality of distributed computing resources, a plurality of cloud computing resources, or combinations thereof. In various embodiments, the one or more software modules comprise, by way of non-limiting examples, a web application, a mobile application, a standalone application, and a distributed or cloud computing application. In some embodiments, software modules are in one computer program or application. In other embodiments, software modules are in more than one computer program or application. In some embodiments, software modules are hosted on one machine. In other embodiments, software modules are hosted on more than one machine. In further embodiments, software modules are hosted on a distributed computing platform such as a cloud computing platform. In some embodiments, software modules are hosted on one or more machines in one location. In other embodiments, software modules are hosted on one or more machines in more than one location.

Databases

In some embodiments, the platforms, systems, media, and methods disclosed herein include one or more databases, or use of the same. In view of the disclosure provided herein, those of skill in the art will recognize that many databases are suitable for storage and retrieval of nucleic acid segment sequences or analysis thereof information. In various embodiments, suitable databases include, by way of non-limiting examples, relational databases, non-relational databases, object oriented databases, object databases, entity-relationship model databases, associative databases, XML databases, document oriented databases, and graph databases. Further non-limiting examples include SQL, PostgreSQL, MySQL, Oracle, DB2, Sybase, and MongoDB. In some embodiments, a database is Internet-based. In further embodiments, a database is web-based. In still further embodiments, a database is cloud computing-based. In a particular embodiment, a database is a distributed database. In other embodiments, a database is based on one or more local computer storage devices.

KITS

[00509] Provided here are compositions related to detecting a target in situ. In some cases, the target may be a ribonucleic acid. In some embodiments, the compositions are included in a kit with instructions for identifying the target *in situ* based on the methods described herein. In some cases, the kit may include a nucleic acid molecule, a detectable nucleotide conjugate, instructions for use, or a combination thereof.

[00510] An aspect of the present disclosure provides a kit for identifying a target (e.g. an RNA or a DNA) in situ. The kit may comprise (a) a nucleic acid molecule comprising (i) an index sequence, and (ii) a target-binding sequence, wherein the target-binding sequence may be complementary to at least a portion of the ribonucleic acid sequence; (b) a nucleotide conjugate (e.g. a detectable nucleotide conjugate) comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in the primed index sequence, wherein the nucleotide conjugate (e.g. detectable nucleotide conjugate) is configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (c) instructions for identifying the ribonucleic acid in situ, wherein the instructions comprise instructions for performing any one of the methods disclosed herein.

[00511] Another aspect of the present disclosure provides a kit for identifying an RNA in situ. The kit may comprise (a) a probe comprising (i) an index sequence, and (ii) a RNA-binding sequence, wherein the RNA-binding sequence recognizes to at least a portion of the target; (b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein the nucleotide conjugate may be configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (c) instructions for identifying the RNA in situ, wherein the instructions comprise instructions for performing any one of the methods disclosed herein.

[00512] Another aspect of the present disclosure provides a kit for identifying an DNA in situ. The kit may comprise (a) a probe comprising (i) an index sequence, and (ii) a DNA-binding sequence, wherein the DNA-binding sequence recognizes to at least a portion of the target; (b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein the nucleotide conjugate may be configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (c) instructions for identifying the

DNA in situ, wherein the instructions comprise instructions for performing any one of the methods disclosed herein.

[00513] Another aspect of the present disclosure provides a kit for identifying a polypeptide in situ. The kit may comprise (a) a polypeptide-binding probe comprising (i) a polypeptide binding moiety, and (ii) an oligonucleotide coupled thereto; (b) a probe comprising (i) an index sequence, and (ii) an oligonucleotide-binding probe, wherein said oligonucleotide-binding probe recognizes to at least a portion of said oligonucleotide; (c) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein the nucleotide conjugate may be configured to form a binding complex between the one or more nucleotide moieties and the nucleotide of the primed index sequence without incorporating the one or more nucleotide moieties into the primer nucleic acid sequence; and (d) instructions for identifying the polypeptide in situ, wherein the instructions comprise instructions for performing any one of the methods disclosed herein.

[00514] In some embodiments, the kit may comprise one or more nucleic acid molecules, as described herein. In some cases, the one or more nucleic acid molecules may comprise an index sequence, as described herein. In some cases, the one or more nucleic acid molecules may comprise a target-binding sequence, as described herein. In some cases, the target-binding sequence may be complementary to at least a portion of a target nucleic acid. The target may be a ribonucleic acid. The one or more nucleic acid molecules may be in a variety of formats, as described above. The one or more nucleic acid molecules may be stored at a variety of temperatures, as described above. The kits disclosed herein may comprise at least about 1, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, at least about 100, at least about 200, at least about 500, at least about 1000, at least about 2000, at least about 5000, at least about 10000, at least about 12000, at least about 20000, at least about 25000, at least about 50000, at least about 100000, or at least about 1000000 nucleic acid molecules. In some embodiments, the nucleic acid molecules may comprise the same sequence. In some embodiments, the nucleic acid molecules may comprise different sequences. The kits described herein may comprise nucleic acid molecules formulated individually. In some embodiments, the kits described herein may comprise at least a subset of the nucleic acid molecules formulated in one container. In some embodiments, the kits described herein may comprise all nucleic acid molecules formulated in one container.

[00515] The kits described herein may comprise a detectable nucleotide conjugate, as described herein. The nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise

a detectable label, as described herein. The nucleotide conjugate (e.g. detectable nucleotide conjugate)may comprise one or more nucleotide moieties, as described herein. The one or more nucleotide moieties may be complementary to a nucleotide in a primed index sequence, as described herein.

[00516] In some embodiments, the kit may comprise one or more primer nucleic acid sequences, as described herein. The one or more primer nucleic acid sequences may be in a variety of formats, as described above. The one or more primer nucleic acid sequences may be stored at a variety of temperatures, as described above. The kits disclosed herein may comprise at least about 1, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, at least about 100, at least about 200, at least about 500, at least about 1000, at least about 2000, at least about 5000, at least about 10000, at least about 12000, at least about 20000, at least about 25000, at least about 50000, at least about 100000, or at least about 1000000 primer nucleic acid sequences. In some embodiments, the primer nucleic acid sequences may comprise the same sequence. In some embodiments, the primer nucleic acid sequences may comprise different sequences. The kits described herein may comprise primer nucleic acid sequences formulated individually. In some embodiments, the kits described herein may comprise at least a subset of the primer nucleic acid sequences formulated in one container. In some embodiments, the kits described herein may comprise all primer nucleic acid sequences formulated in one container.

[00517] In some embodiments, the kit may comprise one or more index-binding molecules, as described herein. The one or more index-binding molecules may be a nucleic acid. In some cases, the one or more index-binding molecule may be a padlock probe. In some embodiments, the one or more index-binding molecule may be configured to bind to an index sequence, as described herein. The one or more index-binding molecules may be in a variety of formats, as described above. The one or more index-binding molecules may be stored at a variety of temperatures, as described above. The kits disclosed herein may comprise at least about 1, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, at least about 100, at least about 200, at least about 500, at least about 1000, at least about 2000, at least about 5000, at least about 10000, at least about 12000, at least about 20000, at least about 25000, at least about 50000, at least about 100000, or at least about 1000000 index-binding molecules. In some embodiments, the index-binding molecules may comprise the same sequence. In some

embodiments, the index-binding molecules may comprise different sequences. The kits described herein may comprise index-binding molecules formulated individually. In some embodiments, the kits described herein may comprise at least a subset of the index-binding molecules formulated in one container. In some embodiments, the kits described herein may comprise all index-binding molecules formulated in one container.

[00518] In some embodiments, the kit may comprise a ligase, as described herein. The ligase may be in a variety of formats, as described herein. The ligase may be stored in a variety of temperatures, as described above. In some embodiments, the kit may include a polymerizing enzyme. The polymerizing enzyme may be in a variety of formats, as described herein. The polymerizing enzyme may be stored in a variety of temperatures, as described above.

[00519] In some embodiments, the kit may include a buffer. The buffer may comprise a variety of components. In some embodiments, the buffer may comprise one or more of any one of the chemicals disclosed herein. The buffer may comprise a salt. In some cases, the salt may comprise NaCl, MgCl2, or CaCl2, or a combination there. In some cases, the buffer may comprise a solvent. The solvent may be polar. In some embodiments, the buffer may comprise a chaotropic reagent, as described herein. In some embodiments, the kit may include a reagent. The reagent may comprise a variety of components. In some embodiments, the reagent may comprise one or more of any one of the chemicals disclosed herein. The reagent may comprise a salt. In some cases, the salt may comprise NaCl, MgCl2, or CaCl2, or a combination there. In some cases, the reagent may comprise a solvent. The solvent may be polar. In some embodiments, the reagent may comprise a chaotropic reagent, as described herein. The buffer and/or reagent may be contained within a vial, a tube, a container, a box, a bottle, or a combination thereof.

[00520] In some embodiments, the kit may include one or more flow cells, as described herein.

EXAMPLES

[00521] The following examples are included for illustrative purposes only and are not intended to limit the scope of the disclosure.

[00522] While preferred embodiments of the present disclosure have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the disclosure. It should be understood that various alternatives to the embodiments of the disclosure described herein may be employed in practicing the disclosure. It is intended that the following claims define the scope of the disclosure

and that methods and structures within the scope of these claims and their equivalents be covered thereby.

Example 1. Preparation of Binding complex

[00523] In this example, one type of multi-armed substrate, as shown in FIG. 5A were made by reacting propargylamine dNTPs with Biotin-PEG-NHS. This aqueous reaction was driven to completion and purified; resulting in a pure Biotin-PEG-dNTP species. In separate reactions, several different PEG lengths were used, varying from 1K to 20K. The Biotin-PEG-dNTP species were mixed with either freshly prepared or commercially sourced dye-labeled streptavidin using a Dye:SA ratio of 3-5:1. Mixing of Biotin-PEG-dNTP with dye-labeled streptavidin was done in the presence of excess biotin-PEG-dNTP to ensure saturation of the biotin binding sites on each streptavidin tetramer. Complete complexes were purified away from excess biotin-PEG-dNTP by size exclusion chromatography. Each nucleotide type was conjugated and purified separately, then mixed together to create a 4 base mix for sequencing.

[00524] Another type of multi-armed substrate as shown in FIG. 5A was made in a single pot by reacting multiarm PEG NHS with excess Dye-NH2 and propargylamine dNTPs. Various multiarm PEG NHS variants were used ranging from 4-16 arms and ranging in molecular weight from 5K to 40K. After reacting, excess small molecule dye and dNTP were removed by size exclusion chromatography. Each nucleotide type was conjugated and purified independently then mixed together to create a 4 base mix for sequencing.

[00525] Class II substrates as shown in FIG. 5B were made using 1 pot reactions to simultaneously conjugate dye and dNTP. Alkyne-PEG-NHS was reacted with excess propargylamine dNTP. This product (Alkyne-PEG-dNTP) was then purified to homogeneity by chromatography. Multiple PEG lengths were used, varying between 1K and 20K. Dendrimer cores containing a variable, discrete number (12, 24, 48, 96) of azide conjugation sites. Conjugation of Alkyne-Dye and Alkyne-PEG-dNTP to the dendrimer core occurred in a one pot reaction containing excess dye and dNTP species via copper mediated click chemistry. After reacting, excess small molecule dye and dNTP were removed by size exclusion chromatography. Each nucleotide type was conjugated and purified independently then mixed together to create a 4 base mix for sequencing. We note that this scheme allows the ready substitution of alternative cores, such as dextrans, other polymers, proteins, etc.

[00526] Class III nucleotide conjugates (e.g. detectable nucleotide conjugates) as shown in **FIG. 5C** were constructed by reacting 4- or 8-arm PEG NHS with a saturating mixture of biotin and propargylamine dNTP. This reaction was then purified by size exclusion chromatography. The result of this reaction was a multiarm PEG containing a discrete distribution of biotin and nucleotide. This heterogeneous population was then reacted with dye-labeled

streptavidin and purified by size exclusion chromatography. Each nucleotide type was conjugated and purified independently then mixed together to create a 4 base mix for sequencing. We note that the distribution of biotin and nucleotide is tunable by the input ration of Biotin-NH2 to propargylamine dNTP.

Example 2. Detection of Binding Complex

[00527] Binding reactions using the binding complex having PEG nucleotide conjugates (e.g. detectable nucleotide conjugates) were analyzed to detect possible formation of binding complex, and the fluorescence images of the various steps are illustrated in FIGS. 9A-9G. In FIG. 9A, red and green fluorescent images post exposure of DNA rolling circle application (RCA) templates (G and A first base) to 500 nM base labeled nucleotides (A-Cy3 and G-Cy5) in exposure buffer containing 20 nM Klenow polymerase and 2.5 mM Sr+2. Multivalent PEG-substrate compositions were prepared using varying ratios of 4-armed PEG-amine (4ArmPEG-NH), biotin-PEG-amine (Biotin-PEG-NH), and nucleotide (Nuc) as follows: Samples PB1 and PB5, 4ArmPEG-NH: Biotin-PEG-NH: Nuc = 0.25: 1: 0.5; Sample PB2, 4ArmPEG-NH: Biotin-PEG-NH: Nuc =0.125: 0.5: 0.25; Sample PB3, 4ArmPEG-NH: Biotin-PEG-NH: Nuc = 0.25: 1: 0.5. Images were collected after washing with imaging buffer with the same composition as the exposure buffer, but containing no nucleotides or polymerase.

[00528] Contrast was scaled to maximize visualization of the dimmest signals, but no signals persisted following washing with imaging buffer (a. inset). In FIGS. 7B-7E, the fluorescence images showing multivalent PEG-nucleotide (base-labeled) ligands at 500 nM after mixing in the exposure buffer and imaging in the imaging buffer as above. (FIG. 9C. PB2; FIG. 9D. PB3; FIG. 9E. PB5). In FIGS. 7G-7I, the fluorescence images showing further base discrimination by exposure of multivalent ligands to inactive mutants of klenow polymerase FIG. 9G. D882H; FIG. 9H. D882E; FIG. 9I. D882A, and the wild type Klenow (control) enzyme is shown in FIG. 9J.

[00529] Using multivalent ligands formulations, the base discrimination can be enabled by providing polymerase-ligand interactions having increased avidity. In addition it is shown that increased concentration of multivalent ligands can generate higher signals as well as various Klenow mutations that knock out catalytic activity can be used for avidity-based sequencing.

Example 3. Sequencing of Target nucleic Acid Based On Binding Complex

[00530] In order to demonstrate sequencing based on multivalent ligand reporters, 4 known templates were amplified using RCA methods on a low binding substrate. Successive cycles were exposed to exposure buffer containing 20 nM Klenow polymerase and 2.5 mM

$Sr^{+2}$and washed with imaging buffer and imaged. After imaging, the substrates were washed with wash buffer (EDTA and high salt) and blocked nucleotides were added to proceed to the next base. The cycle was repeated for 5 cycles. Spots were detected using standard imaging processing and spot detection and the sequences were called using a two color green and red scheme (G-Cy3 and A-Cy5) to identify the templates being cycled. As shown in FIG. 10A and FIG. 10B, multivalent ligands are able to provide base discrimination through all 5 sequencing cycles.

[0001] Example 4. Control of Nucleotide Dissociation From Ternary Complex

[00531] Ternary complexes are prepared and imaged as in Example 2. The complexes are imaged over varying lengths of time to demonstrate the persistence of the ternary complex, e.g., as long as 60 seconds. After a length of time, the complexes are washed with a buffer identical to the buffer used for the formation of the complexes, only lacking any divalent cation, e.g., 10mM Tris pH 8.0, 0.5mM EDTA, 50mM NaCl, 0.016% Triton X100 (without SrOAc), or, alternatively, the complexes are washed with a buffer identical to the buffer used for the formation of the complexes, which contains a chelating agent but otherwise lacks any divalent cation, e.g., 10mM Tris pH 8.0, 0.5mM EDTA, 50mM NaCl, 0.016% Triton X100 (without SrOAc), with 100nm-100mM EDTA. The fluorescence from the complexes is observed over time allowing observation and quantitation of the dissociation of the ternary complexes. A representative time course of this dissolution is shown in Fig. 6.

