

METHODS FOR SPATIAL ANALYSIS USING RNA-TEMPLATED LIGATION

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Patent Application No.

- 5 62/952,736, filed December 23, 2019; U.S. Provisional Patent Application No. 62/969,458, filed February 3, 2020; U.S. Provisional Patent Application No. 63/087,061, filed October 2, 2020; and U.S. Provisional Patent Application No. 63/108,088, filed October 30, 2020. The contents of each of these applications are incorporated herein by reference in their entireties.

BACKGROUND

10

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

15

Spatial heterogeneity has been previously studied using techniques that typically provide data for a handful of analytes in the context of intact tissue or a portion of a tissue (e.g., tissue section), or provide significant analyte data from individual, single cells, but fails to provide information regarding the position of the single cells from the originating biological sample (e.g., tissue).

20

Generally, targeting a particular analyte in a biological sample utilizes a capture probe that targets a common transcript sequence such as a poly(A) mRNA-like tail. However, this approach is capable of detecting a high number of off target analytes. Methods such as RNA-templated ligation offer an alternative to indiscriminant capture of a common transcript sequence. *See, e.g., Yeakley, PLoS One, 25;12(5):e0178302 (2017)*, which is incorporated by reference in its entirety. However, there remains a need to develop an alternative to common transcript sequence (e.g., poly(A) mRNA-like tail) capture of target analytes that is capable of detecting an analyte(s) in an entire transcriptome while providing information regarding the spatial location and abundance of a target analyte.

25

30

SUMMARY

Targeted RNA capture is an attractive alternative to poly(A) mRNA capture in order to interrogate spatial gene expression in a sample (e.g., an FFPE tissue). Compared to

poly(A) mRNA capture, targeted RNA capture as described herein is less affected by RNA degradation associated with FFPE fixation compared to methods dependent on oligo-dT capture and reverse transcription of mRNA. Further targeted RNA capture as described herein allows for sensitive measurement of specific genes of interest that otherwise might be missed with a whole transcriptomic approach. Targeted RNA capture can be used to capture a defined set of RNA molecules of interest, or it can be used at a whole transcriptome level, or anything in between. When combined with the spatial methods disclosed herein, the location and abundance of the RNA targets can be determined.

In one aspect, this disclosure features a method for determining a location of an analyte in a biological sample including: (a) providing the biological sample on an array including a plurality of capture probes, where a capture probe of the plurality includes: (i) a spatial barcode and (ii) a capture domain; (b) contacting a first probe and a second probe with the biological sample, where the first probe and the second probe each include one or more sequences that are substantially complementary to sequences of the analyte, and where the second probe includes a capture probe capture domain; (c) hybridizing the first probe and the second probe to the analyte; (d) generating a ligation product by ligating the first probe and the second probe; (e) releasing the ligated product from the analyte; (f) hybridizing the ligation product to the capture domain; and (g) determining (i) all or a part of the sequence of the ligation product bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample. In some instances, the methods include determining the abundance and location of the analyte in the biological sample. In some instances, the methods include determining the abundance of an analyte at a location in the biological sample.

In some embodiments, the first probe and the second probe are substantially complementary to adjacent sequences of the analyte. In some embodiments, the first probe and the second probe hybridize to sequences that are not adjacent to each other on the analyte.

In some embodiments, the method further includes hybridizing a third probe to the first probe and the second probe, where the third probe includes a first sequence that is substantially complementary to a portion of the first probe and a second sequence that is substantially complementary to a portion of the second probe. In some embodiments, hybridizing the first probe and the second probe to the analyte is performed at a first temperature. In some embodiments, hybridizing the third probe to the first probe and the

second probe is performed at a second temperature. In some embodiments, the first temperature is a higher temperature than the second temperature. In some embodiments, the first temperature is from about 50°C to about 75°C, from about 55°C to about 70°C, or from about 60°C to about 65°C. In some embodiments, the second temperature is from about 15°C to about 35°C, from about 20°C to about 30°C, or from about 25°C to about 30°C. In some embodiments, the first probe is extended with a DNA polymerase, thereby filling in a gap between the first probe and the second probe and generating an extended first probe.

In some embodiments, the first probe includes a sequence that is substantially complementary to a first target sequence of the analyte. In some embodiments, the second probe further includes: (i) a first sequence that is substantially complementary to a second target sequence of the analyte; (ii) a linker sequence; (iii) a second sequence that is substantially complementary to a third target sequence of the analyte; and (iv) a capture probe capture domain that is capable of binding to a capture domain of a capture probe. In some embodiments, the first target sequence of the analyte is directly adjacent to the second target sequence of the analyte. In some embodiments, the second target sequence is not directly adjacent to the third target sequence on the analyte. In some embodiments, the second target sequence and the third target sequence are (i) on different exons of the analyte or (ii) located within the same exon of the analyte but are not adjacent on the analyte.

In some embodiments, the second probe includes a sequence that is substantially complementary to a third target sequence of the analyte. In some embodiments, the first probe includes: (i) a first sequence that is substantially complementary to a first target sequence of the analyte; (ii) a linker sequence; and (iii) second sequence that is substantially complementary to second target sequence of the analyte. In some embodiments, the second target sequence is directly adjacent to the third target sequence. In some embodiments, the first target sequence is not directly adjacent to the second target sequence on the analyte. In some embodiments, the first target sequence and second target sequence are (i) on different exons of the analyte or (ii) located within the same exon but are not directly adjacent on the analyte.

In some embodiments, the linker sequence includes a total of about 1 nucleotide to about 100 nucleotides. In some embodiments, the linker further includes a barcode sequence that serves as a proxy for identifying the analyte.

In some embodiments, the first probe includes at least two ribonucleic acid bases at the 3' end, and where the second probe includes a phosphorylated nucleotide at the 5' end.

In some embodiments, generating a ligation product includes ligating (i) the first probe to the second probe or (ii) the extended first probe to the second probe using enzymatic ligation or chemical ligation, where the enzymatic ligation utilizes a ligase. In some embodiments, the ligase is one or more of a T4 RNA ligase (Rnl2), a splintR ligase, a single stranded DNA ligase, or a T4 DNA ligase.

In some embodiments, the second probe includes a pre-adenylated phosphate group at its 5' end, and where the first probe includes at least two ribonucleic acid bases at the 3' end. In some embodiments, the step of generating a ligation product includes ligating a 3' end of the first probe to the 5' end of the second probe using a ligase that does not require adenosine triphosphate for ligase activity. In some embodiments, the ligase is selected from the group consisting of: thermostable 5' AppDNA/RNA Ligase, truncated T4 RNA Ligase 2, truncated T4 RNA Ligase 2 K227Q, truncated T4 RNA Ligase 2 KQ, Chlorella Virus PBCV-1 DNA Ligase, or any combination thereof.

In some embodiments, the first probe further includes a functional sequence, where the functional sequence is a primer sequence.

In some embodiments, the method further includes providing a capture probe capture domain blocking moiety that interacts with the capture probe capture domain. In some embodiments, the method further includes releasing the capture probe capture domain blocking moiety from the capture probe capture domain prior to step (f). In some embodiments, the capture probe capture domain includes a poly-adenylated (poly(A)) sequence or a complement thereof. In some embodiments, the capture probe capture domain blocking moiety includes a poly-uridine sequence, a poly-thymidine sequence, or both. In some embodiments, releasing the poly-uridine sequence from the poly(A) sequence includes denaturing the ligation product or contacting the ligation product with an endonuclease, exonuclease or ribonuclease. In some embodiments, the capture probe capture domain includes a sequence that is complementary to all or a portion of the capture domain of the capture probe. In some embodiments, the capture probe capture domain includes a degenerate sequence.

In some embodiments, the first probe and/or the second probe is a DNA probe.

In some embodiments, the third probe is a DNA probe.

In some embodiments, the capture probe capture domain blocking moiety is a DNA probe.

In some embodiments, the releasing step (f) includes removing the ligated probe from the analyte.

In some embodiments, the releasing of (i) the ligation product from the analyte or (ii) the capture probe capture domain blocking moiety from the capture domain binding domain, includes contacting the ligated probe with an endoribonuclease. In some embodiments, the endoribonuclease is one or more of RNase H, RNase A, RNase C, or RNase I. In some
5 embodiments, the RNase H includes RNase H1, RNase H2, or RNase H1 and RNase H2.

In some embodiments, the biological sample is a tissue sample. In some
embodiments, the tissue sample is a formalin-fixed, paraffin-embedded (FFPE) tissue sample, a fresh or a frozen tissue sample. In some embodiments, the tissue sample is the FFPE tissue
sample, and the tissue sample is decrosslinked. In some embodiments, the biological sample
10 was previously stained. In some embodiments, the biological sample was previously stained using immunofluorescence or immunohistochemistry. In some embodiments, the biological sample was previously stained using hematoxylin and eosin.

In some embodiments, the method further includes contacting the biological sample with a permeabilization agent, where the permeabilization agent is selected from an organic
15 solvent, a detergent, and an enzyme, or a combination thereof. In some embodiments, the permeabilization agent is selected from the group consisting of: an endopeptidase, a protease sodium dodecyl sulfate (SDS), polyethylene glycol tert-octylphenyl ether, polysorbate 80, and polysorbate 20, N-lauroylsarcosine sodium salt solution, saponin, Triton X-100™, and Tween-20™. In some embodiments, the endopeptidase is pepsin or proteinase K.

20 In some embodiments, the method further includes, prior to step (a), fixing the biological sample. In some embodiments, the step of fixing the biological sample is performed using one or both of methanol and acetone.

In some embodiments, the analyte includes RNA. In some embodiments, the RNA is an mRNA.

25 In some embodiments, the determining step includes amplifying all or part of the ligation product specifically bound to the capture domain. In some embodiments, an amplifying product includes (i) all or part of sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof. In some embodiments, the determining step
30 includes sequencing. In some embodiments, the sequencing step includes *in situ* sequencing, Sanger sequencing methods, next-generation sequencing methods, and nanopore sequencing.

In another aspect, this disclosure features a kit including: (a) a substrate including a plurality of capture probes including a spatial barcode and a capture domain; (b) a system including: a plurality of first probes and second probes, where a first probe and a second

probe each includes sequences that are substantially complementary to an analyte, and where the second probe includes a capture binding domain; and (c) instructions for performing the method of any one of the preceding claims.

In another aspect, this disclosure features a kit including: (a) an array including a plurality of capture probes; (b) a plurality of probes including a first probe and a second, where the first probe and the second probe are substantially complementary to adjacent sequences of an analyte, where the second probe includes (i) a capture probe capture domain that is capable of binding to a capture domain of the capture probe and (ii) a linker sequence; (c) a plurality of enzymes including a ribonuclease and a ligase; and (d) instructions for performing the method of any one of the preceding claims.