Example 5. Extension of Target Nucleic Acid Complementary Sequence

[00532] After preparing, imaging, and dissociating ternary complexes as in Example 4, a deblocking solution is flowed into the chamber containing the bound DNA molecules, sufficient to remove the blocking moiety, such as an O-azidomethyl group, an O-alkyl hydroxylamino group, or an O-amino group, from the 3' end of the elongating DNA strand. Either following or concurrently with this, an extension solution is flowed into the chamber containing the bound DNA molecules. The extension solution contains a buffer, a divalent cation sufficient to support polymerase activity, an active polymerase, and an appropriate amount of all four nucleotides, where the nucleotides are blocked such that they are incapable of supporting further elongation after the addition of a single nucleotide to the elongating DNA strand, such as by incorporation of a 3'-O-azidomnethyl group, a 3'-O-alkyl hydroxylamino group, or a 3'-O-amino group. The elongating strand is thus extended by one and only one base, and the binding of catalytically inactive polymerase and multivalent binding substrate can be used to call the next base in the cycle.

[00533] Alternatively, the nucleotides attached to the multivalent substrate may be attached through a labile bond, such that a buffer may be flowed into the chamber containing the

bound DNA molecules containing a divalent cation or other cofactor sufficient to render the polymerase catalytically active. Prior to, after, or concurrently with this, conditions may be provided that are sufficient to cleave the base from the multivalent substrate such that it may be incorporated into the elongating strand. This cleavage and incorporation causes the dissociation of the label and the polymer backbone of the multivalent substrate while extending the elongating DNA strand by exactly one base. Washing to remove used polymer backbone is carried out, and new multivalent substrate is flowed into the chamber containing the bound DNA molecules, allowing the new base to be called as in Example 1.

Example 6. Use of Detectable nucleotide conjugates with Various Lengths of PEG Branch

[00534] The nucleotide conjugates (e.g. detectable nucleotide conjugates) having varying PEG arm lengths described in Example 3 were subjected to a single sequencing cycle and imaged as described in Example 1. As shown in FIGS. 7A-7G, increasing the length of the PEG branches led to increased signal up to a length corresponding to an apparent average PEG MW of 5K (FIGS. 7A-7D). The use of longer PEG arms than this led to decreases in the fluorescence signal for both Cy3-A and Cy5-G (FIG. 7E-7G). Quantitative measurements of signal intensity are shown graphically in FIG. 8.

Example 7. Use of Detectable nucleotide conjugates with Various Lengths of PEG Branch

[00535] Multivalent substrates were prepared and assembled into binding complexes in the presence and absence of detergent: one set using 10mM Tris pH 8.0, 0.5mM EDTA, 50mM NaCl, 5mM SroAc, 0% TritonX100 (Condition A), and one set using 10mM Tris pH 8.0, 0.5mM EDTA, 50mM NaCl, 5mM SroAc, 0.016% Triton X100. FIG. 11 shows normalized fluorescence from these multivalent substrates bound to DNA clusters, with the substrate complexes formed in the presence (condition B) of Triton-X100 (0.016%) showing clearly enhanced fluorescence intensity.

Example 8. Use of Detectable nucleotide conjugates with Various Lengths of PEG Branch

[00536] Multivalent substrates were prepared and assembled into binding complexes as in Example 2. Complexes were also formed under identical buffer conditions using free labeled nucleotides. Complexes were imaged over the course of 60 min. to characterize the persistence time of the complexes. FIG. 12 shows representative results. Multivalent binding complexes are stable over timescales of >60 minutes (FIG. 12, bottom) while labeled free nucleotides dissociate in less than one minute (FIG. 12, top).

Example 9. Detection of mRNA in human tissue sample using Binding complexes

[00537] In this example, a set of messenger RNAs (mRNAs) will be detected in a tissue sample.

[00538] A human tissue sample will be provided. The tissue sample will be formalin fixed paraffin embedded and will be resected from a human subject. The tissue sample will be sliced into a 5 μm thick section and placed onto the interior surface of a flow cell. An antigen retrieval procedure will be performed using standard procedures, and may include the use of a pressure cooker, microwave, hot pot or heater. The tissue sample will be fixed using formaldehyde and washed with a buffer.

[00539] A set of nucleic acid molecules will be added to the tissue sample and incubated to allow for hybridization of the target-binding sequence of the nucleic acid molecules to mRNA targets in the sample. The set of nucleic acid molecules will target a set of mRNAs in the tissue sample by hybridizing directly to the mRNA with binding sequences that are 15 nucleotides in length and complement the sequence of the mRNA. At least five nucleic acid molecules of the set of nucleic acid molecules will hybridize to each mRNA that will be detected. The five nucleic acid molecules per mRNA will hybridize to different sequences on the mRNA. Each nucleic acid molecule of the set of nucleic acid molecule s will also contain an index sequence. The index sequence will be specific to the corresponding mRNA that the nucleic acid molecule binds to. Thus, each of the five nucleic acid molecules specific to a given mRNA will contain the same index sequence.

[00540] A set of padlock probes will be added to the sample. Each padlock probe will include two index-binding regions that hybridize to a portion of the index sequence. The first index-binding region will be at the 5' end of the padlock probe and the second index-binding region will be at the 3' end of the padlock probe. The set of padlock probes will bind to one of the index sequences included in the oligonucleotides hybridized to the tissue sample. After hybridizing, the padlock probes will be ligated using a T4 ligase to form circular nucleic acids. Rolling circle amplification will then be performed using phi29 polymerase to generate a concatemer based on the sequence of each of the circular nucleic acids. Each concatemer will comprise multiple copies of the reverse complement of the corresponding index sequence.

[00541] A set of primers will be added to the sample and allowed to hybridize to the concatemers. The primers will form a double-stranded region with an extendable overhang region. Next, a set of nucleotide conjugates (e.g. detectable nucleotide conjugates) will be added to the sample. Each nucleotide conjugate (e.g. detectable nucleotide conjugate) will include multiple copies of a given nucleotide, a fluorophore and a polymer core that the multiple copies of a given nucleotide and fluorophore are conjugated to. The set of nucleotide conjugates (e.g.

detectable nucleotide conjugates) will include four different nucleotide conjugates (e.g. detectable nucleotide conjugates) that each include one of A, C, G, and T, and a spectrally distinct fluorophore. Binding complexes will be formed between multiple copies of the overhang sequence in a given concatemer and multiple nucleotide of a given type of a detectable nucleotide conjugate. The fluorophore of each binding complex will correlate with the nucleotide identity of the terminal nucleotide of the overhang within the concatemer. The sample will be imaged using a fluorescent microscope to acquire a two-dimensional scan of the entire tissue sample across each of the four fluorophore channels.

[00542] After imaging, the binding complexes will be disrupted by introducing of a low salt buffer and elevated temperatures. The sample will be washed to remove the nucleotide conjugates (e.g. detectable nucleotide conjugates) and untethered nucleotides and a polymerase will be added to the sample. The terminal nucleotide position of the overhang region within each concatemer will be extended. Multiple rounds of binding complex formation, imaging, binding complex disruption, and single nucleotide extension will be performed until each of the nucleotides associated with the index sequence has been detected. The identity of mRNA targets within the sample will then be determined by comparing the signal at each location across multiple cycles with the corresponding index sequence assigned to each mRNA.

Example 10. *in situ* Detection of RNA and Sequencing

[00543] Culturing Cells on Coated Flowcells

[00544] Cells in fresh cell culture media were seeded on the coated flowcells. The cell culture media included D-MEM high glucose (Thermo Fisher Scientific, catalog No. 11965118), fetal bovine serum (10% FBS; Thermo Fisher Scientific, catalog No. A3160402), MEM non-essential amino acids (0.1 mM MEM, Thermo Fisher Scientific, catalog No. 11140050), L-glutamine (6 mM L-glutamine, Thermo Fisher Scientific, catalog No. A2916801), MEM sodium pyruvate (1 mM sodium pyruvate, Thermo Fisher Scientific, catalog No. 11360070), and an antibiotic (1% penicillin-streptomycin-glutamine, Thermo Fisher, catalog No. 10378016). The seeded cells were cultured for 2 days at 37 °C with a humidified atmosphere of approximately 5-10% carbon dioxide in air. The cells remained as whole cells on the flowcell, and were not embedded in paraffin or sliced.

[00545] Cell Fixation and Permeabilization

[00546] The cells were washed with DPBS and then fixed with 3.7% paraformaldehyde in 1X PBS and water for about 30 minutes at room temperature. The cells were washed once with PBS.

[00547] The fixed cells were permeabilized with 70% ethanol for about 15 minutes at room temperature, then washed with 1X PBS. The fixed and permeabilized cells remained whole and were not sectioned into slices. The cells were not transgenic cells.

[00548] Padlock Probe Hybridization

[00549] The cells on the flowcells were flowed with 1 uM of linear padlock probes in 1X padlock probe hybridization buffer at 55 °C for approximately 90 minutes. The cells were washed with 1X padlock probe hybridization buffer for approximately 15 minutes at 55 °C. The cells were washed with 1X PBS. The 1X padlock probe hybridization buffer comprised: 2X sodium-saline citrate buffer (SSC), pH 7.0, 20%formamide, and 0.5 U/uL RiboLock RNase Inhibitor. The linear padlock probes comprised SEQ ID NO 4: /5phos/ACGACCAAATCCGTTGACTCCGACCTTCACCTTCCCCATGGCAGCCGCATCTT CTTTTTCTATGACATGAGGACTTTTTACTGACTGCGTCTATTT AGTGGAGCCTCCCGTTCTCAGCCTTGACGGTGCCATGGAATTTGCCATGGGTG GAATCATATTGGA -3'.

[00550] Gap Fill-In Reaction and Ligation

[00551] The cells on the flowcells were flowed with reverse transcription reaction mix for approximately 80 minutes at 37 °C. The flowcells were washed with 1X PBS. The reverse transcription reaction mixture included reverse transcriptase and ligase.

[00552] Rolling Circle Amplification

[00553] The cells on the flowcells were flowed with 1 uM soluble amplification primers, and 1X RCA hybridization buffer, and incubated at 55 °C for approximately 30 minutes. The flowcells were washed with 1X RCA hybridization buffer at 55 °C for 10 minutes. The 1X RCA hybridization buffer comprised: 2X SSC, pH 7.0, and 20% formamide.

[00554] The soluble amplification primers comprised the sequence: 5'- GGCTCCACTAAATAGACGCAGTCAGTAA -3'

[00555] The cells on the flowcells were flowed with a first reaction mixture and incubated at 37 °C for approximately 30 minutes. The first reaction mixture comprised a strand-displacing DNA polymerase, a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, and a viscosity agent. The pH of the first reaction mixture was such that the strand-displacing DNA polymerase was inactive.

[00556] Rolling circle amplification was initiated by flowing onto the cells a second reaction mixture and incubated at 40 °C for approximately 2 hours. The rolling circle amplification generated concatemer molecules inside the cells. The second reaction mixture comprised a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions,

potassium ions, ammonium ions, a detergent, a reducing reagent, a viscosity agent and 300 nM compaction oligonucleotides. The pH of the second reaction mixture was such that the strand-displacing DNA polymerase was active.

[00557] The compaction oligonucleotide comprised SEQ ID NO 3: 5'- CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm -3'.

[00558] *In Situ* Sequencing Using 2-Stage Sequencing Method

[00559] The cells on the flowcells were flowed with 1 uM of sequencing primer and 1X sequencing primer hybridization buffer, and incubated at 4 °C overnight, at 40 °C for 20 minutes, and at room temperature for 2 hours. The cells were washed with 1X PBS. The sequencing primer comprised SEQ ID NO 2: 5'- CATGGAATTTGCCATGGGTGGAA -3'. The sequencing primer hybridization buffer comprised: 3X SSC, pH 7.0, and 300 nM compaction oligonucleotides. The sequencing primers were designed to hybridize to the concatemer molecules inside the cells.

[00560] The compaction oligonucleotide comprised SEQ ID NO 3: 5'- CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm -3'.

[00561] A two-stage sequencing method was used to sequence the concatemers inside the cells. Each sequencing cycle included a first stage and a second stage. The first stage employed a first sequencing polymerase and fluorescently labeled multivalent molecules. The second stage employed a second sequencing polymerase and non-labeled nucleotide analogs.

[00562] The first stage sequencing was conducted with fluorescently labeled multivalent molecules. An image of the cells on the flowcells was obtained prior to the start of sequencing. A solution of a first sequencing polymerase was flowed onto the flowcells and incubated at 42 °C for about 10 minutes to form polymerase complexes on the concatemers inside the cells. A trap reagent was flowed onto the cells. The trap reagent included a mixture of fluorescently labeled multivalent molecules (e.g., about 40-100 nM) (see FIGs. 21-24) and a non-catalytic cation (e.g., strontium, barium or calcium). The mixture of fluorescently labeled multivalent molecules included dATP, dGTP, dCTP and dUTP. The flowcells were incubated at 42 °C for about 10 minutes to permit the multivalent molecules to bind the polymerase complexes, without incorporation of the nucleotide moieties, and form avidity complexes on the concatemers inside the cells. The flowcells were washed at room temperature with a trap reagent that lacked the fluorescently labeled multivalent molecules. The flowcells were washed with an imaging buffer. An image was obtained using a fluorescent microscope, for example, 1 minute/FOV, 4 channels, 200 ms/frame, using Z-stack imaging. The multivalent molecules and first sequencing polymerases were removed by washing the flowcells twice at 42 °C with a removal reagent. The flowcells were washed four times at 52 °C with a wash buffer.

[00563] The second stage sequencing was conducted with non-labeled nucleotide analogs. A solution of a second sequencing polymerase, a mixture of non-labeled nucleotide analogs and a catalytic divalent cation (e.g., magnesium) was flowed onto the flowcells and incubated at 52 °C for about 2-5 minutes to permit incorporation of the non-labeled nucleotide analogs. The mixture of non-labeled nucleotide analogs included 3'O-methylazido nucleotides with dATP, dGTP, dCTP and dTTP. The flowcells were washed twice with a removal reagent at 51 °C for about 20 seconds. The 3' blocking moieties were removed from the incorporated nucleotide by washing the flowcells twice with a cleaving reagent at 51 °C for about 40 seconds. The flowcells were washed twice with a wash buffer. One sequencing cycle included a first stage and a second stage.

[00564] The next sequencing cycle was conducted using the two-stage sequencing method described above. The cells on the flowcells were subjected to 100 sequencing cycles which generated a first plurality of sequencing read products. The three graphs shown in FIG. 34 The intensity measurements (top), background measurements (middle), and phasing and pre-phasing measurements (bottom).

Example 11. *in situ* Detection of RNA and Sequencing

[00565] Culturing Cells on Coated Flow cells

[00566] Cells in fresh cell culture media were seeded on the coated flow cells. The cell culture media included D-MEM high glucose (Thermo Fisher Scientific, catalog No. 11965118), fetal bovine serum (10% FBS; Thermo Fisher Scientific, catalog No. A3160402), MEM non-essential amino acids (0.1 mM MEM, Thermo Fisher Scientific, catalog No. 11140050), L-glutamine (6 mM L-glutamine, Thermo Fisher Scientific, catalog No. A2916801), MEM sodium pyruvate (1 mM sodium pyruvate, Thermo Fisher Scientific, catalog No. 11360070), and an antibiotic (1% penicillin-streptomycin-glutamine, Thermo Fisher, catalog No. 10378016). The seeded cells were cultured for 2 days at 37 °C with a humidified atmosphere of approximately 5-10% carbon dioxide in air. The cells remained as whole cells on the flowcell, and were not embedded in paraffin or sliced. The cells were not transgenic cells.

[00567] Cell Fixation and Permeabilization

[00568] The cells were washed with DPBS and then fixed with 3.7% paraformaldehyde in 1X PBS and water for about 30 minutes at room temperature. The cells were washed once with PBS.

[00569] The fixed cells were permeabilized with 70% ethanol for about 15 minutes at room temperature, then washed with 1X PBS. The fixed and permeabilized cells remained whole and were not sectioned into slices.

[00570] Padlock Probe Hybridization

[00571] The cells on the flow cells were flowed with 1 μM of linear padlock probes in 1X padlock probe hybridization buffer at 55 °C for approximately 90 minutes. The linear padlock probes move inside the cells. The cells were washed with 1X padlock probe hybridization buffer for approximately 15 minutes at 55 °C. The cells were washed with 1X PBS. The 1X padlock probe hybridization buffer comprised: 2X sodium-saline citrate buffer (SSC), pH 7.0, 20%formamide, and 0.5 U/μL RiboLock RNase Inhibitor. The linear padlock probes comprised SEQ ID NO 4:

/5phos/ACGACCAAATCCGTTGACTCCGACCTTCACCTTCCCCATGGCAGCCGC ATCTTCTTTTTCTATGACATGAGGACTTTTTACTGACTGCGTCTATTTAGTGG AGCCTCCCGTTCTCAGCCTTGACGGTGCCATGGAATTTGCCATGGGTGGAAT CATATTGGA -3'. The /5phos/ may comprise a 5' phosphorylation.

[00572] Gap Fill-In Reaction and Ligation

[00573] Individual padlock probes comprise left and right binding arms that were designed to hybridize directly to target RNA sequences and form a gap between the left and right binding arms. A reverse transcription reaction mix was used to fill-in the gap and form a nick. The nick was enzymatically closed with ligase to generate covalently closed padlock probes that were hybridized to target RNA. Reverse transcriptase was not used to generate cDNA from RNA.

[00574] The cells on the flow cells were flowed with reverse transcription reaction mix for approximately 80 minutes at 37 °C. The flow cells were washed with 1X PBS. The reverse transcription reaction mixture included reverse transcriptase, a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP) and ligase.

[00575] Rolling Circle Amplification

[00576] The cells on the flow cells were flowed with 1 μM soluble amplification primers, and 1X RCA hybridization buffer, and incubated at 55 °C for approximately 30 minutes. The flow cells were washed with 1X RCA hybridization buffer at 55 °C for 10 minutes. The 1X RCA hybridization buffer comprised: 2X SSC, pH 7.0, and 20% formamide.

[00577] The soluble amplification primers comprised SEQ ID NO 1:

5'- GGCTCCACTAAATAGACGCAGTCAGTAA -3' (Primer 1). The soluble amplification primers were designed to hybridize to a portion of the padlock probe that was not hybridized to the target RNA.

[00578] The cells on the flow cells were flowed with a first reaction mixture and incubated at 37 °C for approximately 30 minutes. The first reaction mixture comprised a strand-displacing DNA polymerase, a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, and a viscosity

agent. The pH of the first reaction mixture was such that the strand-displacing DNA polymerase was inactive.

[00579] Rolling circle amplification was initiated by flowing onto the cells a second reaction mixture and incubated at 40 °C for approximately 2 hours. The rolling circle amplification generated concatemer molecules inside the cells. The second reaction mixture comprised a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, a viscosity agent and 300 nM compaction oligonucleotides. The pH of the second reaction mixture was such that the strand-displacing DNA polymerase was active.

[00580] The compaction oligonucleotide comprised SEQ ID NO 3:

5'-CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm-3'

(Compaction oligonucleotide). The Um nucleotide may comprise a 2'-O-methyluridine.

[00581] The compaction oligonucleotides comprise a 5′ region that binds a first portion of the concatemer molecule and a 3′ region that binds a second portion of the same concatemer molecule to pull together distal portions of the concatemer molecule causing compaction of the concatemer molecule to form a DNA nanoball.

[00582] *in situ* Sequencing Using 2-Stage Sequencing Method

[00583] The cells on the flow cells were flowed with 1 μM of sequencing primer and 1X sequencing primer hybridization buffer, and incubated at 4 °C overnight, at 40 °C for 20 minutes, and at room temperature for 2 hours. The cells were washed with 1X PBS. The sequencing primer comprised SEQ ID NO 2: 5'- CATGGAATTTGCCATGGGTGGAA -3' (Primer 2). The sequencing primer hybridization buffer comprised: 3X SSC, pH 7.0, and 300 nM compaction oligonucleotides. The sequencing primers were designed to hybridize to the concatemer molecules inside the cells.

[00584] The compaction oligonucleotide comprised SEQ ID NO 3:

5'- CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm-3'

(Compaction oligonucleotide).

[00585] A two-stage sequencing method was used to sequence the concatemers inside the cells. Each sequencing cycle included a first stage and a second stage. The first stage employed a first sequencing polymerase and fluorescently labeled multivalent molecules. The second stage employed a second sequencing polymerase and non-labeled nucleotide analogs.

[00586] The first stage sequencing was conducted with fluorescently labeled multivalent molecules. An image of the cells on the flow cells was obtained prior to the start of sequencing. A solution of a first sequencing polymerase was flowed onto the flow cells and incubated at 42 °C for about 10 minutes to form polymerase complexes on the concatemers inside

the cells. A trap reagent was flowed onto the cells. The trap reagent included a mixture of fluorescently labeled multivalent molecules (e.g., about 40-100 nM) (see FIGs. 7-10) and a non-catalytic cation (e.g., strontium, barium or calcium). The mixture of fluorescently labeled multivalent molecules included dATP, dGTP, dCTP and dUTP. The flow cells were incubated at 42 °C for about 10 minutes to permit the multivalent molecules to bind the polymerase complexes, without incorporation of the nucleotide moieties, and form avidity complexes on the concatemers inside the cells. The flow cells were washed at room temperature with a trap reagent that lacked the fluorescently labeled multivalent molecules. The flow cells were washed with an imaging buffer. An image was obtained using a fluorescent microscope, for example, 1 minute/FOV, 4 channels, 200 ms/frame, using Z-stack imaging. The multivalent molecules and first sequencing polymerases were removed by washing the flow cells twice at 42 °C with a removal reagent. The flow cells were washed four times at 52 °C with a wash buffer.

[00587] The second stage sequencing was conducted with non-labeled nucleotide analogs. A solution of a second sequencing polymerase, a mixture of non-labeled nucleotide analogs and a catalytic divalent cation (e.g., magnesium) was flowed onto the flow cells and incubated at 52 °C for about 2-5 minutes to permit incorporation of the non-labeled nucleotide analogs. The mixture of non-labeled nucleotide analogs included 3'O-methylazido nucleotides with dATP, dGTP, dCTP and dTTP. The flow cells were washed twice with a removal reagent at 51 °C for about 20 seconds. The 3' blocking moieties were removed from the incorporated nucleotide by washing the flow cells twice with a cleaving reagent at 51 °C for about 40 seconds. The flow cells were washed twice with a wash buffer. One sequencing cycle included a first stage and a second stage.

[00588] The next sequencing cycle was conducted using the two-stage sequencing method described above. The cells on the flow cells were subjected to 100 sequencing cycles which generated a first plurality of sequencing read products. The three graphs shown in FIG. 20 The intensity measurements (top), background measurements (middle), and phasing and pre-phasing measurements (bottom).

Example 12. *in situ* detection of RNA and sequencing in tissue

[00589] Tissue Sample on Flow cells

[00590] Fresh frozen tissue is resected from a human subject. The tissue sample is sliced into sections that are 5-10 µm thick and placed onto a sequencing flow cell. For example, the tissue sample will be sliced into sections that are 7-8 µm thick. The tissue slices are fixed and permeabilized. For example, the tissue slices are fixed using formaldehyde and washed with a buffer. In another example, the tissue slices are permeabilized using acetone. The tissue slices

will remain on the flow cell during this workflow, including: padlock probe hybridization, gap fill-in and ligation reactions, rolling circle amplification and two-stage sequencing.

[00591] Padlock Probe Hybridization

[00592] A plurality of padlock probes is flowed onto the tissue slices in a manner similar to the method described in Example 10 above. For example, the plurality of padlock probes is flowed onto the tissue slices in 1X padlock probe hybridization buffer and incubated for approximately 15-30 minutes at 37 °C, or incubated for approximately one to three hours at room temperature. The plurality of padlock probes move inside the cells of the sliced tissue. The tissue slices are washed with 1X PBS. The 1X padlock probe hybridization buffer will comprise: 2X sodium-saline citrate buffer (SSC), pH 7.0, 20%formamide, and 0.5 U/μL RiboLock RNase Inhibitor.

[00593] Individual padlock probes in the plurality will comprise single-stranded DNA. In some embodiments, individual padlock probes in the plurality will comprise paired left and right binding arms, and an intervening region between the left and right binding arms (e.g., a backbone region) carrying a barcode index sequence that identifies the target RNA molecule. The left binding arm of a padlock probe will hybridize to a first target sequence within a sub-region of the RNA molecule. The right binding arm of the same padlock probe will hybridize to a second target sequence within the same a sub-region of the same RNA molecule. In some embodiments, the plurality of padlock probes comprise a mixture of padlock probes that will hybridize to different target RNA molecules. In some embodiments, at least two padlock probes in the plurality comprise different barcode index sequences which identify different target RNA molecules, and paired left and right binding arms that hybridize to target sequences of different target RNA molecules.

[00594] In some embodiments, individual padlock probes in the plurality comprise left and right binding arms that will hybridize to a coding RNA sequence or non-coding RNA sequence. In some embodiments, the left and right binding arms of a padlock probe will hybridize to tRNA, rRNA, small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), microRNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), or antisense RNA. In some embodiments, the left and right binding arms of a padlock probe will hybridize to an intron sequence, an exon sequence or an intron-exon junction sequence. In some embodiments, the left and right binding arms of a padlock probe will hybridize to an RNA splice variant sequence.

[00595] In some embodiments, individual padlock probes in the plurality comprise an intervening region that carries a universal binding site for a sequencing primer. In some

embodiments, the intervening region will also carry a universal binding site for a compaction oligonucleotide.

[00596] In some embodiments, individual padlock probes in the plurality comprise a left binding arm that is 10-30 nucleotides in length. In some embodiments, individual padlock probes in the plurality comprise a right binding arm that is 10-30 nucleotides in length. In some embodiments, the left and right binding arms of individual padlock probes will be the same length or different lengths.

[00597] In some embodiments, the plurality of padlock probes will include a set of padlock probes comprising at least a first and a second sub-population of padlock probes. The padlock probes of the first sub-population are designed to hybridize to a first sub-region of an RNA molecule. The padlock probes of the second sub-population are designed to hybridize to a second sub-region of the same RNA molecule. The plurality of padlock probes will further comprise a third, fourth, fifth, sixth, seventh, eighth, ninth, tenth or more sub-populations of padlock probes, wherein individual padlock probes in a given sub-population will hybridize to a designated sub-region of the same RNA molecule. Individual padlock probes in a given sub-population include paired left and right binding arms, and an intervening region between the left and right binding arms (e.g., a backbone region) carrying a barcode index sequence that identifies the target RNA molecule. The left binding arm will hybridize to a first target sequence within a sub-region of the RNA molecule. The right binding arm will hybridize to a second target sequence within the same a sub-region of the same RNA molecule. Thus, individual padlock probes in different sub-populations comprise the same barcode index sequence which identifies the same target RNA molecule, and paired left and right binding arms that hybridize to different sub-regions of the same RNA molecule.

[00598] Gap Fill-In Reaction and Ligation

[00599] The left and right binding arms of individual padlock probes are designed to hybridize directly to RNA and form a gap between the left and right binding arms. In some embodiments, the gap is 2-50 bases in length, or 2-40 bases in length, or 2-30 bases in length, or 2-20 bases in length, or 2-10 bases in length. After hybridizing the plurality of padlock probes to RNA, a reverse transcription reaction mix is used to fill-in the gap and form a nick. The nick is enzymatically closed with a ligase to generate covalently closed padlock probes that are hybridized to target RNA. Reverse transcriptase is not used to generate full length cDNA from RNA.

[00600] The gap fill-in and ligation reaction is conducted in a manner similar to the method described in Example 10 above. For example, the tissue slices on the flow cells are reacted with a reverse transcription reaction mix for approximately 80 minutes at 37 °C.

Alternatively, the tissue slices are reacted with a reverse transcription reaction mix overnight (e.g., at least 8 hours) at 45 °C. The tissue slices are washed with 1X PBS. The reverse transcription reaction mixture will comprise a reverse transcriptase, a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP) and ligase. The reverse transcriptase will comprise a reverse transcriptase from AMV (avian myeloblastosis virus), M-MuLV (moloney murine leukemia virus), or HIV (human immunodeficiency virus). In some embodiment, the reverse transcriptase enzyme will comprise a recombinant enzyme that exhibits reduced RNase H activity, for example REVERTAID (e.g., from Thermo Fisher Scientific, catalog No. EP0441). In some embodiments, the reverse transcriptase will be a commercially-available enzyme, including MULTISCRIBE (e.g., from Thermo Fisher Scientific, catalog # 4311235), THERMOSCRIPT (e.g., from Thermo Fisher Scientific, catalog # 12236-014), or ARRAYSCRIPT (e.g., from Ambion, catalog No. AM2048). In some embodiments, the reverse transcriptase enzyme will comprise SUPERSCRIPT II (e.g., catalog No. 18064014), SUPERSCRIPT III (e.g., catalog No. 18080044), or SUPERSCRIPT IV enzymes (e.g., catalog No. 18090010) (all SUPERSCRIPT enzymes from Invitrogen). In some embodiments, the reverse transcription reaction will include an RNase inhibitor. The ligase will comprise a DNA ligase or an RNA ligase. The ligase will comprise a T4 ligase, a SplintR ligase, a T3 ligase, a T4 ligase, a T7 ligase, an TtcB ligase, a Taq ligase or a combination thereof. After the gap-fill and ligation reactions, the tissue slices will be washed with 1X PBS.

[00601] Rolling Circle Amplification

[00602] A rolling circle amplification reaction will be conducted in a manner similar to the method described in Example 10 above. For example, the sliced tissue on the flow cells will be reacted with 1 μM soluble amplification primers, and 1X RCA hybridization buffer, and incubated at 55 °C for approximately 30 minutes. The sliced tissue will be washed with 1X RCA hybridization buffer at 55 °C for 10 minutes. The 1X RCA hybridization buffer will comprise: 2X SSC, pH 7.0, and 20% formamide. The soluble amplification primer will be designed to hybridize to a portion of the padlock probe that does not hybridize to the target RNA.

[00603] The sliced tissue on the flow cell will be reacted with a first reaction mixture and incubated at 37 °C for approximately 30 minutes. The first reaction mixture will comprise a strand-displacing DNA polymerase, a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, and a viscosity agent. The first reaction mixture will have a pH at which the strand-displacing DNA polymerase is inactive. The strand-displacing DNA polymerase will comprise phi29, Taq polymerase, T4 polymerase, T7 polymerase, or a combination thereof.

[00604] Rolling circle amplification will be initiated by flowing onto the sliced tissue a second reaction mixture and incubated at 40 °C for approximately 2 hours. Alternatively, the sliced tissue will be incubated with the second reaction mixture overnight (e.g., at least 8 hours) at room temperature. The second reaction mixture will comprise a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, a viscosity agent and 300 nM compaction oligonucleotides. The second reaction mixture will have a pH at which the strand-displacing DNA polymerase is active. The rolling circle amplification will generate concatemer molecules inside the sliced tissue.

[00605] The compaction oligonucleotides will comprise a 5′ region that binds a first portion of the concatemer molecule and a 3′ region that binds a second portion of the same concatemer molecule to pull together distal portions of the concatemer molecule causing compaction of the concatemer molecule to form a DNA nanoball.

[00606] *in situ* Sequencing Using 2-Stage Sequencing Method

[00607] The sliced tissue on the flow cells will be subjected to a two-stage sequencing method in a manner similar to the method described in Example 10 above. For example, the sliced tissue on the flow cells will be flowed with 1 μM of sequencing primer and 1X sequencing primer hybridization buffer, and incubated at 4 °C overnight, at 40 °C for 20 minutes, and at room temperature for 2 hours. The sliced tissue will be washed with 1X PBS. The universal sequencing primer will comprise SEQ ID NO 2: 5'-CATGGAATTTGCCATGGGTGGAA -3' (Primer 2). The sequencing primer hybridization buffer will comprise: 3X SSC, pH 7.0, and 300 nM compaction oligonucleotides. The sequencing primers will be designed to hybridize to the concatemer molecules inside the sliced tissue.

[00608] The universal compaction oligonucleotide will comprise SEQ ID NO 3: 5'- CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm -3' (Compaction oligonucleotide).

[00609] A two-stage sequencing method will be used to sequence the concatemers inside the sliced tissue. Each sequencing cycle will include a first stage and a second stage. The first stage will employ a first sequencing polymerase and fluorescently labeled multivalent molecules. The second stage will employ a second sequencing polymerase and non-labeled nucleotide analogs.