In another aspect, this disclosure features a kit including: (a) an array including a plurality of capture probes; (b) a plurality of probes including a first probe and a second probe, where the first probe and the second probe are substantially complementary to adjacent sequences of an analyte, where the first probe includes a linker sequence, where the second probe includes a capture probe capture domain that is capable of binding to a capture domain of the capture probe; (c) a plurality of enzymes including a ribonuclease and a ligase; and (d) instructions for performing the method of any one of the preceding claims.

In some embodiments, the kit includes a second probe including a preadenylated phosphate group at its 5' end, and a ligase that does not require adenosine triphosphate for ligase activity.

In another aspect, this disclosure features a composition including a spatial array including capture probes, where the capture probes include a spatial barcode and a capture domain, a biological sample on the spatial array where the biological sample includes a plurality of analytes of interest, a first probe oligonucleotide and a second probe oligonucleotide hybridized to an analyte and ligated together, where the first probe oligonucleotide and the second probe oligonucleotide each include a sequence that is substantially complementary to adjacent sequences of the analyte and where one of the first probe or the second probe includes a capture probe capture domain.

In some embodiments, the composition further includes an RNase H enzyme. In some embodiments, the composition further includes a ligase. In some embodiments, the probe oligonucleotide that does not include a capture probe capture domain includes a functional domain.

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,

patent application, or item of information was specifically and individually indicated to be incorporated by reference. To the extent publications, patents, patent applications, and items of information 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.

Where values are described in terms of ranges, it should be understood that the description includes the disclosure of all possible sub-ranges within such ranges, as well as specific numerical values that fall within such ranges irrespective of whether a specific numerical value or specific sub-range is expressly stated.

The term “each,” when used in reference to a collection of items, is intended to identify an individual item in the collection but does not necessarily refer to every item in the collection, unless expressly stated otherwise, or unless the context of the usage clearly indicates otherwise.

The singular form “a”, “an”, and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a cell” includes one or more cells, comprising mixtures thereof. “A and/or B” is used herein to include all of the following alternatives: “A”, “B”, “A or B”, and “A and B”.

Various embodiments of the features of this disclosure are described herein. However, it should be understood that such embodiments are provided merely by way of example, and numerous variations, changes, and substitutions can occur to those skilled in the art without departing from the scope of this disclosure. It should also be understood that various alternatives to the specific embodiments described herein are also within the scope of this disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings illustrate certain embodiments of the features and advantages of this disclosure. These embodiments are not intended to limit the scope of the appended claims in any manner. Like reference symbols in the drawings indicate like elements.

FIG. 1 shows an exemplary spatial analysis workflow.

FIG. 2 is a schematic diagram showing an example of a barcoded capture probe, as described herein.

FIG. 3 is a schematic illustrating a cleavable capture probe, wherein the cleaved capture probe can enter into a non-permeabilized cell and bind to target analytes within the sample.

FIG. 4 is a schematic diagram of an exemplary multiplexed spatially-barcoded feature.

FIG. 5 is a schematic showing the exemplary arrangement of barcoded features within an array.

5 **FIG. 6** is a schematic diagram showing an exemplary workflow for RNA-templated ligation.

FIG. 7 is a schematic diagram showing an exemplary workflow for capturing a ligation product on a substrate that includes capture probes.

10 **FIG. 8** is a schematic diagram showing an example of a first probe that includes a linker sequence and a second probe.

FIG. 9 is a schematic diagram showing an example of a second probe that includes a linker sequence and a first probe.

FIG. 10 is a schematic diagram showing an example of a first probe, a second probe, and a spanning probe.

15 **FIG. 11** is a schematic diagram showing an example of an RNA-templated ligation using a first probe oligonucleotide, a second probe oligonucleotide, and a third probe oligonucleotide.

20 **FIGs. 12A-12E** show various approaches for chemically-mediated nucleic acid ligation. **FIG. 12A** illustrates formation of a triazole bond. **FIG. 12B** illustrates formation of a phosphorothioate bond. **FIG. 12C** illustrates formation of an amide bond. **FIG. 12D** illustrates a formation of phosphoramidate bond. **FIG. 12E** illustrates a conjugation reaction.

FIG. 13 shows an exemplary RNA-templated ligation workflow.

25 **FIG. 14** show enrichment PCR results for detection of various probes. R1 = run 1; R2 = run 2. (1) represents band for ligation product with capture probe. (2), (3) and (4) represent non-ligation products.

FIG. 15 shows fraction of reads for various combination of conditions and probes.

FIGs. 16A-16B show matrices of specific probe combinations.

30 **FIGs. 17A-17C** show gene expression patterns in mouse brain tissue using total probe oligonucleotide counts (**FIG. 17A**), nonspecific probe oligonucleotide counts (**FIG. 17B**), and gene-specific probe oligonucleotide counts (**FIG. 17C**).

FIGs. 18A-18E show target genes with spatial expression in mouse brain tissue.

FIGs. 19A-19F show gene expression patterns in mouse brain tissue using total probe oligonucleotide counts (**FIG. 19A**) and gene-specific probe oligonucleotide (**FIGs. 19B-19F**).

FIG. 20A shows median unique molecular identifiers (UMIs) per cell versus mean reads per cell for different hybridization conditions.

FIG. 20B shows median genes per cell versus mean reads per cell for different hybridization conditions.

5 **FIG. 21A** shows hematoxylin staining for a section of a triple positive breast cancer tissue sample.

FIG. 21B shows median UMI per cell count versus mean read per cell for the different conditions.

FIG. 21C shows t-SNE projection of spots in eight different clusters.

10 **FIG. 21D** shows the same tissue from **FIG. 21A** with various clusters (n=8 clusters) expressed in distinct areas of the tissue.

FIG. 21E shows expression of estrogen receptor (ESR1).

FIG. 21F shows expression of estrogen receptor progesterone receptor (PGR).

FIG. 21G shows expression of ERBB2, also known as HER2.

15 **FIG. 22A** shows a hematoxylin stain for a section of Ovarian Cancer sample.

FIG. 22B shows median UMI per cell versus mean read per cell for the different conditions.

FIG. 22C shows the same tissue from **FIG. 22A** with various clusters (n=8 clusters) expressed in distinct areas of the tissue.

20 **FIG. 22D** shows t-SNE projection of spots in eight different clusters.

FIG. 23 shows median UMIs per cell versus reads per cell for different decrosslinking conditions.

FIG. 24 shows median UMIs per cell versus reads per cell (top panel) and median genes per cell versus reads per cell (bottom panel) for different treatment conditions (RNase).

25 **FIG. 25** shows median UMIs per cell versus reads per cell (top panel) and median genes per cell versus reads per cell (bottom panel) for different treatment conditions (time and temperature).

DETAILED DESCRIPTION

30 Targeted RNA capture is an attractive alternative to poly(A) mRNA capture in order to interrogate spatial gene expression in FFPE tissue. Compared to poly(A) mRNA capture, targeted RNA capture is less affected by RNA degradation associated with FFPE fixation than methods dependent on oligo-dT capture and reverse transcription of mRNA; allows for sensitive measurement of specific genes of interest that otherwise might be missed with a

whole transcriptomic approach; and is scalable, with demonstrated probes targeting a large fraction of the transcriptome.

Spatial analysis methodologies and compositions described herein can provide a vast amount of analyte and/or expression data for a variety of analytes within a biological sample at high spatial resolution, while retaining native spatial context. Spatial analysis methods and compositions can include, e.g., the use of a capture probe including a spatial barcode (e.g., a nucleic acid sequence that provides information as to the location or position of an analyte within a cell or a tissue sample (e.g., mammalian cell or a mammalian tissue sample) and a capture domain that is capable of binding to an analyte (e.g., a protein and/or a nucleic acid) produced by and/or present in a cell. Spatial analysis methods and compositions can also include the use of a capture probe having a capture domain that captures an intermediate agent for indirect detection of an analyte. For example, the intermediate agent can include a nucleic acid sequence (e.g., a barcode) associated with the intermediate agent. Detection of the intermediate agent is therefore indicative of the analyte in the cell or tissue sample.

Non-limiting aspects of spatial analysis methodologies and compositions are described in U.S. Patent Nos. 10,774,374, 10,724,078, 10,480,022, 10,059,990, 10,041,949, 10,002,316, 9,879,313, 9,783,841, 9,727,810, 9,593,365, 8,951,726, 8,604,182, 7,709,198, U.S. Patent Application Publication Nos. 2020/239946, 2020/080136, 2020/0277663, 2020/024641, 2019/330617, 2019/264268, 2020/256867, 2020/224244, 2019/194709, 2019/161796, 2019/085383, 2019/055594, 2018/216161, 2018/051322, 2018/0245142, 2017/241911, 2017/089811, 2017/067096, 2017/029875, 2017/0016053, 2016/108458, 2015/000854, 2013/171621, WO 2018/091676, WO 2020/176788, Rodriques et al., Science 363(6434):1463-1467, 2019; Lee et al., Nat. Protoc. 10(3):442-458, 2015; Trejo et al., PLoS ONE 14(2):e0212031, 2019; Chen et al., Science 348(6233):aaa6090, 2015; Gao et al., BMC Biol. 15:50, 2017; and Gupta et al., Nature Biotechnol. 36:1197-1202, 2018; the Visium Spatial Gene Expression Reagent Kits User Guide (e.g., Rev C, dated June 2020), and/or the Visium Spatial Tissue Optimization Reagent Kits User Guide (e.g., Rev C, dated July 2020), both of which are available at the 10x Genomics Support Documentation website, and can be used herein in any combination. Further non-limiting aspects of spatial analysis methodologies and compositions are described herein.

Some general terminology that may be used in this disclosure can be found in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. Typically, a “barcode” is a label, or identifier, that conveys or is capable of conveying information (e.g., information about an analyte in a sample, a bead, and/or a capture probe). A barcode can be

part of an analyte, or independent of an analyte. A barcode can be attached to an analyte. A particular barcode can be unique relative to other barcodes. For the purpose of this disclosure, an “analyte” can include any biological substance, structure, moiety, or component to be analyzed. The term “target” can similarly refer to an analyte of interest.

5 Analytes can be broadly classified into one of two groups: nucleic acid analytes, and non-nucleic acid analytes. Examples of non-nucleic acid analytes include, but are not limited to, lipids, carbohydrates, peptides, proteins, glycoproteins (N-linked or O-linked), lipoproteins, phosphoproteins, specific phosphorylated or acetylated variants of proteins, amidation variants of proteins, hydroxylation variants of proteins, methylation variants of
10 proteins, ubiquitylation variants of proteins, sulfation variants of proteins, viral proteins (e.g., viral capsid, viral envelope, viral coat, viral accessory, viral glycoproteins, viral spike, etc.), extracellular and intracellular proteins, antibodies, and antigen binding fragments. In some embodiments, the analyte(s) can be localized to subcellular location(s), including, for example, organelles, e.g., mitochondria, Golgi apparatus, endoplasmic reticulum,
15 chloroplasts, endocytic vesicles, exocytic vesicles, vacuoles, lysosomes, etc. In some embodiments, analyte(s) can be peptides or proteins, including without limitation antibodies and enzymes. Additional examples of analytes can be found in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. In some embodiments, an analyte can be detected indirectly, such as through detection of an intermediate agent, for example, a ligation
20 product or an analyte capture agent (e.g., an oligonucleotide-conjugated antibody), such as those described herein.

 A “biological sample” is typically obtained from the subject for analysis using any of a variety of techniques including, but not limited to, biopsy, surgery, and laser capture microscopy (LCM), and generally includes cells and/or other biological material from the
25 subject. In some embodiments, a biological sample can be a tissue section. In some embodiments, a biological sample can be a fixed and/or stained biological sample (e.g., a fixed and/or stained tissue section). Non-limiting examples of stains include histological stains (e.g., hematoxylin and/or eosin) and immunological stains (e.g., fluorescent stains). In some embodiments, a biological sample (e.g., a fixed and/or stained biological sample) can
30 be imaged. Biological samples are also described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

 In some embodiments, a biological sample is permeabilized with one or more permeabilization reagents. For example, permeabilization of a biological sample can facilitate

analyte capture. Exemplary permeabilization agents and conditions are described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

Array-based spatial analysis methods involve the transfer of one or more analytes from a biological sample to an array of features on a substrate, where each feature is associated with a unique spatial location on the array. Subsequent analysis of the transferred analytes includes determining the identity of the analytes and the spatial location of the analytes within the biological sample. The spatial location of an analyte within the biological sample is determined based on the feature to which the analyte is bound (e.g., directly or indirectly) on the array, and the feature's relative spatial location within the array.

A "capture probe" refers to any molecule capable of capturing (directly or indirectly) and/or labelling an analyte (e.g., an analyte of interest) in a biological sample. In some embodiments, the capture probe is a nucleic acid or a polypeptide. In some embodiments, the capture probe includes a barcode (e.g., a spatial barcode and/or a unique molecular identifier (UMI)) and a capture domain). In some embodiments, a capture probe can include a cleavage domain and/or a functional domain (e.g., a primer-binding site, such as for next-generation sequencing (NGS)). See, e.g., WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. Generation of capture probes can be achieved by any appropriate method, including those described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

In some embodiments, more than one analyte type (e.g., nucleic acids and proteins) from a biological sample can be detected (e.g., simultaneously or sequentially) using any appropriate multiplexing technique, such as those described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

In some embodiments, detection of one or more analytes (e.g., protein analytes) can be performed using one or more analyte capture agents. As used herein, an "analyte capture agent" refers to an agent that interacts with an analyte (e.g., an analyte in a biological sample) and with a capture probe (e.g., a capture probe attached to a substrate or a feature) to identify the analyte. In some embodiments, the analyte capture agent includes: (i) an analyte binding moiety (e.g., that binds to an analyte), for example, an antibody or antigen-binding fragment thereof; (ii) analyte binding moiety barcode; and (iii) an analyte capture sequence. As used herein, the term "analyte binding moiety barcode" refers to a barcode that is associated with or otherwise identifies the analyte binding moiety. As used herein, the term "analyte capture sequence" refers to a region or moiety configured to hybridize to, bind to, couple to, or otherwise interact with a capture domain of a capture probe. In some cases, an analyte

binding moiety barcode (or portion thereof) may be able to be removed (e.g., cleaved) from the analyte capture agent. Additional description of analyte capture agents can be found in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

There are at least two methods to associate a spatial barcode with one or more
5 neighboring cells, such that the spatial barcode identifies the one or more cells, and/or contents of the one or more cells, as associated with a particular spatial location. One method is to promote analytes or analyte proxies (e.g., intermediate agents) out of a cell and towards a spatially-barcoded array (e.g., including spatially-barcoded capture probes). In some instances, the spatially-barcoded array populated with capture probes (as described further
10 herein) is contacted with a biological sample, and the biological sample is permeabilized, allowing the analyte to migrate away from the sample and toward the array. The analyte interacts with a capture probe on the spatially-barcoded array. Another method is to cleave spatially-barcoded capture probes from an array and promote the spatially-barcoded capture probes towards and/or into or onto the biological sample. In some instances, the spatially-
15 barcoded array populated with capture probes (as described further herein) can be contacted with a sample. The spatially-barcoded capture probes are cleaved and then interact with cells within the provided biological sample. The interaction can be a covalent or non-covalent cell-surface interaction. The interaction can be an intracellular interaction facilitated by a delivery system or a cell penetration peptide. Once the spatially-barcoded capture probe is associated
20 with a particular cell, the sample can be optionally removed for analysis. The sample can be optionally dissociated before analysis. Once the tagged cell is associated with the spatially-barcoded capture probe, the capture probes can be analyzed to obtain spatially-resolved information about the tagged cell.

In some instances, sample preparation may include placing the sample on a slide,
25 fixing the sample, and/or staining the biological sample for imaging. The stained sample can be then imaged on the array using both brightfield (to image the sample hematoxylin and eosin stain) and/or fluorescence (to image features) modalities. Optionally, the sample can be destained prior to permeabilization. In some embodiments, analytes are then released from the sample and capture probes forming the spatially-barcoded array hybridize or bind the
30 released analytes. The sample is then removed from the array and the capture probes cleaved from the array. The biological sample and array are then optionally imaged a second time in one or both modalities while the analytes are reverse transcribed into cDNA, and an amplicon library is prepared and sequenced. Images are then spatially-overlaid in order to correlate spatially-identified biological sample information. When the sample and array are not imaged

a second time, a spot coordinate file is supplied instead. The spot coordinate file replaces the second imaging step. Further, amplicon library preparation can be performed with a unique PCR adapter and sequenced.

In some instances, disclosed is another exemplary workflow that utilizes a spatially-
5 barcoded array on a substrate, where spatially-barcoded capture probes are clustered at areas called features. The spatially-barcoded capture probes can include a cleavage domain, one or more functional domains, a spatial barcode, a unique molecular identifier, and a capture domain. The spatially-barcoded capture probes can also include a 5' end modification for reversible attachment to the substrate. The spatially-barcoded array is contacted with a
10 biological sample, and the sample is permeabilized through application of permeabilization reagents. Permeabilization reagents may be administered by placing the array/sample assembly within a bulk solution. Alternatively, permeabilization reagents may be administered to the sample via a diffusion-resistant medium and/or a physical barrier such as a lid, wherein the sample is sandwiched between the diffusion-resistant medium and/or
15 barrier and the array-containing substrate. The analytes are migrated toward the spatially-barcoded capture array using any number of techniques disclosed herein. For example, analyte migration can occur using a diffusion-resistant medium lid and passive migration. As another example, analyte migration can be active migration, using an electrophoretic transfer system, for example. Once the analytes are in close proximity to the spatially-barcoded
20 capture probes, the capture probes can hybridize or otherwise bind a target analyte. The biological sample can be optionally removed from the array.

The capture probes can be optionally cleaved from the array, and the captured analytes can be spatially-barcoded by performing a reverse transcriptase first strand cDNA reaction. A first strand cDNA reaction can be optionally performed using template switching
25 oligonucleotides. For example, a template switching oligonucleotide can hybridize to a poly(C) tail added to a 3' end of the cDNA by a reverse transcriptase enzyme in a template independent manner. The original mRNA template and template switching oligonucleotide can then be denatured from the cDNA and the spatially-barcoded capture probe can then hybridize with the cDNA and a complement of the cDNA can be generated. The first strand
30 cDNA can then be purified and collected for downstream amplification steps. The first strand cDNA can be amplified using PCR, where the forward and reverse primers flank the spatial barcode and analyte regions of interest, generating a library associated with a particular spatial barcode. In some embodiments, the library preparation can be quantitated and/or quality controlled to verify the success of the library preparation steps. In some embodiments,

the cDNA comprises a sequencing by synthesis (SBS) primer sequence. The library amplicons are sequenced and analyzed to decode spatial information.

In some cases, capture probes may be configured to prime, replicate, and consequently yield optionally barcoded extension products from a template (e.g., a DNA or RNA template, such as an analyte or an intermediate agent (e.g., a ligation product or an analyte capture agent), or a portion thereof), or derivatives thereof (see, e.g., WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663 regarding extended capture probes). In some cases, capture probes may be configured to form ligation products with a template (e.g., a DNA or RNA template, such as an analyte or an intermediate agent, or portion thereof), thereby creating ligations products that serve as proxies for a template.

A “capture probe” refers to any molecule capable of capturing (directly or indirectly) and/or labelling an analyte (e.g., an analyte of interest) in a biological sample. In some embodiments, the capture probe is a nucleic acid or a polypeptide. In some embodiments, the capture probe includes a barcode (e.g., a spatial barcode and/or a unique molecular identifier (UMI)) and a capture domain. In some instances, the capture probe can include functional sequences that are useful for subsequent processing. In some instances, a capture probe can be reversibly attached to a substrate via a linker. The capture probe can include one or more functional sequences, which can include a sequencer specific flow cell attachment sequence, e.g., a P5 or P7 sequence, as well as functional sequence, which can include sequencing primer sequences, e.g., a R1 primer binding site, a R2 primer binding site. In some embodiments, sequence is a P7 sequence and sequence is a R2 primer binding site. A capture probe can additionally include a spatial barcode and/or unique molecular identifier and a capture domain. The different sequences of the capture probe need not be in the sequential manner as depicted in this example, however the capture domain should be placed in a location on the barcode wherein analyte capture and extension of the capture domain to create a copy of the analyte can occur.

FIG. 2 is a schematic diagram showing an example of a capture probe, as described herein. As shown, the capture probe **202** is optionally coupled to a feature **201** by a cleavage domain **203**, such as a disulfide linker. The capture probe can include functional sequences that are useful for subsequent processing, such as functional sequence **204**, which can include a sequencer specific flow cell attachment sequence, e.g., a P5 or P7 sequence, as well as functional sequence **205**, which can include sequencing primer sequences, e.g., a R1 primer binding site. In some embodiments, sequence **204** is a P7 sequence and sequence **205** is a R1

primer binding site. A spatial barcode **206** can be included within the capture probe for use in barcoding the target analyte. The functional sequences can generally be selected for compatibility with any of a variety of different sequencing systems, e.g., Ion Torrent Proton or PGM, Illumina sequencing instruments, PacBio, Oxford Nanopore, etc., and the requirements thereof. In some embodiments, functional sequences can be selected for compatibility with non-commercialized sequencing systems. Examples of such sequencing systems and techniques, for which suitable functional sequences can be used, include (but are not limited to) Ion Torrent Proton or PGM sequencing, Illumina sequencing, PacBio SMRT sequencing, and Oxford Nanopore sequencing. Further, in some embodiments, functional sequences can be selected for compatibility with other sequencing systems.