[00610] The first stage sequencing is conducted with fluorescently labeled multivalent molecules. An image of the cells on the flow cells is obtained prior to the start of sequencing. A solution of a first sequencing polymerase is flowed onto the sliced tissue and

incubated at 42 °C for about 10 minutes to form polymerase complexes on the concatemers inside the sliced tissue. A trap reagent is flowed onto the cells. The trap reagent will include a mixture of fluorescently labeled multivalent molecules (e.g., about 40-100 nM) (see FIGs. 7-10) and a non-catalytic cation (e.g., strontium, barium and/or calcium). The mixture of fluorescently labeled multivalent molecules will include dATP, dGTP, dCTP and dUTP. The flow cells are incubated at 42 °C for about 10 minutes to permit the multivalent molecules to bind the polymerase complexes, without incorporation of the nucleotide moieties, and form avidity complexes on the concatemers inside the sliced tissue. The flow cells are washed at room temperature with a trap reagent that lacks the fluorescently labeled multivalent molecules. The flow cells are washed with an imaging buffer. An image is obtained using a fluorescent microscope, for example, 1 minute/FOV, 4 channels, 200 ms/frame, using Z-stack imaging. The fluorescent microscope will be used to obtain two-dimensional scans of the sliced tissue of the fluorescent signals emitted from the labeled multivalent molecules during the first stage of sequencing. Four fluorescent channels will be used to conduct four color detection of the fluorescent signals emitted from the labeled multivalent molecules during the first stage of sequencing. The two-dimensional scans will include the entire sliced tissue or a portion of the sliced tissue. The two-dimensional scans will be conducted in a single z-plane to obtain images at a single depth of the sliced tissue. The two-dimensional scans will be conducted at two or more z-planes to obtain images at different depths of the sliced tissue. The sequences of the target RNA will be determined from the images. The multivalent molecules and first sequencing polymerases are removed by washing the sliced tissue twice at 42 °C with a removal reagent. The flow cells are washed four times at 52 °C with a wash buffer.

[00611] The second stage sequencing is conducted with non-labeled nucleotide analogs. A solution of a second sequencing polymerase, a mixture of non-labeled nucleotide analogs and a catalytic divalent cation (e.g., magnesium and/or manganese) is flowed onto the sliced tissue and incubated at 52 °C for about 2-5 minutes to permit incorporation of the non-labeled nucleotide analogs. The mixture of non-labeled nucleotide analogs will include 3'O-methylazido nucleotides with dATP, dGTP, dCTP and dTTP. The flow cells are washed twice with a removal reagent at 51 °C for about 20 seconds. The 3' blocking moieties are removed from the incorporated nucleotide by washing the sliced tissue twice with a cleaving reagent at 51 °C for about 40 seconds. The flow cells are washed twice with a wash buffer. One sequencing cycle included a first stage and a second stage.

[00612] The next sequencing cycle was conducted using the two-stage sequencing method described above. The cells on the flow cells will be subjected to 100 sequencing cycles which will generate a first plurality of sequencing read products. At each sequencing cycle,

intensity measurement, background measurements, and phasing and pre-phasing measurements will be obtained (e.g., see the three graphs shown in FIG. 34).

Example 13. *in situ* Detection of DNA and sequencing in tissue

[00613] Tissue Sample on Flow cells

[00614] Tissue samples will be prepared in a manner similar to the method described in Example 11 above. For example, fresh frozen tissue is resected from a human subject. The tissue will comprise a tumor sample and patient-matched tissue from the same human subject. The tissue sample is sliced into sections that are 5-10 μm thick and placed onto a sequencing flow cell. For example, the tissue sample will be sliced into sections that are 7-8 μm thick. The tissue slices are fixed and permeabilized. For example, the tissue slices are fixed using formaldehyde and washed with a buffer. In another example, the tissue slices are permeabilized using acetone. The tissue slices will remain on the flow cell during this workflow, including: padlock probe hybridization, gap fill-in and ligation reactions, rolling circle amplification and two-stage sequencing.

[00615] Padlock Probe Hybridization

[00616] The DNA in the tissue slices will be pre-treated with a restriction enzyme. For example, the DNA will be digested at 37 °C for 1-2 hours with MseI, SfiI, MluI, SacII, NotI, SdaI, AscI, EagI, PacI or any combination thereof. A plurality of padlock probes is flowed onto the tissue slices in a manner similar to the method described in Example 10 above. For example, the plurality of padlock probes is flowed onto the tissue slices in 1X padlock probe hybridization buffer and incubated for 15-30 minutes at 37 °C, or incubated for approximately one to three hours at room temperature. The plurality of padlock probes move inside the cells of the sliced tissue. The tissue slices are washed with 1X PBS. The 1X padlock probe hybridization buffer will comprise: 2X sodium-saline citrate buffer (SSC), pH 7.0, and 20%formamide.

[00617] Individual padlock probes in the plurality will comprise single-stranded DNA. In some embodiments, individual padlock probes in the plurality will comprise paired left and right binding arms, and an intervening region between the left and right binding arms (e.g., a backbone region) carrying a barcode index sequence that identifies the target DNA molecule. The left binding arm of a padlock probe will hybridize to a first target sequence within a sub-region of the DNA molecule. The right binding arm of the same padlock probe will hybridize to a second target sequence within the same a sub-region of the same DNA molecule. In some embodiments, the plurality of padlock probes comprise a mixture of padlock probes that will hybridize to different target DNA molecules. In some embodiments, at least two padlock probes in the plurality comprise different barcode index sequences which identify different target DNA

molecules, and paired left and right binding arms that hybridize to target sequences of different target DNA molecules.

[00618] In some embodiments, individual padlock probes in the plurality comprise left and right binding arms that will hybridize to DNA that encodes a polypeptide or will hybridize to non-coding DNA. The left and right binding arms of a padlock probe will hybridize to a 5' untranslated sequence on DNA.

[00619] In some embodiments, individual padlock probes in the plurality comprise an intervening region that carries a universal binding site for a sequencing primer. In some embodiments, the intervening region will also carry a universal binding site for a compaction oligonucleotide.

[00620] In some embodiments, individual padlock probes in the plurality comprise a left binding arm that is 10-30 nucleotides in length. In some embodiments, individual padlock probes in the plurality comprise a right binding arm that is 10-30 nucleotides in length. In some embodiments, the left and right binding arms of individual padlock probes will be the same length or different lengths.

[00621] In some embodiments, the plurality of padlock probes will include a set of padlock probes comprising at least a first and a second sub-population of padlock probes. The padlock probes of the first sub-population are designed to hybridize to a first sub-region of an DNA molecule. The padlock probes of the second sub-population are designed to hybridize to a second sub-region of the same DNA molecule. The plurality of padlock probes will further comprise a third, fourth, fifth, sixth, seventh, eighth, ninth, tenth or more sub-populations of padlock probes, wherein individual padlock probes in a given sub-population will hybridize to a designated sub-region of the same DNA molecule. Individual padlock probes in a given sub-population include paired left and right binding arms, and an intervening region between the left and right binding arms (e.g., a backbone region) carrying a barcode index sequence that identifies the target DNA molecule. The left binding arm will hybridize to a first target sequence within a sub-region of the DNA molecule. The right binding arm will hybridize to a second target sequence within the same a sub-region of the same DNA molecule. Thus, individual padlock probes in different sub-populations comprise the same barcode index sequence which identifies the same target DNA molecule, and paired left and right binding arms that hybridize to different sub-regions of the same DNA molecule.

[00622] Ligation Reaction

[00623] The left and right binding arms of individual padlock probes are designed to hybridize directly to DNA and form a nick between the left and right binding arms. After

hybridizing the plurality of padlock probes to target DNA, the nick is enzymatically closed with a ligase to generate covalently closed padlock probes that are hybridized to target DNA.

[00624] The ligation reaction is conducted by reacting the tissue slices with a ligation reaction mix for approximately 15-30 minutes at 37 °C. The ligation reaction mix will comprise a DNA ligase and a ligase buffer. The ligase buffer will comprise bovine serum albumin (BSA), Tris-acetate (pH 7.9), Mg-acetate, K-acetate, Tween 20, DTT, and ATP. The DNA ligase will comprise T4 DNA ligase. After the ligation reaction the tissue slices will be washed with 1X PBS.

[00625] Rolling Circle Amplification

[00626] A rolling circle amplification reaction will be conducted in a manner similar to the method described in Example 10 above. For example, the sliced tissue on the flow cells will be flowed with 1 μM soluble amplification primers, and 1X RCA hybridization buffer, and incubated at 55 °C for approximately 30 minutes. The sliced tissue will be washed with 1X RCA hybridization buffer at 55 °C for 10 minutes. The 1X RCA hybridization buffer will comprise: 2X SSC, pH 7.0, and 20% formamide. The soluble amplification primer will be designed to hybridize to a portion of the padlock probe that does not hybridize to the target DNA sequence.

[00627] The sliced tissue on the flow cell will be flowed with a first reaction mixture and incubated at 37 °C for approximately 30 minutes. The first reaction mixture will comprise a strand-displacing DNA polymerase, a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, and a viscosity agent. The first reaction mixture will have a pH at which the strand-displacing DNA polymerase is inactive. The strand-displacing DNA polymerase will comprise phi29, Taq polymerase, T4 polymerase, T7 polymerase, or a combination thereof.

[00628] Rolling circle amplification will be initiated by flowing onto the sliced tissue a second reaction mixture and incubated at 40 °C for approximately 2 hours. Alternatively, the sliced tissue will be incubated with the second reaction mixture overnight (e.g., at least 8 hours) at room temperature. The second reaction mixture will comprise a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, a viscosity agent and 300 nM compaction oligonucleotides. The second reaction mixture will have a pH at which the strand-displacing DNA polymerase is active. The rolling circle amplification will generate concatemer molecules inside the sliced tissue.

[00629] The compaction oligonucleotides will comprise a 5′ region that binds a first portion of the concatemer molecule and a 3′ region that binds a second portion of the same

concatemer molecule to pull together distal portions of the concatemer molecule causing compaction of the concatemer molecule to form a DNA nanoball.

[00630] *in situ* Sequencing Using 2-Stage Sequencing Method

[00631] The sliced tissue on the flow cells will be subjected to a two-stage sequencing method in a manner similar to the method described in Example 10 above. For example, the sliced tissue on the flow cells will be flowed with 1 μM of sequencing primer and 1X sequencing primer hybridization buffer, and incubated at 4 °C overnight, at 40 °C for 20 minutes, and at room temperature for 2 hours. The sliced tissue will be washed with 1X PBS. The universal sequencing primer will comprise SEQ ID NO 2: 5'-CATGGAATTTGCCATGGGTGGAA -3' (Primer 2). The sequencing primer hybridization buffer will comprise: 3X SSC, pH 7.0, and 300 nM compaction oligonucleotides. The sequencing primers will be designed to hybridize to the concatemer molecules inside the sliced tissue.

[00632] The universal compaction oligonucleotide will comprise SEQ ID NO 3: 5'- CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm -3' (Compaction oligonucleotide).

[00633] A two-stage sequencing method will be used to sequence the concatemers inside the sliced tissue. Each sequencing cycle will include a first stage and a second stage. The first stage will employ a first sequencing polymerase and fluorescently labeled multivalent molecules. The second stage will employ a second sequencing polymerase and non-labeled nucleotide analogs.

[00634] The first stage sequencing is conducted with fluorescently labeled multivalent molecules. An image of the cells on the flow cells is obtained prior to the start of sequencing. A solution of a first sequencing polymerase is flowed onto the sliced tissue and incubated at 42 °C for about 10 minutes to form polymerase complexes on the concatemers inside the sliced tissue. A trap reagent is flowed onto the cells. The trap reagent will include a mixture of fluorescently labeled multivalent molecules (e.g., about 40-100 nM) (see FIGs. 7-10) and a non-catalytic cation (e.g., strontium, barium and/or calcium). The mixture of fluorescently labeled multivalent molecules will include dATP, dGTP, dCTP and dUTP. The flow cells are incubated at 42 °C for about 10 minutes to permit the multivalent molecules to bind the polymerase complexes, without incorporation of the nucleotide moieties, and form avidity complexes on the concatemers inside the sliced tissue. The flow cells are washed at room temperature with a trap reagent that lacks the fluorescently labeled multivalent molecules. The flow cells are washed with an imaging buffer. An image is obtained using a fluorescent microscope, for example, 1 minute/FOV, 4 channels, 200 ms/frame, using Z-stack imaging. The

fluorescent microscope will be used to obtain two-dimensional scans of the sliced tissue of the fluorescent signals emitted from the labeled multivalent molecules during the first stage of sequencing. Four fluorescent channels will be used to conduct four color detection of the fluorescent signals emitted from the labeled multivalent molecules during the first stage of sequencing. The two-dimensional scans will include the entire sliced tissue or a portion of the sliced tissue. The two-dimensional scans will be conducted in a single z-plane to obtain images at a single depth of the sliced tissue. The two-dimensional scans will be conducted at two or more z-planes to obtain images at different depths of the sliced tissue. The sequences of the target DNA will be determined from the images. The multivalent molecules and first sequencing polymerases are removed by washing the sliced tissue twice at 42 °C with a removal reagent. The flow cells are washed four times at 52 °C with a wash buffer.

[00635] The second stage sequencing is conducted with non-labeled nucleotide analogs. A solution of a second sequencing polymerase, a mixture of non-labeled nucleotide analogs and a catalytic divalent cation (e.g., magnesium and/or manganese) is flowed onto the sliced tissue and incubated at 52 °C for about 2-5 minutes to permit incorporation of the non-labeled nucleotide analogs. The mixture of non-labeled nucleotide analogs will include 3'O-methylazido nucleotides with dATP, dGTP, dCTP and dTTP. The flow cells are washed twice with a removal reagent at 51 °C for about 20 seconds. The 3' blocking moieties are removed from the incorporated nucleotide by washing the sliced tissue twice with a cleaving reagent at 51 °C for about 40 seconds. The flow cells are washed twice with a wash buffer. One sequencing cycle included a first stage and a second stage.

[00636] The next sequencing cycle was conducted using the two-stage sequencing method described above. The cells on the flow cells will be subjected to 100 sequencing cycles which will generate a first plurality of sequencing read products. At each sequencing cycle, intensity measurement, background measurements, and phasing and pre-phasing measurements will be obtained (e.g., see the three graphs shown in FIG. 34).

Example 14. *in situ* Detection of Polypeptides and Sequencing in Tissue

[00637] Preparing antibody-oligonucleotide conjugates and antibody-oligonucleotide padlock probe complexes

[00638] Primary antibodies will be selected to selectively bind target polypeptides. In some embodiments, antibody-oligonucleotide conjugates will be prepared by conjugating the primary antibodies to oligonucleotides using disuccinimidyl suberate (DSS) linking chemistry. Briefly, amine-modified oligonucleotides are mixed with DSS linkers, acetonitrile and triethylamine, and the mixture is incubated at room temperature for 30 minutes. The mixture is ethanol precipitated by adding sodium acetate (pH 5.2) and pure ethanol, vortexed, and stored at

-80 °C for at least 2 hours. The mixture is centrifuged at 14,000 RPM at 4 °C for 30 minutes. The supernatant is removed and the remaining pellet is air dried at room temperature for at least 10 minutes. The dried pellet is resuspended in a resuspension buffer comprising HEPES (pH 7.4) and sodium chloride.

[00639] The amine-modified oligonucleotides will comprise single-stranded DNA. The amine-modified oligonucleotides will be designed to include a first sub-region that hybridizes to a left binding arm of a padlock probe, and a second sub-region that hybridizes to the right binding arm of the same padlock probe. The terminal 3' end of an amine-modified oligonucleotides will comprise a 3' OH moiety that will be extended in a polymerase-catalyzed extension reaction.

[00640] In some embodiments, a plurality of antibody-oligonucleotide conjugates comprises a mixture of antibody-oligonucleotide conjugates that will selectively bind to different target polypeptides. In some embodiments, the plurality of antibody-oligonucleotide conjugates comprises at least two antibody-oligonucleotide conjugates that will selectively bind their cognate target polypeptides.