In some embodiments, the spatial barcode **206**, functional sequences **204** (e.g., flow cell attachment sequence) and **205** (e.g., sequencing primer sequences) can be common to all of the probes attached to a given feature. The spatial barcode can also include a capture domain **207** to facilitate capture of a target analyte.

In some cases, capture probes are introduced into the cell using a cell-penetrating peptide. **FIG. 3** is a schematic illustrating a cleavable capture probe that includes a cell-penetrating peptide, wherein the cleaved capture probe can enter into a non-permeabilized cell and bind to analytes within the sample. The capture probe **301** contains a cleavage domain **302**, a cell penetrating peptide **303**, a reporter molecule **304**, and a disulfide bond (-S-S-). **305** represents all other parts of a capture probe, for example a spatial barcode and a capture domain.

In some instances, the disclosure provides multiplexed spatially-barcoded features. **FIG. 4** is a schematic diagram of an exemplary multiplexed spatially-barcoded feature. In **FIG. 4**, the feature **401** (e.g., a bead, a location on a slide or other substrate, a well on a slide or other substrate, a partition on a slide or other substrate, etc.) can be coupled to spatially-barcoded capture probes, wherein the spatially-barcoded probes of a particular feature can possess the same spatial barcode, but have different capture domains designed to associate the spatial barcode of the feature with more than one target analyte. For example, a feature may be coupled to four different types of spatially-barcoded capture probes, each type of spatially-barcoded capture probe possessing the spatial barcode **402**. One type of capture probe associated with the feature includes the spatial barcode **402** in combination with a poly(T) capture domain **403**, designed to capture mRNA target analytes. A second type of capture probe associated with the feature includes the spatial barcode **402** in combination with a random N-mer capture domain **404** for gDNA analysis. A third type of capture probe

associated with the feature includes the spatial barcode **402** in combination with a capture domain complementary to the analyte capture agent of interest **405**. A fourth type of capture probe associated with the feature includes the spatial barcode **402** in combination with a capture probe that can specifically bind a nucleic acid molecule **406** that can function in a CRISPR assay (e.g., CRISPR/Cas9). While only four different capture probe-barcode constructs are shown in **FIG. 4**, capture-probe barcoded constructs can be tailored for analyses of any given analyte associated with a nucleic acid and capable of binding with such a construct. For example, the schemes shown in **FIG. 4** can also be used for concurrent analysis of other analytes disclosed herein, including, but not limited to: (a) mRNA, a lineage tracing construct, cell surface or intracellular proteins and metabolites, and gDNA; (b) mRNA, accessible chromatin (e.g., ATAC-seq, DNase-seq, and/or MNase-seq) cell surface or intracellular proteins and metabolites, and a perturbation agent (e.g., a CRISPR crRNA/sgRNA, TALEN, zinc finger nuclease, and/or antisense oligonucleotide as described herein); (c) mRNA, cell surface or intracellular proteins and/or metabolites, a barcoded labelling agent (e.g., the MHC multimers described herein), and a V(D)J sequence of an immune cell receptor (e.g., T-cell receptor). In some embodiments, a perturbation agent can be a small molecule, an antibody, a drug, an aptamer, a miRNA, a physical environmental (e.g., temperature change), or any other known perturbation agents.

Additional features of capture probes are described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663, each of which is incorporated by reference in its entirety. Generation of capture probes can be achieved by any appropriate method, including those described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663, each of which is incorporated by reference in its entirety.

As used herein, an “extended capture probe” refers to a capture probe having additional nucleotides added to the terminus (e.g., 3’ or 5’ end) of the capture probe thereby extending the overall length of the capture probe. For example, an “extended 3’ end” indicates additional nucleotides were added to the most 3’ nucleotide of the capture probe to extend the length of the capture probe, for example, by polymerization reactions used to extend nucleic acid molecules including templated polymerization catalyzed by a polymerase (e.g., a DNA polymerase or a reverse transcriptase). In some embodiments, extending the capture probe includes adding to a 3’ end of a capture probe a nucleic acid sequence that is complementary to a nucleic acid sequence of an analyte or intermediate agent specifically bound to the capture domain of the capture probe. In some embodiments, the capture probe is extended using reverse transcription. In some embodiments, the capture probe is extended

using one or more DNA polymerases. The extended capture probes include the sequence of the capture probe and the sequence of the spatial barcode of the capture probe.

In some embodiments, extended capture probes are amplified (e.g., in bulk solution or on the array) to yield quantities that are sufficient for downstream analysis, e.g., via DNA sequencing. In some embodiments, extended capture probes (e.g., DNA molecules) act as
5 templates for an amplification reaction (e.g., a polymerase chain reaction).

Additional variants of spatial analysis methods, including in some embodiments, an imaging step, are described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. Analysis of captured analytes (and/or intermediate agents or portions thereof), for example, including sample removal, extension of capture probes, sequencing (e.g., of a cleaved extended capture probe and/or a cDNA molecule complementary to an extended capture probe), sequencing on the array (e.g., using, for example, in situ hybridization or in situ ligation approaches), temporal analysis, and/or proximity capture, is described in WO 2020/176788 and/or U.S. Patent Application Publication No.
10 2020/0277663. Some quality control measures are described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

Spatial information can provide information of biological and/or medical importance. For example, the methods and compositions described herein can allow for: identification of one or more biomarkers (e.g., diagnostic, prognostic, and/or for determination of efficacy of a treatment) of a disease or disorder; identification of a candidate drug target for treatment of a disease or disorder; identification (e.g., diagnosis) of a subject as having a disease or disorder; identification of stage and/or prognosis of a disease or disorder in a subject; identification of a subject as having an increased likelihood of developing a disease or disorder; monitoring of progression of a disease or disorder in a subject; determination of efficacy of a treatment of a disease or disorder in a subject; identification of a patient subpopulation for which a treatment is effective for a disease or disorder; modification of a treatment of a subject with a disease or disorder; selection of a subject for participation in a clinical trial; and/or selection of a treatment for a subject with a disease or disorder.
20 25

Spatial information can provide information of biological importance. For example, the methods and compositions described herein can allow for: identification of transcriptome and/or proteome expression profiles (e.g., in healthy and/or diseased tissue); identification of multiple analyte types in close proximity (e.g., nearest neighbor analysis); determination of up- and/or down-regulated genes and/or proteins in diseased tissue; characterization of tumor microenvironments; characterization of tumor immune responses; characterization of cells
30

types and their co-localization in tissue; and identification of genetic variants within tissues (e.g., based on gene and/or protein expression profiles associated with specific disease or disorder biomarkers).

Typically, for spatial array-based methods, a substrate functions as a support for direct or indirect attachment of capture probes to features of the array. A “feature” is an entity that acts as a support or repository for various molecular entities used in spatial analysis. In some embodiments, some or all of the features in an array are functionalized for analyte capture. Exemplary substrates are described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. Exemplary features and geometric attributes of an array can be found in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

FIG. 5 depicts an exemplary arrangement of barcoded features within an array. From left to right, **FIG. 5** shows (left) a slide including six spatially-barcoded arrays, (center) an enlarged schematic of one of the six spatially-barcoded arrays, showing a grid of barcoded features in relation to a biological sample, and (right) an enlarged schematic of one section of an array, showing the specific identification of multiple features within the array (labelled as ID578, ID579, ID560, etc.).

Generally, analytes and/or intermediate agents (or portions thereof) can be captured when contacting a biological sample with a substrate including capture probes (e.g., a substrate with capture probes embedded, spotted, printed, fabricated on the substrate, or a substrate with features (e.g., beads, wells, areas on a substrate) comprising capture probes). As used herein, “contact,” “contacted,” and/or “contacting,” a biological sample with a substrate refers to any contact (e.g., direct or indirect) such that capture probes can interact (e.g., bind covalently or non-covalently (e.g., hybridize)) with analytes from the biological sample. Capture can be achieved actively (e.g., using electrophoresis) or passively (e.g., using diffusion). Analyte capture is further described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

In some cases, spatial analysis can be performed by attaching and/or introducing a molecule (e.g., a peptide, a lipid, or a nucleic acid molecule) having a barcode (e.g., a spatial barcode) to a biological sample (e.g., to a cell in a biological sample). In some embodiments, a plurality of molecules (e.g., a plurality of nucleic acid molecules) having a plurality of barcodes (e.g., a plurality of spatial barcodes) are introduced to a biological sample (e.g., to a plurality of cells in a biological sample) for use in spatial analysis. In some embodiments, after attaching and/or introducing a molecule having a barcode to a biological sample, the biological sample can be physically separated (e.g., dissociated) into single cells or cell

groups for analysis. Some such methods of spatial analysis are described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

In some cases, spatial analysis can be performed by detecting multiple oligonucleotides that hybridize to an analyte. In some instances, for example, spatial analysis
5 can be performed using RNA-templated ligation (RTL). Methods of RTL have been described previously. See, e.g., Credle *et al.*, *Nucleic Acids Res.* 2017 Aug 21;45(14):e128. Typically, RTL includes hybridization of two oligonucleotides to adjacent sequences on an analyte (e.g., an RNA molecule, such as an mRNA molecule). In some instances, the oligonucleotides are DNA molecules. In some instances, one of the oligonucleotides includes
10 at least two ribonucleic acid bases at the 3' end and/or the other oligonucleotide includes a phosphorylated nucleotide at the 5' end. In some instances, one of the two oligonucleotides includes a capture domain (e.g., a poly(A) sequence, a non-homopolymeric sequence). After hybridization to the analyte, a ligase (e.g., SplintR ligase) ligates the two oligonucleotides together, creating a ligation product. In some instances, the two oligonucleotides hybridize to
15 sequences that are not adjacent to one another. For example, hybridization of the two oligonucleotides creates a gap between the hybridized oligonucleotides. In some instances, a polymerase (e.g., a DNA polymerase) can extend one of the oligonucleotides prior to ligation. After ligation, the ligation product is released from the analyte. In some instances, the ligation product is released using an endonuclease (e.g., RNase H). The released ligation
20 product can then be captured by capture probes (e.g., instead of direct capture of an analyte) on an array, optionally amplified, and sequenced, thus determining the location and optionally the abundance of the analyte in the biological sample.

During analysis of spatial information, sequence information for a spatial barcode associated with an analyte is obtained, and the sequence information can be used to provide
25 information about the spatial distribution of the analyte in the biological sample. Various methods can be used to obtain the spatial information. In some embodiments, specific capture probes and the analytes they capture are associated with specific locations in an array of features on a substrate. For example, specific spatial barcodes can be associated with specific array locations prior to array fabrication, and the sequences of the spatial barcodes
30 can be stored (e.g., in a database) along with specific array location information, so that each spatial barcode uniquely maps to a particular array location.

Alternatively, specific spatial barcodes can be deposited at predetermined locations in an array of features during fabrication such that at each location, only one type of spatial barcode is present so that spatial barcodes are uniquely associated with a single feature of the

array. Where necessary, the arrays can be decoded using any of the methods described herein so that spatial barcodes are uniquely associated with array feature locations, and this mapping can be stored as described above.

When sequence information is obtained for capture probes and/or analytes during
5 analysis of spatial information, the locations of the capture probes and/or analytes can be determined by referring to the stored information that uniquely associates each spatial barcode with an array feature location. In this manner, specific capture probes and captured analytes are associated with specific locations in the array of features. Each array feature location represents a position relative to a coordinate reference point (e.g., an array location, a
10 fiducial marker) for the array. Accordingly, each feature location has an “address” or location in the coordinate space of the array.

Some exemplary spatial analysis workflows are described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. See, for example, the Exemplary embodiment starting with “In some non-limiting examples of the workflows described herein,
15 the sample can be immersed...” of WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. See also, e.g., the Visium Spatial Gene Expression Reagent Kits User Guide (e.g., Rev C, dated June 2020), and/or the Visium Spatial Tissue Optimization Reagent Kits User Guide (e.g., Rev C, dated July 2020).

In some embodiments, spatial analysis can be performed using dedicated hardware
20 and/or software, such as any of the systems described in WO 2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663, or any of one or more of the devices or methods described in WO 2020/123320.

Suitable systems for performing spatial analysis can include components such as a chamber (e.g., a flow cell or sealable, fluid-tight chamber) for containing a biological sample.
25 The biological sample can be mounted for example, in a biological sample holder. One or more fluid chambers can be connected to the chamber and/or the sample holder via fluid conduits, and fluids can be delivered into the chamber and/or sample holder via fluidic pumps, vacuum sources, or other devices coupled to the fluid conduits that create a pressure gradient to drive fluid flow. One or more valves can also be connected to fluid conduits to
30 regulate the flow of reagents from reservoirs to the chamber and/or sample holder.

The systems can optionally include a control unit that includes one or more electronic processors, an input interface, an output interface (such as a display), and a storage unit (e.g., a solid state storage medium such as, but not limited to, a magnetic, optical, or other solid state, persistent, writeable and/or re-writeable storage medium). The control unit can

optionally be connected to one or more remote devices via a network. The control unit (and components thereof) can generally perform any of the steps and functions described herein. Where the system is connected to a remote device, the remote device (or devices) can perform any of the steps or features described herein. The systems can optionally include one or more detectors (e.g., CCD, CMOS) used to capture images. The systems can also optionally include one or more light sources (e.g., LED-based, diode-based, lasers) for illuminating a sample, a substrate with features, analytes from a biological sample captured on a substrate, and various control and calibration media.

The systems can optionally include software instructions encoded and/or implemented in one or more of tangible storage media and hardware components such as application specific integrated circuits. The software instructions, when executed by a control unit (and in particular, an electronic processor) or an integrated circuit, can cause the control unit, integrated circuit, or other component executing the software instructions to perform any of the method steps or functions described herein.

In some cases, the systems described herein can detect (e.g., register an image) the biological sample on the array. Exemplary methods to detect the biological sample on an array are described in PCT Application No. 2020/061064 and/or U.S. Patent Application Serial No. 16/951,854.

Prior to transferring analytes from the biological sample to the array of features on the substrate, the biological sample can be aligned with the array. Alignment of a biological sample and an array of features including capture probes can facilitate spatial analysis, which can be used to detect differences in analyte presence and/or level within different positions in the biological sample, for example, to generate a three-dimensional map of the analyte presence and/or level. Exemplary methods to generate a two- and/or three-dimensional map of the analyte presence and/or level are described in PCT Application No. 2020/053655 and spatial analysis methods are generally described in WO 2020/061108 and/or U.S. Patent Application Serial No. 16/951,864.

In some cases, a map of analyte presence and/or level can be aligned to an image of a biological sample using one or more fiducial markers, e.g., objects placed in the field of view of an imaging system which appear in the image produced, as described in WO 2020/123320, PCT Application No. 2020/061066, and/or U.S. Patent Application Serial No. 16/951,843. Fiducial markers can be used as a point of reference or measurement scale for alignment (e.g., to align a sample and an array, to align two substrates, to determine a location of a

sample or array on a substrate relative to a fiducial marker) and/or for quantitative measurements of sizes and/or distances.

I. RNA Capture Using RNA-Templated Ligation

5 (a) General Background

Although techniques such as whole genome sequencing and whole exome sequencing are available, these techniques have drawbacks in that they provide a lot of information and increase costs for an experiment. In situations where one prefers to examine a more limited number of analytes, methods herein are provided for targeted RNA capture. Capturing a
10 derivative of an analyte (e.g., a ligation product) provides enhanced specificity with respect to detection of an analyte. This is because at least two probes specific for a target are required to hybridize to the target in order to facilitate ligation and ultimate capture of the nucleic acid.

Referring to **FIG. 1**, in an exemplary embodiment of the disclosure, provided are methods for identifying a location of an analyte in a biological sample. In some instances, the
15 methods include **101** contacting a biological sample with array of spatially-barcoded capture probes. In some instances, the array is on a substrate and the array includes a plurality of capture probes, wherein a capture probe of the plurality includes: (i) a spatial barcode and (ii) a capture domain. After placing the biological sample on the array, the biological sample **102** is contacted with a first probe and a second probe, wherein the first probe and the second
20 probe each include one or more sequences that are substantially complementary to sequences of the analyte, and wherein the second probe includes a capture probe capture domain; the first probe and the second probe **103** hybridize to complementary sequences in the analyte. After hybridization a ligation product comprising the first probe and the second probe **104** is generated, and the ligation product is released from the analyte. The liberated ligation product
25 is then freed **105** to hybridize to the capture domain of a probe on the array. After capture, (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof **106** can be determined, and then one can use the determined sequence of (i) and (ii) **107** to identify the location of the analyte in the biological sample.

30 Referring to **FIG. 13**, in another non-limiting example, a biological sample is deparaffinized, stained, and imaged **1301**. After destaining and decrosslinking **1302**, probes are added to the sample and hybridize to an analyte **1303**. In some instances, the probes are DNA probes. In some instances, the probes are diribo-containing probes. Probes are ligated **1304** and then released using an endonuclease such as RNase H **1305**. Ligated probes are

captured on an array by a capture probe **1306**, extended using a polymerase **1307** and denatured **1308**. After quality control cleanup **1309**, the abundance and location of an analyte is determined.

5 A non-limiting example of the methods disclosed herein is depicted in **FIG. 6**. After a biological sample is contacted with a substrate including a plurality of capture probes and contacted with (a) a first probe **601** having a target-hybridization sequence **603** and a primer sequence **602** and (b) a second probe **604** having a target-hybridization sequence **605** and a capture domain (e.g., a poly-A sequence) **606**, the first probe **601** and a second probe **604** hybridize **610** to an analyte **607**. A ligase **621** ligates **620** the first probe to the second probe
10 thereby generating a ligation product **622**. The ligation product is released **630** from the analyte **631** by digesting the analyte using an endoribonuclease **632**. The sample is permeabilized **640** and the ligation product **641** is able to hybridize to a capture probe on the substrate.

Also provided herein are methods for identifying a location of an analyte in a
15 biological sample that includes a second probe including a pre-adenylated phosphate group at its 5' end, which enables the ligating to use a ligase that does not require adenosine triphosphate for ligase activity.

Also provided herein are methods for identifying a location of an analyte in a biological sample that includes one or more spanning probes, in addition to the first and
20 second probes. Using a spanning probe enables greater flexibility in designing RTL probes, primarily by increasing the sequences within the analyte that can be used as optional target sequences.

Also provided herein are methods for identifying a location of an analyte in a biological sample that includes optimized hybridizing, washing, and releasing steps.

25 In some embodiments, as shown in **FIG. 7**, the ligation product **701** includes a capture probe capture domain **702**, which can bind to a capture probe **703** (e.g., a capture probe immobilized, directly or indirectly, on a substrate **704**). In some embodiments, methods provided herein include contacting **705** a biological sample with a substrate **704**, wherein the capture probe **703** is affixed to the substrate (e.g., immobilized to the substrate, directly or
30 indirectly). In some embodiments, the capture probe capture domain **702** of the ligated product specifically binds to the capture domain **706**. The capture probe can also include a unique molecular identifier (UMI) **707**, a spatial barcode **708**, a functional sequence **709**, and a cleavage domain **710**.

In some embodiments, methods provided herein include permeabilization of the biological sample such that the capture probe can more easily bind to the captured ligated probe (i.e., compared to no permeabilization). In some embodiments, reverse transcription (RT) reagents can be added to permeabilized biological samples. Incubation with the RT reagents can extend the capture probes **711** to produce spatially-barcoded full-length cDNA **712** and **713** from the captured analytes (e.g., polyadenylated mRNA). Second strand reagents (e.g., second strand primers, enzymes) can be added to the biological sample on the slide to initiate second strand synthesis.

In some embodiments, cDNA can be denatured **714** from the capture probe template and transferred (e.g., to a clean tube) for amplification, and/or library construction. The spatially-barcoded, full-length cDNA can be amplified **715** via PCR prior to library construction. The cDNA can then be enzymatically fragmented and size-selected in order to optimize the cDNA amplicon size. **P5 716**, **i5 717**, **i7 718**, and **P7 719**, and can be used as sample indexes, and TruSeq Read 2 can be added via End Repair, A-tailing, Adaptor Ligation, and PCR. The cDNA fragments can then be sequenced using paired-end sequencing using TruSeq Read 1 and TruSeq Read 2 as sequencing primer sites.

(b) Probes for RNA-Templated Ligation

The methods provided herein utilize probe pairs (or sets; the terms are interchangeable). In some instances, the probe pairs are designed so that each probe hybridizes to a sequence in an analyte that is specific to the analyte (e.g., compared to the entire genome). That is, in some instances, a single probe pair can be specific to a single analyte.

In other embodiments, probes can be designed so that one of the probes of a pair is a probe that hybridizes to a specific sequence. Then, the other probe can be designed to detect a mutation of interest. Accordingly, in some instances, multiple second probes can be designed and can vary so that each binds to a specific sequence. For example, one second probe can be designed to hybridize to a wild-type sequence, and another second probe can be designed to detect a mutated sequence. Thus, in some instances, a probe set can include one first probe and two second probes (or vice versa).