[00641] In some embodiments, the plurality of antibody-oligonucleotide conjugates will include at least a first and a second sub-population of antibody-oligonucleotide conjugates. In some embodiments, the antibody portion of antibody-oligonucleotide conjugates of the first sub-population will selectively bind to a first target polypeptide and the oligonucleotide portion will selectively hybridize to a first padlock probe carrying a first barcode index sequence that identifies the first target polypeptide. In some embodiments, the antibody portion of antibody-oligonucleotide conjugates of the second sub-population will selectively bind to a second target polypeptide and the oligonucleotide portion will selectively hybridize to a second padlock probe carrying a second barcode index sequence that identifies the second target polypeptide.

[00642] In some embodiments, a plurality of antibody-oligonucleotide padlock probe complexes will be formed by hybridizing/binding a padlock probe to the oligonucleotide portion of an antibody-oligonucleotide conjugate. Individual padlock probes will include paired left and right binding arms and an intervening region between the left and right binding arms (e.g., a backbone region). The left and right binding arms will be designed to hybridize to an oligonucleotide that is conjugated to an antibody. Hybridization between the oligonucleotide and the left and right binding arms of a padlock probe will form a nick which is enzymatically ligatable. The intervening region between the left and right binding arms (e.g., a backbone region) will carrying a barcode index sequence that identifies the target polypeptide to which the primary antibody will bind. In some embodiments, the left binding arm of individual padlock

probes will be 10-30 nucleotides in length. In some embodiments, the right binding arm of individual padlock probes will be 10-30 nucleotides in length. In some embodiments, the left and right binding arms of individual padlock probes will be the same length or different lengths. In some embodiments, the intervening region of a padlock probe will also carry a universal binding site for a sequencing primer. In some embodiments, the intervening region will also carry a universal binding site for a compaction oligonucleotide.

[00643] Binding an antibody-oligonucleotide conjugate to a padlock probe will form an antibody-oligonucleotide-padlock probe complex. Individual antibody-oligonucleotide-padlock probe complexes comprise (i) a primary antibody that binds a target polypeptide, (ii) an oligonucleotide joined to the primary antibody, and (iii) a padlock probe hybridized to the oligonucleotide where the padlock probe carries a barcode index sequence which identifies the target polypeptide.

[00644] In some embodiments, a plurality of antibody-oligonucleotide padlock probe complexes comprises a mixture of antibody-oligonucleotide padlock probe complexes that will selectively bind to different target polypeptides. In some embodiments, the plurality of antibody-oligonucleotide padlock probe complexes comprises at least two antibody-oligonucleotide padlock probe complexes that will selectively bind their cognate target polypeptides.

[00645] In some embodiments, a plurality of antibody-oligonucleotide padlock probe complexes comprises at least a first and a second sub-population of antibody-oligonucleotide padlock probe complexes. Individual antibody-oligonucleotide-padlock probe complexes in a given sub-population will bind the same type of target polypeptide and carry the same barcode index sequence. Individual antibody-oligonucleotide-padlock probe complexes in the first sub-population will selectively bind a first target polypeptide. Individual antibody-oligonucleotide-padlock probe complexes in the second sub-population will selectively bind a second target polypeptide. The first and second target polypeptides will be different polypeptides. The plurality of antibody-oligonucleotide-padlock probe complexes will further comprise a third, fourth, fifth, sixth, seventh, eighth, ninth, tenth or more different antibody-oligonucleotide-padlock probe complexes, wherein individual antibody-oligonucleotide-padlock probe complexes in a given sub-population will bind a designated target polypeptide.

[00646] Examples of primary antibodies that will be used to prepare the antibody-oligonucleotide conjugates or the antibody-oligonucleotide padlock probe complexes include anti-alpha tubulin antibody (e.g., from Abcam, catalog No. ab52866), anti-histone antibody (e.g., from Abcam, catalog No. ab176842), anti-c-myc antibody (e.g., from Abcam, catalog No. ab185656), anti-HMOX-1 (heme oxygenase-1) antibody (e.g., from Abcam, catalog No.

ab189491), anti-claudin-18 antibody (e.g., from Abcam, catalog No. ab203563) and anti-GM130 (e.g., from Abcam, catalog No. ab52649).

[00647] Tissue Sample on Flow cells

[00648] Tissue samples will be prepared in a manner similar to the method described in Example 11 above. For example, fresh frozen tissue is resected from a human subject. The tissue will comprise a tumor sample and patient-matched tissue from the same human subject. The tissue sample is sliced into sections that are 5-10 μm thick and placed onto a sequencing flow cell. For example, the tissue sample will be sliced into sections that are 7-8 μm thick. The tissue slices are fixed and permeabilized. For example, the tissue slices are fixed using formaldehyde and washed with a buffer. In another example, the tissue slices are permeabilized using acetone. The tissue slices will remain on the flow cell during this workflow, including: contact with antibody-oligonucleotide conjugates, padlock probe hybridization, ligation reactions, rolling circle amplification and two-stage sequencing.

[00649] Forming Antibody-Oligonucleotide-Padlock Probe Complexes in Tissue Slices

[00650] A plurality of antibody-oligonucleotide conjugates is flowed onto the tissue slices and incubated for approximately one to three hours at room temperature so that the antibody-oligonucleotide conjugates bind their cognate target polypeptides. The plurality of antibody-oligonucleotide conjugates move inside the cells of the sliced tissue. The tissue slices are washed with a buffer.

[00651] A plurality of padlock probes is flowed onto the tissue slices in 1X padlock probe buffer and incubated for 15-30 minutes at 37 °C, or incubated for approximately one to three hours at room temperature, so that the padlock probes hybridize to their cognate oligonucleotide of an antibody-oligonucleotide conjugate. The plurality of padlock probes move inside the cells of the sliced tissue. Hybridization of the left and right binding arms of the padlock probes to their cognate oligonucleotide forms a nick which is enzymatically ligatable. The tissue slices are washed with 1X PBS. The 1X padlock probe hybridization buffer will comprise: 2X sodium-saline citrate buffer (SSC), pH 7.0, and 20%formamide.

[00652] Ligation Reaction

[00653] The ligation reaction is conducted by reacting the tissue slices with a ligation reaction mix for approximately 15-30 minutes at 37 °C. The ligation reaction mix will comprise a DNA ligase and a ligase buffer. The ligase buffer will comprise bovine serum albumin (BSA), Tris-acetate (pH 7.9), Mg-acetate, K-acetate, Tween 20, DTT, and ATP. After the ligation reaction the tissue slices will be washed with 1X PBS. In some embodiments, the DNA ligase will comprise a T4 DNA ligase or Hi-T4 DNA ligase.

[00654] Rolling Circle Amplification

[00655] A rolling circle amplification reaction will be conducted in a manner similar to the method described in Example 10 above. For example, the sliced tissue on the flow cells will be flowed with 1X RCA hybridization buffer and incubated at 55 °C for approximately 30 minutes. The sliced tissue will be washed with 1X RCA hybridization buffer at 55 °C for 10 minutes. The 1X RCA hybridization buffer will comprise: 2X SSC, pH 7.0, and 20% formamide.

[00656] The sliced tissue on the flow cell will be flowed with a first reaction mixture and incubated at 37 °C for approximately 30 minutes. The first reaction mixture will comprise a strand-displacing DNA polymerase, a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, and a viscosity agent. The first reaction mixture will have a pH at which the strand-displacing DNA polymerase is inactive. The strand-displacing DNA polymerase will comprise phi29, Taq polymerase, T4 polymerase, T7 polymerase, or a combination thereof.

[00657] Rolling circle amplification will be initiated by flowing onto the sliced tissue a second reaction mixture and incubated at 40 °C for approximately 2 hours. Alternatively, the sliced tissue will be incubated with the second reaction mixture overnight (e.g., at least 8 hours) at room temperature. The second reaction mixture will comprise a mixture of nucleotides (e.g., dATP, dCTP, dGTP and dTTP), magnesium ions, potassium ions, ammonium ions, a detergent, a reducing reagent, a viscosity agent and 300 nM compaction oligonucleotides. The second reaction mixture will have a pH at which the strand-displacing DNA polymerase is active. The rolling circle amplification reaction will be initiated at the terminal 3' end of the oligonucleotides that are conjugated to an antibody. The rolling circle amplification will generate concatemer molecules inside the sliced tissue. Individual concatemer molecules will remain joined (conjugated) to their cognate antibody. The concatemer molecules will serve as a surrogate for detecting the target polypeptides.

[00658] The compaction oligonucleotides will comprise a 5′ region that binds a first portion of the concatemer molecule and a 3′ region that binds a second portion of the same concatemer molecule to pull together distal portions of the concatemer molecule causing compaction of the concatemer molecule to form a DNA nanoball.

[00659] *in situ* Sequencing Using 2-Stage Sequencing Method

[00660] The sliced tissue on the flow cells will be subjected to a two-stage sequencing method in a manner similar to the method described in Example 10 above. For example, the sliced tissue on the flow cells will be flowed with 1 μM of sequencing primer and 1X sequencing primer hybridization buffer, and incubated at 4 °C overnight, at 40 °C for 20

minutes, and at room temperature for 2 hours. The sliced tissue will be washed with 1X PBS. The universal sequencing primer will comprise SEQ ID NO 2: 5'-CATGGAATTTGCCATGGGTGGAA -3' (Primer 2). The sequencing primer hybridization buffer will comprise: 3X SSC, pH 7.0, and 300 nM compaction oligonucleotides. The sequencing primers will be designed to hybridize to the concatemer molecules inside the sliced tissue.

[00661] The universal compaction oligonucleotide will comprise SEQ ID NO 3: 5'- CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm -3' (Compaction oligonucleotide).

[00662] A two-stage sequencing method will be used to sequence the concatemers inside the sliced tissue. Each sequencing cycle will include a first stage and a second stage. The first stage will employ a first sequencing polymerase and fluorescently labeled multivalent molecules. The second stage will employ a second sequencing polymerase and non-labeled nucleotide analogs.

[00663] The first stage sequencing is conducted with fluorescently labeled multivalent molecules. An image of the cells on the flow cells is obtained prior to the start of sequencing. A solution of a first sequencing polymerase is flowed onto the sliced tissue and incubated at 42 °C for about 10 minutes to form polymerase complexes on the concatemers inside the sliced tissue. A trap reagent is flowed onto the cells. The trap reagent will include a mixture of fluorescently labeled multivalent molecules (e.g., about 40-100 nM) (see FIGs. 7-10) and a non-catalytic cation (e.g., strontium, barium and/or calcium). The mixture of fluorescently labeled multivalent molecules will include dATP, dGTP, dCTP and dUTP. The flow cells are incubated at 42 °C for about 10 minutes to permit the multivalent molecules to bind the polymerase complexes, without incorporation of the nucleotide moieties, and form avidity complexes on the concatemers inside the sliced tissue. The flow cells are washed at room temperature with a trap reagent that lacks the fluorescently labeled multivalent molecules. The flow cells are washed with an imaging buffer. An image is obtained using a fluorescent microscope, for example, 1 minute/FOV, 4 channels, 200 ms/frame, using Z-stack imaging. The fluorescent microscope will be used to obtain two-dimensional scans of the sliced tissue of the fluorescent signals emitted from the labeled multivalent molecules during the first stage of sequencing. Four fluorescent channels will be used to conduct four color detection of the fluorescent signals emitted from the labeled multivalent molecules during the first stage of sequencing. The two-dimensional scans will include the entire sliced tissue or a portion of the sliced tissue. The two-dimensional scans will be conducted in a single z-plane to obtain images at a single depth of the sliced tissue. The two-dimensional scans will be conducted at two or

more z-planes to obtain images at different depths of the sliced tissue. The sequences of the concatemer molecules will be determined from the images. The sequences of the concatemer molecules will serve as a surrogate for detecting and identifying the target polypeptides. The multivalent molecules and first sequencing polymerases are removed by washing the sliced tissue twice at 42 °C with a removal reagent. The flow cells are washed four times at 52 °C with a wash buffer.

[00664] The second stage sequencing is conducted with non-labeled nucleotide analogs. A solution of a second sequencing polymerase, a mixture of non-labeled nucleotide analogs and a catalytic divalent cation (e.g., magnesium and/or manganese) is flowed onto the sliced tissue and incubated at 52 °C for about 2-5 minutes to permit incorporation of the non-labeled nucleotide analogs. The mixture of non-labeled nucleotide analogs will include 3'O-methylazido nucleotides with dATP, dGTP, dCTP and dTTP. The flow cells are washed twice with a removal reagent at 51 °C for about 20 seconds. The 3' blocking moieties are removed from the incorporated nucleotide by washing the sliced tissue twice with a cleaving reagent at 51 °C for about 40 seconds. The flow cells are washed twice with a wash buffer. One sequencing cycle included a first stage and a second stage.

[00665] The next sequencing cycle was conducted using the two-stage sequencing method described above. The cells on the flow cells will be subjected to 100 sequencing cycles which will generate a first plurality of sequencing read products. At each sequencing cycle, intensity measurement, background measurements, and phasing and pre-phasing measurements will be obtained (e.g., see the three graphs shown in FIG. 34). Table 1 provides sequences for use in the methods, systems, and kits, as described herein.

Table 1: Table of Sequences

SEQ ID NO:	Sequence	Description
1	GGCTCCACTAAATAGACGCAGTCAGTAA	Primer 1
2	CATGGAATTTGCCATGGGTGGA	Primer 2
3	CAGCCGCATCTTCTTTTGCTTTCAGCCGCATCTTCTTTTGCUmUmUm	Compaction oligonucleotide
4	/5phos/ACGACCAAATCCGTTGACTCCGACCTTCACCTTCCCCATGGCAGCCGCATCTTCTTTTTCTATGACATGAGGACTTTTTACTGACTGCGTCTATTTAGTGGAGCCTCCCGTTCTCAGCCTTGACGGTGCCA	Padlock probe

	TGGAATTTGCCATGGGTGGAATCATAT TGGA	

Table 1 provides nucleotide sequences associated with the methods, systems, and kits as described herein. Um means the sequence includes a 2'-O-methyluridine at the listed position. /5phos/ means a 5'-phosphoyrl group is included at the listed position.

EMBODIMENTS

[00666] The following are examples of embodiments of the present disclosure and are not meant to be limiting in any way.

1. A method for identifying a ribonucleic acid (RNA) *in situ*, the method comprising:

(a) providing a cell or a tissue that has been permeabilized, wherein said cell or said tissue comprises said RNA;

(b) directing a probe to said cell or tissue, wherein said probe comprises (i) an RNA-binding sequence and (ii) an index sequence, under conditions sufficient to couple said RNA-binding sequence to at least a portion of said RNA;

(c) amplifying said index sequence to generate one or more amplicons, wherein said one or more amplicons comprises multiple copies of said index sequence or derivative thereof;

(d) directing a primer nucleic acid sequence to said cell or said tissue under conditions sufficient to bind said primer nucleic acid sequence to at least a portion of said one or more amplicons to produce a primed index sequence;

(e) directing said cell or said tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence or said derivative thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating said one or more nucleotide moieties into said primer nucleic acid sequence; and

(f) detecting said binding complex within said cell or tissue, thereby identifying said RNA.

2. The method of embodiment 1, wherein said method does not comprise performing a reverse transcription reaction of said RNA.

3. The method of any one of embodiments 1 or 2, further comprising providing said cell in (a).

4. The method of embodiment 1, wherein said cell is a cancer cell.

5. The method of any one of embodiments 1 or 2, further comprising providing said tissue in (a).

6. The method of embodiment 5, wherein said tissue is a formalin-fixed paraffin embedded tissue.

7. The method of embodiment 5, wherein said tissue is a fresh-frozen tissue.

8. The method of any one of embodiments 1 or 2, wherein said cell or said tissue is extracted from a subject.

9. The method of embodiment 8, wherein said subject is a human subject.

10. The method of any one of embodiments 1-9, wherein said index sequence comprises a nucleic acid.

11. The method of embodiment 10, wherein said index sequence is 4-50 nucleotides in length.

12. The method of any one of embodiments 1-11, wherein at least a portion of said index sequence or derivative thereof recognizes said primer nucleic acid sequence.

13. The method of any one of embodiments 1-11, wherein said primer nucleic acid sequence comprises a blocking group.

14. The method of embodiment 13, wherein said blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence.

15. The method of embodiment 13, wherein said blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group.

16. The method of any one of embodiments 1-15, wherein said primer nucleic acid sequence does not include a 3' hydroxyl group.

17. The method of any one of embodiments 1-16, wherein said primer nucleic acid sequence binds to at least a portion of said index sequence.

18. The method of any one of embodiments 1-17, wherein said primer nucleic acid sequence binds to at least a portion of a reverse complement of said index sequence.

19. The method of any one of embodiments 1-18, wherein said primed index sequence comprises a double-stranded nucleic acid.

20. The method of any one of embodiments 1-19, wherein said primed index sequence comprises a single-stranded nucleic acid.

21. The method of any one of embodiments 1-20, wherein said cell or said tissue is immobilized on a surface of a substrate.

22. The method of embodiment 21, wherein said surface of said substrate has a water contact angle of less than or equal to 45 degrees.

23. The method of embodiment 21, wherein said surface of said substrate is an interior surface of a flow cell.

24. The method of embodiment 23, wherein said interior surface of said flow cell comprises one or more hydrophilic polymer layers.

25. The method of embodiment 24, wherein said one or more hydrophilic polymer layers comprises a branched polymer.

26. The method of embodiment 24, wherein said one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol.

27. The method of any one of embodiments 1-26, wherein said nucleotide conjugate comprises a common core.

28. The method of embodiment 27, wherein said one or more nucleotide moieties are coupled to said common core.

29. The method of embodiment 27 or 28, wherein said common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot.

30. The method of embodiment 29, wherein said polymer comprises a protein.

31. The method of any one of embodiments 1-30, wherein in (e) said nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of said plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other.

32. The method of embodiment 31, wherein said different types of said nucleotide moieties comprise at least three different types of said nucleotide moieties.

33. The method of any one of embodiments 1-32, wherein said one or more nucleotide moieties do not comprise a blocking group coupled thereto.

34. The method of any one of embodiments 1-33, wherein said detecting in (f) comprises imaging said cell or said tissue using one or more image sensors.

35. The method of embodiment 34, wherein said one or more image sensors comprises a photodetector array.

36. The method of any one of embodiments 1-35, further comprising washing said cell or said tissue following (f) to remove said nucleotide conjugate from said cell or said tissue.

37. The method of embodiment 36, further comprising repeating (e) and (f) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in said primed index sequence or derivative thereof.

38. The method of any one of embodiments 1-37, wherein said detectable label comprises a fluorophore.

39. The method of any one of embodiments 1-38, wherein said detectable label comprises a quantum dot.

40. The method of any one of embodiments 1-39, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises an incubation temperature, wherein said incubation temperature is about 20-45°C.

41. The method of any one of embodiments 1-40, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises a salt.

42. The method of any one of embodiments 1-41, wherein said RNA is a messenger RNA.

43. The method of any one of embodiments 1-41, wherein said RNA is a micro RNA.

44. The method of any one of embodiments 1-41, wherein said RNA is a ribosomal RNA.

45. The method of any one of embodiments 1-41, wherein said RNA is a transfer RNA.

46. The method of any one of embodiments 1-41, wherein said RNA is a transfer-messenger RNA.

47. The method of any one of embodiments 1-46, wherein said RNA comprises a modification.

48. The method of embodiment 47, wherein said modification is a 5' phosphorylation.

49. The method of embodiment 47, wherein said modification is an inter-nucleotide linkage.

50. The method of any one of embodiments 1-49, wherein said probe comprises nucleic acid.

51. The method of embodiment 50, wherein said nucleic acid comprises deoxyribonucleic acid (DNA).

52. The method of embodiment 50, wherein said nucleic acid comprises RNA.

53. The method of any one of embodiments 1-49, wherein said RNA-binding sequence comprises nucleic acid.

54. The method of embodiment 53, wherein said nucleic acid comprises deoxyribonucleic acid (DNA).

55. The method of embodiment 53, wherein said nucleic acid comprises RNA.

56. The method of any one of embodiments 1-55, wherein said probe is a padlock probe.

57. The method of embodiment 56, further comprising, before (c), ligating a first end of said probe to a second end of said probe using a ligase.

58. The method of embodiment 57, wherein said ligase is a T4 ligase.

59. The method of embodiment 57, wherein said ligase is a SplintR ligase.

60. The method of any one of embodiments 1-59, wherein performing said amplification reaction in (c) comprises performing a rolling circle amplification reaction using a polymerizing enzyme.

61. The method of embodiment 60, wherein said polymerizing enzyme is a phi29 polymerase.

62. The method of any one of embodiments 1-61, wherein said cell or said tissue is permeabilized using a solvent.

63. The method of embodiment 62, wherein said solvent is methanol.

64. The method of embodiment 62, wherein said solvent is acetone.

65. The method of any one of embodiments 1-61, wherein said cell or said tissue is permeabilized using a detergent.

66. The method of embodiment 65, wherein said detergent is saponin.

67. The method of embodiment 65, wherein said detergent is Triton X-100.

68. A method for identifying a polypeptide *in situ*, the method comprising:

(a) providing a cell or a tissue that has been permeabilized, wherein said cell or said tissue comprises said polypeptide;

(b) directing a polypeptide-binding probe to said cell or tissue, wherein said polypeptide-binding probe comprises a polypeptide binding moiety and an oligonucleotide coupled thereto, under conditions sufficient to couple said polypeptide-binding probe to said polypeptide;

(c) directing a probe to said cell or said tissue, wherein said with a probe comprising (i) an oligonucleotide-binding probe and (ii) an index sequence under conditions sufficient to couple said oligonucleotide-binding probe to at least a portion of said oligonucleotide of said polypeptide-binding probe;

(d) amplifying said index sequence to generate one or more amplicons, wherein said one or more amplicons comprises multiple copies of said index sequence or derivative thereof;

(e) directing a primer nucleic acid sequence to said cell or said tissue under conditions sufficient to bind said primer nucleic acid sequence to at least a portion of said one or more amplicons to produce a primed index sequence;

(f) directing said cell or said tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said

primed index sequence or said one or more amplicons thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating said one or more nucleotide moieties into said primer nucleic acid sequence; and

(g) detecting said binding complex within said cell or tissue, thereby identifying said polypeptide.

69. The method of embodiment 68, further comprising providing said cell in (a).

70. The method of embodiment 69, wherein said cell is a cancer cell.

71. The method of embodiment 68, further comprising providing said tissue in (a).

72. The method of embodiment 71, wherein said tissue is a formalin-fixed paraffin embedded tissue.

73. The method of embodiment71, wherein said tissue is a fresh-frozen tissue.

74. The method of embodiment 68, wherein said cell or said tissue is extracted from a subject.

75. The method of embodiment 74, wherein said subject is a human subject.

76. The method of any one of embodiments 68-75, wherein said index sequence comprises a nucleic acid.

77. The method of embodiment 76, wherein said index sequence is 4-50 nucleotides in length.

78. The method of any one of embodiments 68-77, wherein at least a portion of said index sequence or derivative thereof recognizes said primer nucleic acid sequence.

79. The method of any one of embodiments 68-78, wherein said primer nucleic acid sequence comprises a blocking group.

80. The method of embodiment 79, wherein said blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence.

81. The method of embodiment 79, wherein said blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group.

82. The method of any one of embodiments 68-81, wherein said primer nucleic acid sequence does not include a 3' hydroxyl group.

83. The method of any one of embodiments 68-82, wherein said primer nucleic acid sequence binds to at least a portion of said index sequence.

84. The method of any one of embodiments 68-83, wherein said primer nucleic acid sequence binds to at least a portion of a reverse complement of said index sequence.

85. The method of any one of embodiments 68-84, wherein said primed index sequence comprises a double-stranded nucleic acid.

86. The method of any one of embodiments 68-85, wherein said primed index sequence comprises a single-stranded nucleic acid.

87. The method of any one of embodiments 68-86, wherein said cell or said tissue is immobilized on a surface of a substrate.

88. The method of embodiment 87, wherein said surface of said substrate has a water contact angle of less than or equal to 45 degrees.

89. The method of embodiment 87, wherein said surface of said substrate is an interior surface of a flow cell.

90. The method of embodiment 89, wherein said interior surface of said flow cell comprises one or more hydrophilic polymer layers.

91. The method of embodiment 90, wherein said one or more hydrophilic polymer layers comprises a branched polymer.

92. The method of embodiment 90, wherein said one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol.

93. The method of any one of embodiments 68-92, wherein said nucleotide conjugate comprises a common core.

94. The method of embodiment 93, wherein said one or more nucleotide moieties are coupled to said common core.

95. The method of embodiment 93 or 94, wherein said common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot.

96. The method of embodiment 95, wherein said polymer comprises a protein.

97. The method of any one of embodiments 68-96, wherein in (e) said nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of said plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other.

98. The method of embodiment 97, wherein said different types of said nucleotide moieties comprise at least three different types of said nucleotide moieties.

99. The method of any one of embodiments 68-98, wherein said one or more nucleotide moieties do not comprise a blocking group coupled thereto.

100. The method of any one of embodiments 68-99, wherein said detecting in (f) comprises imaging said cell or said tissue using one or more image sensors.

101. The method of embodiment 100, wherein said one or more image sensors comprises a photodetector array.

102. The method of any one of embodiments 68-101, further comprising washing said cell or said tissue following (f) to remove said nucleotide conjugate from said cell or said tissue.

103. The method of embodiment 102, further comprising repeating (e) and (f) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in said primed index sequence or derivative thereof.

104. The method of any one of embodiments 68-103, wherein said detectable label comprises a fluorophore.

105. The method of any one of embodiments 68-104, wherein said detectable label comprises a quantum dot.

106. The method of any one of embodiments 68-105, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises an incubation temperature, wherein said incubation temperature is about 20-45°C.

107. The method of any one of embodiments 68-106, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises a salt.

108. The method of any one of embodiment 68-107, wherein said polypeptide comprises a protein.

109. The method of embodiment 108, wherein said protein is a cell-surface protein.

110. The method of embodiment 108, wherein said protein is an intracellular protein.

111. The method of embodiment 108, wherein said protein is a signaling protein.

112. The method of any one of embodiments 68-111, wherein said polypeptide comprises a protein complex.

113. The method of any one of embodiments 68-112, wherein said polypeptide binding moiety comprises an antibody or antibody fragment.

114. The method of any one of embodiments 68-112, wherein said polypeptide binding moiety comprises an aptamer.

115. The method of any one of embodiments 68-112, wherein said polypeptide binding moiety comprises a nanobody.

116. The method of any one of embodiments 68-115, wherein said oligonucleotide comprises DNA.

117. The method of any one of embodiments 68-116, wherein said oligonucleotide is 4-50 nucleotides in length.

118. The method of any one of embodiments 68-117, wherein said polypeptide binding moiety is covalently attached to said oligonucleotide.

119. The method of any one of embodiments 68-117, wherein said polypeptide binding moiety is non-covalently attached to said oligonucleotide.

120. The method of any one of embodiments 68-119, wherein said probe comprises nucleic acid.

121. The method of embodiment 120, wherein said nucleic acid comprises DNA.

122. The method of embodiment 120, wherein said nucleic acid comprises RNA.

123. The method of any one of embodiments 68-122, wherein said probe is a padlock probe.

124. The method of embodiment 123, further comprising, before (d), ligating a first end of said probe to a second end of said probe using a ligase.

125. The method of embodiment 124, wherein said ligase is a T4 ligase.

126. The method of embodiment 124, wherein said ligase is a SplintR ligase.

127. The method of any one of embodiments 68-126, wherein performing said amplification reaction in (d) comprises performing a rolling circle amplification reaction using a polymerizing enzyme.

128. The method of embodiment 127, wherein said polymerizing enzyme is a phi29 polymerase.

129. The method of any one of embodiments 68-128, wherein said cell or said tissue is permeabilized using a solvent.

130. The method of embodiment 129, wherein said solvent is methanol.

131. The method of embodiment 129, wherein said solvent is acetone.

132. The method of any one of embodiments 68-131, wherein said cell or said tissue is permeabilized using a detergent.

133. The method of embodiment 132, wherein said detergent is saponin.

134. The method of embodiment 132, wherein said detergent is Triton X-100.

135. A method for identifying a DNA *in situ*, the method comprising:

(a) providing a cell or a tissue that has been permeabilized, wherein said cell or said tissue comprises said DNA, where said DNA is not cDNA;

(b) directing a probe to said cell or said tissue, wherein said probe comprises a (i) DNA-binding sequence and (ii) an index sequence, under conditions sufficient to couple said DNA-binding sequence to at least a portion of said DNA;

(c) amplifying said index sequence to generate one or more amplicons, wherein said one or more amplicons comprises multiple copies of said index sequence or derivative thereof;

(d) directing a primer nucleic acid sequence to said cell or said tissue under conditions sufficient to bind said primer nucleic acid sequence to at least a portion of said one or more amplicons to produce a primed index sequence;

(e) directing said cell or said tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence or said derivatives thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating said one or more nucleotide moieties into said primer nucleic acid sequence; and

(f) detecting said binding complex within said cell or tissue, thereby identifying said DNA.

136. The method of embodiment 135, further comprising providing said cell in (a).

137. The method of embodiment 136, wherein said cell is a cancer cell.

138. The method of embodiment 135, further comprising providing said tissue in (a).

139. The method of embodiment 138, wherein said tissue is a formalin-fixed paraffin embedded tissue.

140. The method of embodiment 138, wherein said tissue is a fresh-frozen tissue.

141. The method of embodiment 135, wherein said cell or said tissue is extracted from a subject.

142. The method of embodiment 141, wherein said subject is a human subject.

143. The method of any one of embodiments 135-142, wherein said index sequence comprises a nucleic acid.

144. The method of embodiment 143, wherein said index sequence is 4-50 nucleotides in length.

145. The method of any one of embodiments 135-144, wherein at least a portion of said index sequence or derivative thereof recognizes said primer nucleic acid sequence.

146. The method of any one of embodiments 135-145, wherein said primer nucleic acid sequence comprises a blocking group.

147. The method of embodiment 146, wherein said blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence.

148. The method of embodiment146, wherein said blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group.

149. The method of any one of embodiments 135-148, wherein said primer nucleic acid sequence does not include a 3' hydroxyl group.

150. The method of any one of embodiments 135-149, wherein said primer nucleic acid sequence binds to at least a portion of said index sequence.

151. The method of any one of embodiments 135-150, wherein said primer nucleic acid sequence binds to at least a portion of a reverse complement of said index sequence.

152. The method of any one of embodiments 135-151, wherein said primed index sequence comprises a double-stranded nucleic acid.

153. The method of any one of embodiments 135-152, wherein said primed index sequence comprises a single-stranded nucleic acid.

154. The method of any one of embodiments 135-153, wherein said cell or said tissue is immobilized on a surface of a substrate.

155. The method of embodiment 154, wherein said surface of said substrate has a water contact angle of less than or equal to 45 degrees.

156. The method of embodiment 154, wherein said surface of said substrate is an interior surface of a flow cell.

157. The method of embodiment 156, wherein said interior surface of said flow cell comprises one or more hydrophilic polymer layers.

158. The method of embodiment 157, wherein said one or more hydrophilic polymer layers comprises a branched polymer.

159. The method of embodiment 157, wherein said one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol.

160. The method of any one of embodiments 135-159, wherein said nucleotide conjugate comprises a common core.

161. The method of embodiment 160, wherein said one or more nucleotide moieties are coupled to said common core.

162. The method of embodiment 160 or 161, wherein said common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot.

163. The method of embodiment 162, wherein said polymer comprises a protein.

164. The method of any one of embodiments 135-163, wherein in (e) said nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of said plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other.

165. The method of embodiment 164, wherein said different types of said nucleotide moieties comprise at least three different types of said nucleotide moieties.

166. The method of any one of embodiments 135-165, wherein said one or more nucleotide moieties do not comprise a blocking group coupled thereto.

167. The method of any one of embodiments 135-166, wherein said detecting in (f) comprises imaging said cell or said tissue using one or more image sensors.

168. The method of embodiment 167, wherein said one or more image sensors comprises a photodetector array.

169. The method of any one of embodiments 135-168, further comprising washing said cell or said tissue following (f) to remove said nucleotide conjugate from said cell or said tissue.

170. The method of embodiment 169, further comprising repeating (e) and (f) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in said primed index sequence or derivative thereof.

171. The method of any one of embodiments 135-170, wherein said detectable label comprises a fluorophore.

172. The method of any one of embodiments 135-171, wherein said detectable label comprises a quantum dot.

173. The method of any one of embodiments 135-172, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises an incubation temperature, wherein said incubation temperature is about 20-45°C.