On the other hand, in some instances, probes can be designed so that they cover conserved regions of an analyte. Thus, in some instances, a probe (or probe pair can hybridize to similar analytes in a biological sample (e.g., to detect conserved or similar analytes) or in different biological samples (e.g., across different species).

In some embodiments, probe sets cover all or nearly all of a genome (e.g., human genome). In instances where probe sets are designed to cover an entire genome (e.g., the human genome), the methods disclosed herein can detect analytes in an unbiased manner. In some instances, one probe oligonucleotide pair is designed to cover one analyte (e.g., transcript). In some instances, more than one probe oligonucleotide pair (e.g., a probe pair comprising a first probe and a second probe) is designed to cover one analyte (e.g., transcript). For example, at least two, three, four, five, six, seven, eight, nine, ten, or more probe sets can be used to hybridize to a single analyte. Factors to consider when designing probes is presence of variants (e.g., SNPs, mutations) or multiple isoforms expressed by a single gene. In some instances, the probe oligonucleotide pair does not hybridize to the entire analyte (e.g., a transcript), but instead the probe oligonucleotide pair hybridizes to a portion of the entire analyte (e.g., transcript).

In some instances, about 5000, 10,000, 15,000, 20,000, or more probe oligonucleotides pair (e.g., a probe pair comprising a first probe and a second probe) are used in the methods described herein. In some instances, about 20,000 probe oligonucleotides pair are used in the methods described herein

In some instances, RNA capture is targeted RNA capture. Targeted RNA capture using the methods disclosed herein allows for examination of a subset of RNA analytes from the entire transcriptome. In some embodiments, the subset of analytes includes an individual target RNA. In some embodiments, the subset of analytes includes two or more targeted RNAs. In some embodiments, the subset of analytes includes one or more mRNAs transcribed by one or more targeted genes. In some embodiments, the subset of analytes includes one or more mRNA splice variants of one or more targeted genes. In some embodiments, the subset of analytes includes non-polyadenylated RNAs in a biological sample. In some embodiments, the subset of analytes includes detection of mRNAs having one or more single nucleotide polymorphisms (SNPs) in a biological sample.

In some embodiments, the subset of analytes includes mRNAs that mediate expression of a set of genes of interest. In some embodiments, the subset of analytes includes mRNAs that share identical or substantially similar sequences, which mRNAs are translated into polypeptides having similar functional groups or protein domains. In some embodiments, the subset of analytes includes mRNAs that do not share identical or substantially similar sequences, which mRNAs are translated into proteins that do not share similar functional groups or protein domains. In some embodiments, the subset of analytes includes mRNAs that are translated into proteins that function in the same or similar biological pathways. In

some embodiments, the biological pathways are associated with a pathologic disease. For example, targeted RNA capture can detect genes that are overexpressed or underexpressed in cancer.

In some embodiments, the subset of analytes includes 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 95, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 225, about 250, about 275, about 300, about 325, about 350, about 375, about 400, about 425, about 450, about 475, about 500, about 600, about 700, about 800, about 900, or about 1000 analytes.

In some instances, the methods disclosed herein can detect the abundance and location of at least 5,000, 10,000, 15,000, 20,000, or more different analytes.

In some embodiments, the subset of analytes detected by targeted RNA capture methods provided herein includes a large proportion of the transcriptome of one or more cells. For example, the subset of analytes detected by targeted RNA capture methods provided herein can include at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or more of the mRNAs present in the transcriptome of one or more cells.

In some instances, the probes are DNA probes. In some instances, the probes are diribo-containing probes.

Additional embodiments of probe(s) and probe set(s) are described herein.

(i) First Probe

In some embodiments, the methods described herein include a first probe. As used herein, a “first probe” can refer to a probe that hybridizes to all or a portion of an analyte and can be ligated to one or more additional probes (e.g., a second probe or a spanning probe). In some embodiments, “first probe” can be used interchangeably with “first probe oligonucleotide.”

In some embodiments, the first probe includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or

analogous base pair interactions. In some embodiments, the first probe includes deoxyribonucleotides. In some embodiments, the first probe includes deoxyribonucleotides and ribonucleotides. In some embodiments, the first probe includes a deoxyribonucleic acid that hybridizes to an analyte, and includes a portion of the oligonucleotide that is not a deoxyribonucleic acid. For example, in some embodiments, the portion of the first oligonucleotide that is not a deoxyribonucleic acid is a ribonucleic acid or any other non-deoxyribonucleic acid nucleic acid as described herein. In some embodiments where the first probe includes deoxyribonucleotides, hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the first probe includes only deoxyribonucleotides and upon hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid.

In some embodiments, the method includes a first probe that includes one or more sequences that are substantially complementary to one or more sequences of an analyte. In some embodiments, a first probe includes a sequence that is substantially complementary to a first target sequence in the analyte. In some embodiments, the sequence of the first probe that is substantially complementary to the first target sequence in the analyte is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the first target sequence in the analyte.

In some embodiments, a first probe includes a sequence that is about 10 nucleotides to about 100 nucleotides (e.g., a sequence of about 10 nucleotides to about 90 nucleotides, about 10 nucleotides to about 80 nucleotides, about 10 nucleotides to about 70 nucleotides, about 10 nucleotides to about 60 nucleotides, about 10 nucleotides to about 50 nucleotides, about 10 nucleotides to about 40 nucleotides, about 10 nucleotides to about 30 nucleotides, about 10 nucleotides to about 20 nucleotides, about 20 nucleotides to about 100 nucleotides, about 20 nucleotides to about 90 nucleotides, about 20 nucleotides to about 80 nucleotides, about 20 nucleotides to about 70 nucleotides, about 20 nucleotides to about 60 nucleotides, about 20 nucleotides to about 50 nucleotides, about 20 nucleotides to about 40 nucleotides, about 20 nucleotides to about 30 nucleotides, about 30 nucleotides to about 100 nucleotides, about 30 nucleotides to about 90 nucleotides, about 30 nucleotides to about 80 nucleotides, about 30 nucleotides to about 70 nucleotides, about 30 nucleotides to about 60 nucleotides, about 30 nucleotides to about 50 nucleotides, about 30 nucleotides to about 40 nucleotides, about 40 nucleotides to about 100 nucleotides, about 40 nucleotides to about 90 nucleotides, about 40 nucleotides to about 80 nucleotides, about 40 nucleotides to about 70 nucleotides, about

40 nucleotides to about 60 nucleotides, about 40 nucleotides to about 50 nucleotides, about 50 nucleotides to about 100 nucleotides, about 50 nucleotides to about 90 nucleotides, about 50 nucleotides to about 80 nucleotides, about 50 nucleotides to about 70 nucleotides, about 50 nucleotides to about 60 nucleotides, about 60 nucleotides to about 100 nucleotides, about 60 nucleotides to about 90 nucleotides, about 60 nucleotides to about 80 nucleotides, about 60 nucleotides to about 70 nucleotides, about 70 nucleotides to about 100 nucleotides, about 70 nucleotides to about 90 nucleotides, about 70 nucleotides to about 80 nucleotides, about 80 nucleotides to about 100 nucleotides, about 80 nucleotides to about 90 nucleotides, or about 90 nucleotides to about 100 nucleotides).

10 In some embodiments, a sequence of the first probe that is substantially complementary to a sequence in the analyte includes a sequence that is about 5 nucleotides to about 50 nucleotides (e.g., about 5 nucleotides to about 45 nucleotides, about 5 nucleotides to about 40 nucleotides, about 5 nucleotides to about 35 nucleotides, about 5 nucleotides to about 30 nucleotides, about 5 nucleotides to about 25 nucleotides, about 5 nucleotides to about 20 nucleotides, about 5 nucleotides to about 15 nucleotides, about 5 nucleotides to about 10 nucleotides, about 10 nucleotides to about 50 nucleotides, about 10 nucleotides to about 45 nucleotides, about 10 nucleotides to about 40 nucleotides, about 10 nucleotides to about 35 nucleotides, about 10 nucleotides to about 30 nucleotides, about 10 nucleotides to about 25 nucleotides, about 10 nucleotides to about 20 nucleotides, about 10 nucleotides to about 15 nucleotides, about 15 nucleotides to about 50 nucleotides, about 15 nucleotides to about 45 nucleotides, about 15 nucleotides to about 40 nucleotides, about 15 nucleotides to about 35 nucleotides, about 15 nucleotides to about 30 nucleotides, about 15 nucleotides to about 25 nucleotides, about 15 nucleotides to about 20 nucleotides, about 20 nucleotides to about 50 nucleotides, about 20 nucleotides to about 45 nucleotides, about 20 nucleotides to about 40 nucleotides, about 20 nucleotides to about 35 nucleotides, about 20 nucleotides to about 30 nucleotides, about 20 nucleotides to about 25 nucleotides, about 25 nucleotides to about 50 nucleotides, about 25 nucleotides to about 45 nucleotides, about 25 nucleotides to about 40 nucleotides, about 25 nucleotides to about 35 nucleotides, about 25 nucleotides to about 30 nucleotides, about 30 nucleotides to about 50 nucleotides, about 30 nucleotides to about 45 nucleotides, about 30 nucleotides to about 40 nucleotides, about 30 nucleotides to about 35 nucleotides, about 35 nucleotides to about 50 nucleotides, about 35 nucleotides to about 45 nucleotides, about 35 nucleotides to about 40 nucleotides, about 40 nucleotides to about 50 nucleotides, about 40 nucleotides to about 45 nucleotides, or about 45 nucleotides to about 50 nucleotides).

In some embodiments, a first probe includes a functional sequence. In some embodiments, a functional sequence includes a primer sequence.

In some embodiments, a first probe includes at least two ribonucleic acid bases at the 3' end. In such cases, a second probe oligonucleotide comprises a phosphorylated nucleotide at the 5' end. In some embodiments, a first probe includes at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least ten ribonucleic acid bases at the 3' end.

As shown in **FIG. 6**, a non-limiting example of a first probe **601**, which can be referred to as an LHS probe, includes a functional sequence **602**, a sequence **603** that is substantially complementary to a first target sequence in the analyte **607**, and two ribonucleic acid bases at the 3' end.

In some embodiments, a first probe includes an auxiliary sequence that does not hybridize to an analyte. In some embodiments, the auxiliary sequence can be used to hybridize to additional probes.

(ii) Second Probe

In some embodiments, the methods described herein include a second probe. As used herein, a "second probe" can refer to a probe that hybridizes to all or a portion of an analyte and can be ligated to one or more additional probes (e.g., a first probe or a spanning probe).

In some embodiments, "second probe" can be used interchangeably with "second probe oligonucleotide." One of skill in the art will appreciate that the order of the probes is arbitrary, and thus the contents of the first probe and/or second probe as disclosed herein are interchangeable.

In some embodiments, the second probe includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some embodiments, the second probe includes deoxyribonucleotides. In some embodiments, the second probe includes deoxyribonucleotides and ribonucleotides. In some embodiments, the second probe includes a deoxyribonucleic acid that hybridizes to an analyte and includes a portion of the oligonucleotide that is not a deoxyribonucleic acid. For example, in some embodiments, the portion of the second probe that is not a deoxyribonucleic acid is a ribonucleic acid or any other non-deoxyribonucleic acid nucleic acid as described herein. In some embodiments where the second probe includes deoxyribonucleotides, hybridization of the second probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the second probe

includes only deoxyribonucleotides and upon hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid.

In some embodiments, the method includes a second probe that includes one or more sequences that are substantially complementary to one or more sequences of an analyte. In some embodiments, a second probe includes a sequence that is substantially complementary to a second target sequence in the analyte. In some embodiments, the sequence of the second probe that is substantially complementary to the second target sequence in the analyte is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the second target sequence in the analyte.

In some embodiments, a second probe includes a sequence that is about 10 nucleotides to about 100 nucleotides (e.g., a sequence of about 10 nucleotides to about 90 nucleotides, about 10 nucleotides to about 80 nucleotides, about 10 nucleotides to about 70 nucleotides, about 10 nucleotides to about 60 nucleotides, about 10 nucleotides to about 50 nucleotides, about 10 nucleotides to about 40 nucleotides, about 10 nucleotides to about 30 nucleotides, about 10 nucleotides to about 20 nucleotides, about 20 nucleotides to about 100 nucleotides, about 20 nucleotides to about 90 nucleotides, about 20 nucleotides to about 80 nucleotides, about 20 nucleotides to about 70 nucleotides, about 20 nucleotides to about 60 nucleotides, about 20 nucleotides to about 50 nucleotides, about 20 nucleotides to about 40 nucleotides, about 20 nucleotides to about 30 nucleotides, about 30 nucleotides to about 100 nucleotides, about 30 nucleotides to about 90 nucleotides, about 30 nucleotides to about 80 nucleotides, about 30 nucleotides to about 70 nucleotides, about 30 nucleotides to about 60 nucleotides, about 30 nucleotides to about 50 nucleotides, about 30 nucleotides to about 40 nucleotides, about 40 nucleotides to about 100 nucleotides, about 40 nucleotides to about 90 nucleotides, about 40 nucleotides to about 80 nucleotides, about 40 nucleotides to about 70 nucleotides, about 40 nucleotides to about 60 nucleotides, about 40 nucleotides to about 50 nucleotides, about 50 nucleotides to about 100 nucleotides, about 50 nucleotides to about 90 nucleotides, about 50 nucleotides to about 80 nucleotides, about 50 nucleotides to about 70 nucleotides, about 50 nucleotides to about 60 nucleotides, about 60 nucleotides to about 100 nucleotides, about 60 nucleotides to about 90 nucleotides, about 60 nucleotides to about 80 nucleotides, about 60 nucleotides to about 70 nucleotides, about 70 nucleotides to about 100 nucleotides, about 70 nucleotides to about 90 nucleotides, about 70 nucleotides to about 80 nucleotides, about 80 nucleotides to about 100 nucleotides, about 80 nucleotides to about 90 nucleotides, or about 90 nucleotides to about 100 nucleotides).

In some embodiments, a sequence of the second probe that is substantially complementary to a sequence in the analyte includes a sequence that is about 5 nucleotides to about 50 nucleotides (e.g., about 5 nucleotides to about 45 nucleotides, about 5 nucleotides to about 40 nucleotides, about 5 nucleotides to about 35 nucleotides, about 5 nucleotides to about 30 nucleotides, about 5 nucleotides to about 25 nucleotides, about 5 nucleotides to about 20 nucleotides, about 5 nucleotides to about 15 nucleotides, about 5 nucleotides to about 10 nucleotides, about 10 nucleotides to about 50 nucleotides, about 10 nucleotides to about 45 nucleotides, about 10 nucleotides to about 40 nucleotides, about 10 nucleotides to about 35 nucleotides, about 10 nucleotides to about 30 nucleotides, about 10 nucleotides to about 25 nucleotides, about 10 nucleotides to about 20 nucleotides, about 10 nucleotides to about 15 nucleotides, about 15 nucleotides to about 50 nucleotides, about 15 nucleotides to about 45 nucleotides, about 15 nucleotides to about 40 nucleotides, about 15 nucleotides to about 35 nucleotides, about 15 nucleotides to about 30 nucleotides, about 15 nucleotides to about 25 nucleotides, about 15 nucleotides to about 20 nucleotides, about 20 nucleotides to about 50 nucleotides, about 20 nucleotides to about 45 nucleotides, about 20 nucleotides to about 40 nucleotides, about 20 nucleotides to about 35 nucleotides, about 20 nucleotides to about 30 nucleotides, about 20 nucleotides to about 25 nucleotides, about 25 nucleotides to about 50 nucleotides, about 25 nucleotides to about 45 nucleotides, about 25 nucleotides to about 40 nucleotides, about 25 nucleotides to about 35 nucleotides, about 25 nucleotides to about 30 nucleotides, about 30 nucleotides to about 50 nucleotides, about 30 nucleotides to about 45 nucleotides, about 30 nucleotides to about 40 nucleotides, about 30 nucleotides to about 35 nucleotides, about 35 nucleotides to about 50 nucleotides, about 35 nucleotides to about 45 nucleotides, about 35 nucleotides to about 40 nucleotides, about 40 nucleotides to about 50 nucleotides, about 40 nucleotides to about 45 nucleotides, or about 45 nucleotides to about 50 nucleotides).

In some embodiments, a second probe includes a capture probe capture domain sequence. As used herein, a “capture probe capture domain” is a sequence, domain, or moiety that can bind specifically to a capture domain of a capture probe. In some embodiments, “capture domain capture domain” can be used interchangeably with “capture probe binding domain.” In some embodiments, a second probe includes a sequence from 5' to 3': a sequence that is substantially complementary to a sequence in the analyte and a capture probe capture domain.

In some embodiments, a capture probe capture domain includes a poly(A) sequence. In some embodiments, the capture probe capture domain includes a poly-uridine sequence, a

poly-thymidine sequence, or both. In some embodiments, the capture probe capture domain includes a random sequence (e.g., a random hexamer or octamer). In some embodiments, the capture probe capture domain is complementary to a capture domain in a capture probe that detects a particular target(s) of interest. In some embodiments, a capture probe capture domain blocking moiety that interacts with the capture probe capture domain is provided. In some embodiments, a capture probe capture domain blocking moiety includes a sequence that is complementary or substantially complementary to a capture probe capture domain. In some embodiments, a capture probe capture domain blocking moiety prevents the capture probe capture domain from binding the capture probe when present. In some embodiments, a capture probe capture domain blocking moiety is removed prior to binding the capture probe capture domain (e.g., present in a ligated probe) to a capture probe. In some embodiments, a capture probe capture domain blocking moiety includes a poly-uridine sequence, a poly-thymidine sequence, or both. In some embodiments, the capture probe capture domain sequence includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some embodiments, the capture probe binding domain sequence includes at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides. In some embodiments, the capture probe binding domain sequence includes at least 25, 30, or 35 nucleotides.

In some embodiments, a second probe includes a phosphorylated nucleotide at the 5' end. The phosphorylated nucleotide at the 5' end can be used in a ligation reaction to ligate the second probe to the first probe.

As shown in **FIG. 6**, a non-limiting example of a second probe **604**, which can be referred to a RHS probe, includes a sequence **605** that is substantially complementary to a second target sequence on the analyte **607** and a capture probe capture domain **606**.

In some embodiments, a second probe includes an auxiliary sequence that does not hybridize to an analyte. In some embodiments, the auxiliary sequence can be used to hybridize to additional probes.

(iii) Multiple Probes

In some embodiments, the methods of target RNA capture as disclosed herein include multiple probe oligonucleotides. In some embodiments, the methods include 2, 3, 4, or more probe oligonucleotides. In some embodiments, each of the probe oligonucleotides includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some

embodiments, each of the probe oligonucleotides includes deoxyribonucleotides. In some embodiments, each of the probe oligonucleotides includes deoxyribonucleotides and ribonucleotides.

5 In some instances, the multiple probes span different target sequences, and multiple, serial ligation steps are carried out to determine the location and abundance of an analyte.

In some instances, the methods include a first probe and multiple second probes (or vice versa) are used, with the multiple second probes hybridizing to different sequences (e.g., wild-type versus mutant sequence, different isoforms, splice variants) in order to identify the sequence of an analyte. It is appreciated that this method can be utilized to detect single
10 mutations (e.g., point mutations, SNPs, splice variants, etc.) or can multi-nucleotide mutations (e.g., insertions, deletions, etc.).

Methods provided herein may be applied to a single nucleic acid molecule or a plurality of nucleic acid molecules. A method of analyzing a sample comprising a nucleic acid molecule may comprise providing a plurality of nucleic acid molecules (e.g., RNA
15 molecules), where each nucleic acid molecule comprises a first target region (e.g., a first target sequence) and a second target region (e.g., a second target sequence), a plurality of first probe oligonucleotides, and a plurality of second probe oligonucleotides. In some cases, one or more target regions of nucleic acid molecules of the plurality of nucleic acid molecules may comprise the same sequence. The first and second target regions (e.g., the first and
20 second target sequences) of a nucleic acid molecule of the plurality of nucleic acid molecules may be adjacent to one another.

(iv) First Probe Having a Linker Sequence

25 Also provided herein are methods for identifying a location of an analyte in a biological sample where the method includes a first probe that includes a linker and a second probe. Using a pair of probes where the first probe includes a linker sequence enables greater flexibility in designing RTL probes, primarily by increasing the sequences within the analyte that can be used as optional target sequences.

30 As used herein, a “linker sequence” can refer to one or more nucleic acids sequences on a probe (e.g., a first probe, a second probe, or a spanning probe that are disposed between sequences that hybridize to an analyte, sequences that link together the analyte specific sequences of a probe). In some embodiments, a linker includes a sequence that is not substantially complementary to either the sequence of the target analyte or to the analyte

specific sequences of a first probe, a second probe, or a spanning probe. In some embodiments, the linker sequence includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides, where the sequence within the linker is not substantially complementary to the target analyte or the analyte specific sequences of a first probe, a second probe, or a spanning probe.

In some embodiments where a first and/or a second probe include a linker sequence, the linker sequence can include a total of about 10 nucleotides to about 100 nucleotides, or any of the subranges described herein.

In some embodiments, a linker sequence includes a barcode sequence that serves as a proxy for identifying the analyte. In some embodiments, the barcode sequence is a sequence that is at least 70% identical (e.g., at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, or at least 99% identical) to a sequence in the analyte. In some embodiments where a linker sequence includes a barcode sequence, the barcode sequence is located 5' to the linker sequence. In some embodiments where a linker sequence includes a barcode sequence, the barcode sequence is located 3' to the linker sequence. In some embodiments, the barcode sequence is disposed between two linker sequences. In such cases, the two linker sequences flanking the barcode sequence can be considered to be a part of the same linker sequence.