174. The method of any one of embodiments 135-173, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises a salt.

175. The method of any one of embodiments 135-174, wherein said DNA comprises a modification.

176. The method of embodiment 175, wherein said modification is a methyl modification.

177. The method of embodiment 175, wherein said modification is a phosphoryl modification.

178. The method of any one of embodiments 135-177, wherein said probe comprises nucleic acid.

179. The method of embodiment 178, wherein said nucleic acid comprises DNA.

180. The method of embodiment 178, wherein said nucleic acid comprises RNA.

181. The method of any one of embodiments 135-180, wherein said DNA-binding sequence comprises nucleic acid.

182. The method of embodiment 181, wherein said nucleic acid comprises DNA.

183. The method of embodiment 181, wherein said nucleic acid comprises RNA.

184. The method of any one of embodiments 135-183, wherein said probe is a padlock probe.

185. The method of embodiment 184, further comprising before (c), ligating a first end of said probe to a second end of said probe using a ligase.

186. The method of embodiment 185, wherein said ligase is a T4 ligase.

187. The method of embodiment 185, wherein said ligase is a SplintR ligase.

188. The method of any one of embodiments 135-187, wherein performing said amplification reaction in (c) comprises performing a rolling circle amplification reaction using a polymerizing enzyme.

189. The method of embodiment 188, wherein said polymerizing enzyme is a phi29 polymerase.

190. The method of any one of embodiments 135-189, wherein said cell or said tissue is permeabilized using a solvent.

191. The method of embodiment 190, wherein said solvent is methanol.

192. The method of embodiment 190, wherein said solvent is acetone.

193. The method of any one of embodiments 135-192, wherein said cell or said tissue is permeabilized using a detergent.

194. The method of embodiment 193, wherein said detergent is saponin.

195. The method of embodiment 193, wherein said detergent is Triton X-100.

196. A method for identifying a ribonucleic acid in situ, the method comprising:

(a) providing a cell or tissue that has been permeabilized, wherein said cell or said tissue comprises said ribonucleic acid;

(b) directing a nucleic acid molecule to said cell or said tissue, wherein said nucleic acid molecule comprises a target-binding sequence, under conditions sufficient to couple said target-binding sequence to at least a portion of said ribonucleic acid, wherein said nucleic acid molecule further comprises an index sequence;

(c) directing a primer nucleic acid sequence to said cell or said tissue under conditions sufficient to bind said primer nucleic acid sequence to at least a portion of said one or more amplicons to produce a primed index sequence;

(d) directing said cell or said tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence or derivative thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence;

(e) detecting said binding complex within said cell or tissue, thereby identifying said ribonucleic acid.

197. The method of embodiment 196, further comprising providing said cell in (a).

198. The method of embodiment 197, wherein said cell is a cancer cell.

199. The method of embodiment 196, further comprising providing said tissue in (a).

200. The method of embodiment 199, wherein said tissue is a formalin-fixed paraffin embedded tissue.

201. The method of embodiment 199, wherein said tissue is a fresh-frozen tissue.

202. The method of embodiment 196, wherein said cell or said tissue is extracted from a subject.

203. The method of embodiment 202, wherein said subject is a human subject.

204. The method of any one of embodiments 196-203, wherein said ribonucleic acid is a messenger ribonucleic acid.

205. The method of any one of embodiments 196-204, wherein said ribonucleic acid is a micro ribonucleic acid.

206. The method of any one of embodiments 196-204, wherein said ribonucleic acid is a ribosomal ribonucleic acid.

207. The method of any one of embodiments 196-204, wherein said ribonucleic acid is a transfer ribonucleic acid.

208. The method of any one of embodiments 196-204, wherein said ribonucleic acid is a transfer-messenger ribonucleic acid.

209. The method of any one of embodiments 196-208, wherein said ribonucleic acid comprises a modification.

210. The method of embodiment 209, wherein said modification comprises a methyl group.

211. The method of any one of embodiments 196-210, wherein said nucleic acid molecule is a deoxyribonucleic acid.

212. The method of any one of embodiments 196-211, wherein said nucleic acid molecule is about 10-80 nucleotides in length.

213. The method of any one of embodiments 196-211, wherein said nucleic acid molecule is about 20-60 nucleotides in length.

214. The method of any one of embodiments 196-211, wherein said nucleic acid molecule is about 30-50 nucleotides in length.

215. The method of any one of embodiments 196-211, wherein said nucleic acid molecule comprises a modification.

216. The method of embodiment 215, wherein said modification is a 5' phosphorylation.

217. The method of embodiment 215, wherein said modification is an inter-nucleotide linkage.

218. The method of any one of embodiments 196-217, wherein said target-binding sequence is about 5-25 nucleotides in length.

219. The method of any one of embodiments 196-218, further comprising introducing an index-binding molecule under conditions sufficient to couple said index-binding molecule to said index sequence.

220. The method of embodiment 219, wherein said index-binding molecule comprises a nucleic acid.

221. The method of embodiment 220, wherein said index-binding molecule is a padlock probe.

222. The method of embodiment 220, further comprising ligating said index-binding molecule using a ligase.

223. The method of embodiment 222, wherein said ligase is a T4 ligase.

224. The method of embodiment 222, wherein said ligase is a SplintR ligase.

225. The method of any one of embodiments 196-224, further comprising amplifying said index sequence.

226. The method of embodiment 225, wherein said amplifying comprises performing rolling circle amplification using a polymerizing enzyme.

227. The method of embodiment 226, wherein said polymerizing enzyme is a phi29 polymerase.

228. The method of any one of embodiments 225-227, wherein said amplifying generates multiple copies of said index sequence or a reverse complement of said index sequence.

229. The method of any one of embodiments 196-228, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said ribonucleic acid in (b) comprises a salt.

230. The method of any one of embodiments 196-229, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said ribonucleic acid

in (b) comprises an incubation temperature, wherein said incubation temperature is about 20-45°C.

231. The method of any one of embodiments 196-229, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said ribonucleic acid in (b) comprises a chaotropic reagent.

232. The method of any one of embodiments 196-231, wherein said index sequence is uniquely associated with said ribonucleic acid.

233. The method of any one of embodiments 196-231, wherein said index sequence is associated with said ribonucleic acid and at least one other analyte.

234. The method of embodiment 233, wherein said other analyte is a nucleic acid.

235. The method of embodiment 234, wherein said nucleic acid is a ribonucleic acid.

236. The method of embodiment 234, wherein said nucleic acid is a deoxyribonucleic acid.

237. The method of embodiment 233, wherein said other analyte is a polypeptide.

238. The method of embodiment 237, wherein said polypeptide is a protein.

239. The method of any one of embodiments 196-238, wherein said index sequence is 4-50 nucleotides in length.

240. The method of any one of embodiments 196-239, wherein at least a portion of said index sequence recognizes said primer nucleic acid sequence.

241. The method of any one of embodiments 196-240, wherein said primer nucleic acid sequence comprises a blocking group.

242. The method of embodiment 241, wherein said blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence.

243. The method of embodiment 241, wherein said blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group.

244. The method of any one of embodiments 196-243, wherein said primer nucleic acid sequence does not include a 3' hydroxyl group.

245. The method of any one of embodiments 196-244, wherein said primer nucleic acid sequence binds to at least a portion of said index sequence.

246. The method of any one of embodiments 196-245, wherein said primer nucleic acid sequence binds to at least a portion of the reverse complement of said index sequence.

247. The method of any one of embodiments 196-246, wherein said primed index sequence comprises a double-stranded nucleic acid.

248. The method of any one of embodiments 196-247, wherein said primed index sequence comprises a single-stranded nucleic acid.

249. The method of any one of embodiments 196-248, wherein said cell or said tissue is immobilized on a surface of a substrate.

250. The method of embodiment 249, wherein said surface has a water contact angle of less than or equal to 45 degrees.

251. The method of embodiment 249, wherein said surface is an interior surface of a flow cell.

252. The method of embodiment 251, wherein said interior surface of said flow cell comprises one or more hydrophilic polymer layers.

253. The method of embodiment 252, wherein said one or more hydrophilic polymer layers comprises a branched polymer.

254. The method of embodiment 253, wherein said one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol.

255. The method of any one of embodiments 196-254, wherein said nucleotide conjugate comprises a common core.

256. The method of embodiment 255, wherein said one or more nucleotide moieties are coupled to said common core.

257. The method of embodiment 255 or 256, wherein said common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot.

258. The method of embodiment 257, wherein said polymer comprises a protein.

259. The method of any one of embodiments 196-258, wherein in (d) said nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of said plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other.

260. The method of embodiment 259, wherein said different types of said nucleotide moieties comprise at least three different types of said nucleotide moieties.

261. The method of any one of embodiments 196-260, wherein said one or more nucleotide moieties do not comprise a blocking group coupled thereto.

262. The method of any one of embodiments 196-261, wherein said detecting in (e) comprises imaging said cell or said tissue using one or more image sensors.

263. The method of embodiment 262, wherein said one or more image sensors comprises a photodetector array.

264. The method of any one of embodiments 196-263, further comprising washing said cell or said tissue following (e) to remove said nucleotide conjugate from said cell or said tissue.

265. The method of embodiment 264, further comprising repeating steps (d) and (e) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in said primed index sequence or derivative thereof.

266. The method of any one of embodiments 196-265, wherein said detectable label comprises a fluorophore.

267. The method of any one of embodiments 196-266, wherein said detectable label comprises a quantum dot.

268. A system for identifying a RNA *in situ*, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of any one of clams 1-67.

269. The system of embodiment 268, further comprising said RNA.

270. A system for identifying a polypeptide *in situ*, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of any one of clams 68-134.

271. The system of embodiment 270, further comprising said polypeptide.

272. The system of any one of embodiments 270 or 271, further comprising said polypeptide-binding probe.

273. A system for identifying a DNA *in situ*, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of any one of clams 135-195.

274. The system of embodiment 273, further comprising said DNA.

275. The system of any one of embodiments 268-274, further comprising a cell.

276. The system of embodiment 275, wherein said cell is a cancer cell.

277. The system of any one of embodiments 268-274, further comprising a tissue.

278. The system of embodiment 277, wherein said tissue is a formalin-fixed paraffin embedded tissue.

279. The system of embodiment 277, wherein said tissue is a fresh-frozen tissue.

280. The system of any one of embodiments 268-274, wherein said cell or tissue is extracted from a subject.

281. The system of embodiment 280, wherein said subject is a human subject.

282. The system of any one of embodiments 268-281, further comprising said probe.

283. The system of embodiment 282, wherein said probe is a padlock probe.

284. The system of embodiment 283, wherein said padlock probe comprises a nucleic acid.

285. The system of any one of embodiments 268-284, further comprising said primer nucleic acid molecule sequence.

286. The system of any one of embodiments 268-285, further comprising a ligase.

287. The system of embodiment 286, wherein said ligase is a T4 ligase.

288. The system of embodiment 286, wherein said ligase is a SplintR ligase.

289. The system of any one of embodiments 268-288, further comprising a polymerizing enzyme.

290. The system of embodiment 289, wherein said polymerizing enzyme is a phi29 polymerizing enzyme.

291. The system of any one of embodiments 268-290, further comprising a salt.

292. The system of embodiment 291, wherein said salt comprises NaCl, $MgCl_2$, or $CaCl_2$.

293. The system of any one of embodiments 268-292, further comprising a solvent.

294. The system of embodiment 293, wherein said solvent is a polar solvent.

295. The system of embodiment 293, wherein said solvent is methanol.

296. The system of embodiment 293, wherein said solvent is acetone.

297. The system of any one of embodiments 268-296, further comprising a detergent.

298. The system of embodiment 297, wherein said detergent is saponin.

299. The system of embodiment 297, wherein said detergent is Triton X-100.

300. The system of any one of embodiments 268-299, further comprising a substrate.

301. The system of embodiment 300, wherein said substrate is an interior surface of a flow cell.

302. The system of any one of embodiments 268-301, wherein said one or more computer processors comprises a central processing unit.

303. The system of any one of embodiments 268-302, wherein said one or more computer processors comprises a graphic processing unit.

304. The system of any one of embodiments 268-303, wherein said one or more computer processors comprises a general-purpose processing unit.

305. The system of any one of embodiments 268-304, wherein said one or more processors comprises a field-programmable gate array.

306. The system of any one of embodiments 268-305, further comprising a fluidic system.

307. The system of embodiment 306, wherein said fluidic system is configured to provide fluid flow control capability for delivering reagents to a sample in a flow cell.

308. A system for identifying a ribonucleic acid in situ, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of any one of clams 196-267.

309. The system of embodiment 308, further comprising said cell.

310. The system of embodiment 309, wherein said cell is a cancer cell.

311. The system of embodiment 308, further comprising said tissue.

312. The system of embodiment 311, wherein said tissue is a formalin-fixed paraffin embedded tissue.

313. The system of embodiment 312, wherein said tissue is a fresh-frozen tissue.

314. The system of embodiment 308, wherein said cell or said tissue is extracted from a subject.

315. The system of embodiment 314, wherein said subject is a human subject.

316. The system of any one of embodiments 308-315, further comprising said ribonucleic acid.

317. The system of any one of embodiments 308-316, further comprising said nucleic acid molecule.

318. The system of any one of embodiments 308-317, further comprising a primer nucleic acid sequence.

319. The system of any one of embodiment 308-318, further comprising an index-binding molecule.

320. The system of embodiment 319, wherein said index-binding molecule comprises a nucleic acid.

321. The system of embodiment 319, wherein said index-binding molecule is a padlock probe.

322. The system of embodiment 319, wherein said index-binding molecule is configured to bind to said index sequence.

323. The system of any one of embodiments 308-322, further comprising a ligase.

324. The system of embodiment 323, wherein said ligase is a T4 ligase.

325. The system of embodiment 323, wherein said ligase is a SplintR ligase.

326. The system of any one of embodiments 308-325, further comprising a polymerizing enzyme.

327. The system of embodiment 326, wherein said polymerizing enzyme is a phi29 polymerase.

328. The system of any one of embodiments 308-327, further comprising a salt.

329. The system of embodiment 328, wherein said salt comprises NaCl, MgCl2, or CaCl2.

330. The system of any one of embodiments 308-329, further comprising a solvent.

331. The system of embodiment 330, wherein said solvent is a polar solvent.

332. The system of any one of embodiments 308-331, further comprising a chaotropic reagent.

333. The system of any one of embodiments 308-332, further comprising a substrate.

334. The system of embodiment 333, wherein said substrate is an interior surface of a flow cell.

335. The system of any one of embodiments 308-334, wherein said one or more computer processors comprises a central processing unit.

336. The system of any one of embodiments 308-335, wherein said one or more computer processors comprises a graphic processing unit.

337. The system of any one of embodiments 308-336, wherein said one or more computer processors comprises a general-purpose processing unit.

338. The system of any one of embodiments 308-337, wherein said one or more processors comprises a field-programmable gate array.

339. The system of any one of embodiments 308-338, further comprising a fluidic system.

340. The system of embodiment 339, wherein said fluidic system is configured to provide fluid flow control capability for delivering reagents to a sample in a flow cell.

341. A kit for identifying a RNA *in situ*, the kit comprising:

(a) a probe comprising (i) an index sequence, and (ii) an RNA-binding sequence, wherein said RNA-binding sequence recognizes to at least a portion of said target;

(b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(c) instructions for identifying said RNA *in situ*, wherein said instructions comprise the method of any one of embodiments 1-67.

342. A kit for identifying a polypeptide *in situ*, the kit comprising:

(a) a polypeptide-binding probe comprising (i) a polypeptide binding moiety, and (ii) an oligonucleotide coupled thereto;

(b) a probe comprising (i) an index sequence, and (ii) an oligonucleotide-binding probe, wherein said oligonucleotide-binding probe recognizes to at least a portion of said oligonucleotide;

(c) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(d) instructions for identifying said polypeptide *in situ*, wherein said instructions comprise the method of any one of embodiments 68-134.

343. A kit for identifying a DNA *in situ*, the kit comprising:

(a) a probe comprising (i) an index sequence, and (ii) an RNA-binding sequence, wherein said RNA-binding sequence recognizes to at least a portion of said target;

(b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(c) instructions for identifying said RNA *in situ*, wherein said instructions comprise the method of any one of embodiments 135-195.

344. The kit of any one of embodiments 341-343, further comprising a primer nucleic acid sequence configured to bind to at least a portion of said index sequence or derivative thereof to produce said primed index sequence.

345. The kit of any one of embodiments 341-344, wherein said probe is a padlock probe.

346. The kit of embodiment 345, wherein said padlock probe comprises a nucleic acid.

347. The kit of any one of embodiments 341-346, further comprising a ligase.

348. The kit of embodiment 347, wherein said ligase is a T4 ligase.

349. The kit of embodiment 347, wherein said ligase is a SplintR ligase.

350. The kit of any one of embodiments 341-349, further comprising a polymerizing enzyme.

351. The kit of embodiment 350, wherein said polymerizing enzyme is a phi29 polymerase.

352. The kit of any one of embodiments 341-351, further comprising a buffer.

353. The kit of embodiment 352, wherein said buffer comprises a salt.

354. The kit of embodiment 353, wherein said salt comprises NaCl, $MgCl_2$, or $CaCl_2$.

355. The kit of any one of embodiments 341-354, wherein said buffer comprises a solvent.

356. The kit of embodiment 355, wherein said solvent is polar.

357. The kit of any one of embodiments 341-356, wherein said buffer comprises a chaotropic agent.

358. The kit of any one of embodiment 341-357, further comprising a solvent.

359. The kit of embodiment 358, wherein said solvent is methanol.

360. The kit of embodiment 358, wherein said solvent is acetone.

361. The kit of any one of embodiments 341-360, further comprising a detergent.

362. The kit of embodiment 361, wherein said detergent is saponin.

363. The kit of embodiment 361, wherein said detergent is Triton X-100.

364. The kit of any one of embodiments 341-363, further comprising a flow cell.

365. The kit of embodiment 364, wherein said flow cell comprises a substrate configured to be coupled to said cell or said tissue.

366. A kit for identifying a ribonucleic acid in situ, the kit comprising:

(a) a nucleic acid molecule comprising (i) an index sequence, and (ii) a target-binding sequence, wherein said target-binding sequence is complementary to at least a portion of said ribonucleic acid sequence;

(b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(c) instructions for identifying said ribonucleic acid in situ, wherein said instructions comprise the method of any one of embodiments 196-267.

367. The kit of embodiment 366, further comprising a primer nucleic acid sequence configured to bind to at least a portion of said index sequence or derivative thereof to produce a primed index sequence.

368. The kit of any one of embodiments 366 or 367, further comprising an index-binding molecule.

369. The kit of embodiment 368, wherein said index-binding molecule comprises a nucleic acid.

370. The kit of embodiment 368, wherein said index-binding molecule is a padlock probe.

371. The kit of embodiment 368, wherein said index-binding molecule is configured to bind to said index sequence.

372. The kit of any one of embodiments 366-371, further comprising a ligase.

373. The kit of embodiment 372, wherein said ligase is a T4 ligase.

374. The kit of embodiment 372, wherein said ligase is a SplintR ligase.

375. The kit of any one of embodiments 366-374, further comprising a polymerizing enzyme.

376. The kit of embodiment 375, wherein said polymerizing enzyme is a phi29 polymerase.

377. The kit of any one of embodiments 366-376, further comprising a buffer.

378. The kit of embodiment 377, wherein said buffer comprises a salt.

379. The kit of embodiment 378, wherein said salt comprises NaCl, MgCl2, or CaCl2.

380. The kit of any one of embodiments 377-379, wherein said buffer comprises a solvent.

381. The kit of embodiment 380, wherein said solvent is polar.

382. The kit of any one of embodiments 377-381, wherein said buffer comprises a chaotropic agent.

383. The kit of any one of embodiments 366-382, further comprising a flow cell.

384. The kit of embodiment 383, wherein said flow cell comprises a substrate configured to be coupled to said cell or said tissue.

CLAIMS

What is Claimed:

1. A method for identifying a ribonucleic acid (RNA) *in situ*, the method comprising:

(a) providing a cell or a tissue that has been permeabilized, wherein said cell or said tissue comprises said RNA;

(b) directing a probe to said cell or tissue, wherein said probe comprises (i) an RNA-binding sequence and (ii) an index sequence, under conditions sufficient to couple said RNA-binding sequence to at least a portion of said RNA;

(c) amplifying said index sequence to generate one or more amplicons, wherein said one or more amplicons comprises multiple copies of said index sequence or derivative thereof;

(d) directing a primer nucleic acid sequence to said cell or tissue under conditions sufficient to bind said primer nucleic acid sequence to a portion of said one or more amplicons to produce a primed index sequence;

(e) directing said cell or tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence or said derivative thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating said one or more nucleotide moieties into said primer nucleic acid sequence; and

(f) detecting said binding complex within said cell or tissue, thereby identifying said RNA.

2. The method of claim 1, wherein said method does not comprise performing a reverse transcription reaction of said RNA.

3. The method of any one of claims 1 or 2, further comprising providing said cell in (a).

4. The method of claim 1, wherein said cell is a cancer cell.

5. The method of any one of claims 1 or 2, further comprising providing said tissue in (a).

6. The method of claim 5, wherein said tissue is a formalin-fixed paraffin embedded tissue.

7. The method of claim 5, wherein said tissue is a fresh-frozen tissue.

8. The method of any one of claims 1 or 2, wherein said cell or said tissue is extracted from a subject.

9. The method of claim 8, wherein said subject is a human subject.

10. The method of any one of claims 1-9, wherein said index sequence comprises a nucleic acid.

11. The method of claim 10, wherein said index sequence is 4-50 nucleotides in length.

12. The method of any one of claims 1-11, wherein at least a portion of said index sequence or derivative thereof recognizes said primer nucleic acid sequence.
13. The method of any one of claims 1-11, wherein said primer nucleic acid sequence comprises a blocking group.
14. The method of claim 13, wherein said blocking group inhibits incorporation of another nucleotide into the primer nucleic acid sequence.
15. The method of claim 13, wherein said blocking group comprises a 3'-0-azido group, a 3'-0-azidomethyl group, a 3' -O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-0-malonyl group, or a 3'-0-benzyl group.
16. The method of any one of claims 1-15, wherein said primer nucleic acid sequence does not include a 3' hydroxyl group.
17. The method of any one of claims 1-16, wherein said primer nucleic acid sequence binds to at least a portion of said index sequence.
18. The method of any one of claims 1-17, wherein said primer nucleic acid sequence binds to at least a portion of a reverse complement of said index sequence.
19. The method of any one of claims 1-18, wherein said primed index sequence comprises a double-stranded nucleic acid.
20. The method of any one of claims 1-19, wherein said primed index sequence comprises a single-stranded nucleic acid.
21. The method of any one of claims 1-20, wherein said cell or said tissue is immobilized on a surface of a substrate.
22. The method of claim 21, wherein said surface of said substrate has a water contact angle of less than or equal to 45 degrees.
23. The method of claim 21, wherein said surface of said substrate is an interior surface of a flow cell.
24. The method of claim 23, wherein said interior surface of said flow cell comprises one or more hydrophilic polymer layers.
25. The method of claim 24, wherein said one or more hydrophilic polymer layers comprises a branched polymer.
26. The method of claim 24, wherein said one or more hydrophilic polymer layers comprises a polymer comprising polyethylene glycol.
27. The method of any one of claims 1-26, wherein said nucleotide conjugate comprises a common core.
28. The method of claim 27, wherein said one or more nucleotide moieties are coupled to said common core.

29. The method of claim 27 or 28, wherein said common core comprises a polymer, a micelle, a liposome, a microparticle, a nanoparticle, or a quantum dot.
30. The method of claim 29, wherein said polymer comprises a protein.
31. The method of any one of claims 1-30, wherein in (e) said nucleotide conjugate is included in a mixture of a plurality of nucleotide conjugates, wherein each of said plurality of nucleotide conjugates comprises different types of nucleotide moieties from each other.
32. The method of claim 31, wherein said different types of said nucleotide moieties comprise at least three different types of said nucleotide moieties.
33. The method of any one of claims 1-32, wherein said one or more nucleotide moieties do not comprise a blocking group coupled thereto.
34. The method of any one of claims 1-33, wherein said detecting in (f) comprises imaging said cell or said tissue using one or more image sensors.
35. The method of claim 34, wherein said one or more image sensors comprises a photodetector array.
36. The method of any one of claims 1-35, further comprising washing said cell or said tissue following (f) to remove said nucleotide conjugate from said cell or said tissue.
37. The method of claim 36, further comprising repeating (e) and (f) with a second nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a second nucleotide in said primed index sequence or derivative thereof.
38. The method of any one of claims 1-37, wherein said detectable label comprises a fluorophore.
39. The method of any one of claims 1-38, wherein said detectable label comprises a quantum dot.
40. The method of any one of claims 1-39, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises an incubation temperature, wherein said incubation temperature is about 20-45°C.
41. The method of any one of claims 1-40, wherein said conditions sufficient to couple said target-binding sequence to at least a portion of said target in (b) comprises a salt.
42. The method of any one of claims 1-41, wherein said RNA is a messenger RNA.
43. The method of any one of claims 1-41, wherein said RNA is a micro RNA.
44. The method of any one of claims 1-41, wherein said RNA is a ribosomal RNA.
45. The method of any one of claims 1-41, wherein said RNA is a transfer RNA.
46. The method of any one of claims 1-41, wherein said RNA is a transfer-messenger RNA.

47. The method of any one of claims 1-46, wherein said RNA comprises a modification.
48. The method of claim 47, wherein said modification is a 5' phosphorylation.
49. The method of claim 47, wherein said modification is an inter-nucleotide linkage.
50. The method of any one of claims 1-49, wherein said probe comprises nucleic acid.
51. The method of claim 50, wherein said nucleic acid comprises deoxyribonucleic acid (DNA).
52. The method of claim 50, wherein said nucleic acid comprises RNA.
53. The method of any one of claims 1-49, wherein said RNA-binding sequence comprises nucleic acid.
54. The method of claim 53, wherein said nucleic acid comprises deoxyribonucleic acid (DNA).
55. The method of claim 53, wherein said nucleic acid comprises RNA.
56. The method of any one of claims 1-55, wherein said probe is a padlock probe.
57. The method of claim 56, further comprising, before (c), ligating a first end of said probe to a second end of said probe using a ligase.
58. The method of claim 57, wherein said ligase is a T4 ligase.
59. The method of claim 57, wherein said ligase is a SplintR ligase.
60. The method of any one of claims 1-59, wherein performing said amplification reaction in (c) comprises performing a rolling circle amplification reaction using a polymerizing enzyme.
61. The method of claim 60, wherein said polymerizing enzyme is a phi29 polymerase.
62. The method of any one of claims 1-61, wherein said cell or said tissue is permeabilized using a solvent.
63. The method of claim 62, wherein said solvent is methanol.
64. The method of claim 62, wherein said solvent is acetone.
65. The method of any one of claims 1-61, wherein said cell or said tissue is permeabilized using a detergent.
66. The method of claim 65, wherein said detergent is saponin.
67. The method of claim 65, wherein said detergent is Triton X-100.
68. A method for identifying a polypeptide in situ, the method comprising:

(a) providing a cell or a tissue that has been permeabilized, wherein said cell or said tissue comprises said polypeptide;

(b) directing a polypeptide-binding probe to said cell or tissue, wherein said polypeptide-binding probe comprises a polypeptide binding moiety and an oligonucleotide

coupled thereto, under conditions sufficient to couple said polypeptide-binding probe to said polypeptide;

(c) directing a probe to said cell or said tissue, wherein said with a probe comprising (i) an oligonucleotide-binding probe and (ii) an index sequence under conditions sufficient to couple said oligonucleotide-binding probe to at least a portion of said oligonucleotide of said polypeptide-binding probe;

(d) amplifying said index sequence to generate one or more amplicons, wherein said one or more amplicons comprises multiple copies of said index sequence or derivative thereof;

(e) directing a primer nucleic acid sequence to said cell or said tissue under conditions sufficient to bind said primer nucleic acid sequence to at least a portion of said one or more amplicons to produce a primed index sequence;

(f) directing said cell or said tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence or said one or more amplicons thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating said one or more nucleotide moieties into said primer nucleic acid sequence; and

(g) detecting said binding complex within said cell or tissue, thereby identifying said polypeptide.

69. A method for identifying a DNA *in situ*, the method comprising:

(a) providing a cell or a tissue that has been permeabilized, wherein said cell or said tissue comprises said DNA, where said DNA is not cDNA;

(b) directing a probe to said cell or said tissue, wherein said probe comprises a (i) DNA-binding sequence and (ii) an index sequence, under conditions sufficient to couple said DNA-binding sequence to at least a portion of said DNA;

(c) amplifying said index sequence to generate one or more amplicons, wherein said one or more amplicons comprises multiple copies of said index sequence or derivative thereof;

(d) directing a primer nucleic acid sequence to said cell or said tissue under conditions sufficient to bind said primer nucleic acid sequence to at least a portion of said one or more amplicons to produce a primed index sequence;

(e) directing said cell or said tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence or said derivatives thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index

sequence without incorporating said one or more nucleotide moieties into said primer nucleic acid sequence; and

(f) detecting said binding complex within said cell or tissue, thereby identifying said DNA.

70. A method for identifying a ribonucleic acid in situ, the method comprising:

(a) providing a cell or tissue that has been permeabilized, wherein said cell or said tissue comprises said ribonucleic acid;

(b) directing a nucleic acid molecule to said cell or said tissue, wherein said nucleic acid molecule comprises a target-binding sequence, under conditions sufficient to couple said target-binding sequence to at least a portion of said ribonucleic acid, wherein said nucleic acid molecule further comprises an index sequence;

(c) directing a primer nucleic acid sequence to said cell or said tissue under conditions sufficient to bind said primer nucleic acid sequence to at least a portion of said one or more amplicons to produce a primed index sequence;

(d) directing said cell or said tissue with a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence or derivative thereof, under conditions suitable to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence;

(e) detecting said binding complex within said cell or tissue, thereby identifying said ribonucleic acid.

71. A system for identifying a RNA *in situ*, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of any one of clams 1-67.

72. A system for identifying a polypeptide *in situ*, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of claim 68.

73. A system for identifying a DNA *in situ*, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of claim 69.

74. A system for identifying a ribonucleic acid in situ, the system comprising:

a computing device comprising one or more computer processors, an operating system configured to perform executable instructions, a memory, and a computer program including instructions comprising the method of claim 70.

75. A kit for identifying a RNA *in situ*, the kit comprising:

(a) a probe comprising (i) an index sequence, and (ii) an RNA-binding sequence, wherein said RNA-binding sequence recognizes to at least a portion of said target;

(b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(c) instructions for identifying said RNA *in situ*, wherein said instructions comprise the method of any one of claims 1-67.

76. A kit for identifying a polypeptide *in situ*, the kit comprising:

(a) a polypeptide-binding probe comprising (i) a polypeptide binding moiety, and (ii) an oligonucleotide coupled thereto;

(b) a probe comprising (i) an index sequence, and (ii) an oligonucleotide-binding probe, wherein said oligonucleotide-binding probe recognizes to at least a portion of said oligonucleotide;

(c) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(d) instructions for identifying said polypeptide *in situ*, wherein said instructions comprise the method of claim 68.

77. A kit for identifying a DNA *in situ*, the kit comprising:

(a) a probe comprising (i) an index sequence, and (ii) an RNA-binding sequence, wherein said RNA-binding sequence recognizes to at least a portion of said target;

(b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide of a primed

index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(c) instructions for identifying said RNA *in situ*, wherein said instructions comprise the method of claim 69.

78. A kit for identifying a ribonucleic acid in situ, the kit comprising:

(a) a nucleic acid molecule comprising (i) an index sequence, and (ii) a target-binding sequence, wherein said target-binding sequence is complementary to at least a portion of said ribonucleic acid sequence;

(b) a nucleotide conjugate comprising (i) a detectable label and (ii) one or more nucleotide moieties complementary to a nucleotide in said primed index sequence, wherein said nucleotide conjugate is configured to form a binding complex between said one or more nucleotide moieties and said nucleotide of said primed index sequence without incorporating the one or more nucleotide moieties into said primer nucleic acid sequence; and

(c) instructions for identifying said ribonucleic acid in situ, wherein said instructions comprise the method of claim 70.