In some embodiments where a first and/or a second probe include a linker sequence, the linker sequence can include ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides.

A non-limiting example of a method for identifying a location of an analyte in a biological sample includes a first probe that includes a linker sequence and a second probe comprising: (a) contacting the biological sample with a substrate including a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes a capture domain and a spatial barcode; (b) contacting the biological sample with a first probe and a second probe, wherein a portion of the first probe and a portion of the second probe are substantially complementary to adjacent sequences of the analyte, wherein the first probe includes: (i) a first sequence that is substantially complementary to a first target sequence of the analyte; (ii) a linker sequence; (iii) a second sequence that is substantially complementary to a second target sequence of the analyte; and wherein the second probe includes a sequence that is substantially complementary to a third target sequence of the analyte and a capture probe capture domain that is capable of binding to a capture domain of a capture probe; (c) hybridizing the first probe and the second probe to the analyte; (d) ligating the first probe and

the second probe, thereby creating a ligation product; (e) releasing the ligation product from the analyte; (f) hybridizing the capture probe binding domain to a capture domain; and (g) determining (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample.

A non-limiting example of a method for identifying a location of an analyte in a biological sample where the method includes a first probe that includes a linker and a second probe can include the components as shown **FIG. 8**. A first probe **801** includes a functional sequence **802**, a first sequence **803** that sequence that is substantially complementary to a first target sequence **804** of the analyte, a linker sequence **805**; and a second sequence **806** that is substantially complementary to a second target sequence **807** of the analyte. A second probe **808** includes a sequence **809** that is substantially complementary to a third target sequence **810** of the analyte and a capture probe capture domain **811** that is capable of binding to a capture domain of a capture probe.

1) First probe

In some embodiments where first probe includes a linker sequence, the first probe includes a first sequence that is substantially complementary to a first target sequence of the analyte, a linker sequence, and a second sequence that is substantially complementary to second target sequence of the analyte. In some embodiments, a first probe includes from 5' to 3': a first sequence that is substantially complementary to a first target sequence of the analyte, a linker sequence, and a second sequence that is substantially complementary to second target sequence of the analyte.

In some embodiments where a first probe includes a linker sequence, the first probe includes a functional sequence. In some embodiments, a first probe includes a functional sequence, a first sequence that is substantially complementary to a first target sequence of the analyte, a linker sequence, and a second sequence that is substantially complementary to second target sequence of the analyte. In some embodiments, the functional sequence includes a primer sequence.

In some embodiments where a first probe includes a linker sequence, a first probe includes at least two ribonucleic acid bases at the 3' end. In some embodiments, a first probe includes at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least ten ribonucleic acid bases at the 3' end.

In some embodiments where a first probe includes a linker, the first probe includes a sequence that is about 10 nucleotides to about 300 nucleotides (e.g., a sequence of about 10 nucleotides to about 300 nucleotides, about 10 nucleotides to about 250 nucleotides, about 10 nucleotides to about 200 nucleotides, about 10 nucleotides to about 150 nucleotides, about 10 nucleotides to about 100 nucleotides, about 10 nucleotides to about 50 nucleotides, about 50 nucleotides to about 300 nucleotides, about 50 nucleotides to about 250 nucleotides, about 50 nucleotides to about 200 nucleotides, about 50 nucleotides to about 150 nucleotides, about 50 nucleotides to about 100 nucleotides, about 100 nucleotides to about 300 nucleotides, about 100 nucleotides to about 250 nucleotides, about 100 nucleotides to about 200 nucleotides, about 100 nucleotides to about 150 nucleotides, about 150 nucleotides to about 300 nucleotides, about 150 nucleotides to about 250 nucleotides, about 150 nucleotides to about 200 nucleotides, about 200 nucleotides to about 300 nucleotides, about 200 nucleotides to about 250 nucleotides, or about 250 nucleotides to about 300 nucleotides).

In some embodiments where a first probe includes a linker sequence, the first probe includes a first sequence that is substantially complementary to a first target sequence of the analyte. In some embodiments, the first sequence of the first probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the first target sequence in the analyte. In some embodiments, the first sequence of the first probe that is substantially complementary to a first target sequence can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein.

In some embodiments where a first probe includes a linker, the first probe includes a second sequence that is substantially complementary to a second target sequence of the analyte. In some embodiments, the second sequence of the first probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the second target sequence in the analyte. In some embodiments, the second sequence of the first probe that is substantially complementary to a second target sequence can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein.

In some embodiments, a first probe that includes a linker sequence includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some

embodiments, the first probe that includes a linker sequence includes deoxyribonucleotides. In some embodiments, the first probe that includes a linker sequence includes deoxyribonucleotides and ribonucleotides. In some embodiments where the first probe that includes a linker sequence includes deoxyribonucleotides, hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the first probe that includes a linker sequence includes only deoxyribonucleotides and upon hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid.

2) Second probe

In some embodiments where a first probe includes a linker sequence, a second probe includes a sequence that is substantially complementary to a third target sequence of the analyte and a capture probe capture domain that is capable of binding to a capture domain of a capture probe. In some embodiments where a first probe includes a linker, a second probe includes a sequence that is about 10 nucleotides to about 100 nucleotides, or any of the subranges described herein.

In some embodiments where a first probe includes a linker, the sequence of the second probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the third target sequence in the analyte. In some embodiments, the sequence of the second probe that is substantially complementary to a third target sequence can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein.

In some embodiments where a first probe includes a linker sequence, a first target sequence is not adjacent to a second target sequence. For example, the first target sequence and second target sequences are located on different exons of the same mRNA molecule. In another example, the first target sequence and the second target sequence are located on the same exon of the same mRNA molecule but are not adjacent. In some instances, the first probe and the second probe hybridize to sequences that at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or more nucleotides apart.

In some embodiments where a first probe includes a linker sequence, a second target sequence is directly adjacent to a third target sequence.

In some embodiments where a first probe includes a linker sequence, the second probe includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some

embodiments, the second probe includes deoxyribonucleotides. In some embodiments, the second probe includes deoxyribonucleotides and ribonucleotides. In some embodiments where the second probe includes deoxyribonucleotides, hybridization of the second probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the second probe
 5 includes only deoxyribonucleotides and upon hybridization of the second probe to the mRNA molecule results in a DNA:RNA hybrid.

(v) Second Probe Having a Linker

Also provided herein are methods for identifying a location of an analyte in a
 10 biological sample where the method includes a first probe and a second probe that includes a linker sequence. Using a pair of probes where the second probe includes a linker sequence enables greater flexibility in designing RTL probes, primarily by increasing the sequences within the analyte that can be used as optional target sequences.

A non-limiting example of a method for identifying a location of an analyte in a
 15 biological sample where the method includes a first probe and a second probe that includes a linker includes: (a) contacting the biological sample with a substrate including a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes a capture domain and a spatial barcode; (b) contacting the biological sample with a first probe and a second probe, wherein a portion of the first probe and a portion of the second probe are
 20 substantially complementary to adjacent sequences of the analyte, wherein the first probe includes a sequence that is substantially complementary to a first target sequence of the analyte, wherein the second probe includes: (i) a first sequence that is substantially complementary to a second target sequence of the analyte; (ii) a linker sequence; (iii) a second sequence that is substantially complementary to a third target sequence of the analyte;
 25 and (iv) a capture probe binding domain that is capable of binding to a capture domain of a capture probe; (c) hybridizing the first probe and the second probe to the analyte; (d) ligating the first probe and the second probe, thereby creating a ligation product; (e) releasing the ligation product from the analyte; (f) hybridizing the capture probe binding domain to a capture domain; and (g) determining (i) all or a part of the sequence of the ligation product
 30 specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample.

A non-limiting example of a method for identifying a location of an analyte in a biological sample where the method includes a first probe and a second probe that includes a

linker can include the components as shown **FIG. 9**. A first probe **901** includes a functional sequence **902** and a sequence **903** that is substantially complementary to a first target sequence **904** of the analyte. A second probe **905** includes a first sequence **906** that is substantially complementary to a second target sequence **907** of the analyte, a linker sequence **908**; and a second sequence **909** that is substantially complementary to a second target sequence **910** of the analyte, and a capture probe capture domain **911** that is capable of binding to a capture domain of a capture probe.

1) First Probe

10 In some embodiments where a second probe includes a linker sequence, a first probe includes a sequence that is substantially complementary to a first target sequence of the analyte.

In some embodiments where a second probe includes a linker sequence, a first probe includes a functional sequence. In some embodiments, a first probe includes a functional
15 sequence and a sequence that is substantially complementary to a first target sequence of the analyte. In some embodiments, the functional sequence includes a primer sequence. In some embodiments, the first probe oligonucleotide includes from 5' to 3': a functional sequence, and a sequence that is substantially complementary to a first target sequence.

In some embodiments where a linker is on a second probe, a first probe includes a
20 sequence that is about 10 nucleotides to about 100 nucleotides, or any of the subranges described herein.

In some embodiments where the second probe includes a linker sequence, a sequence of a first probe that is substantially complementary to a first target sequence of the analyte is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%,
25 at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the first target sequence in the analyte. In some embodiments, the sequence that is substantially complementary to a first target sequence can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein.

In some embodiments where a second probe includes a linker sequence, a first probe
30 includes at least two ribonucleic acid bases at the 3' end. In some embodiments, a first probe includes at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least ten ribonucleic acid bases at the 3' end.

In some embodiments where a second probe includes a linker sequence, the first probe includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are

capable of participating in Watson-Crick type or analogous base pair interactions. In some embodiments, the first probe includes deoxyribonucleotides. In some embodiments, the first probe includes deoxyribonucleotides and ribonucleotides. In some embodiments where the first probe includes deoxyribonucleotides, hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the first probe includes only deoxyribonucleotides and upon hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid.

2) Second Probe

In some embodiments where a linker is on a second probe, the second probe includes (i) a first sequence that is substantially complementary to a second target sequence of the analyte; (ii) a linker sequence (e.g., any of the exemplary linker sequences described herein); (iii) a second sequence that is substantially complementary to third target sequence of the analyte; and (iv) a capture probe capture domain (e.g., any of the exemplary capture probe capture domains described herein) that is capable of binding to a capture domain of a capture probe.

In some embodiments where a second probe includes a linker sequence, the second probe includes a sequence that is about 10 nucleotides to about 300 nucleotides, or any of the subranges described herein.

In some embodiments where a second probe includes a linker sequence, the second probe includes a first sequence that is substantially complementary to a second target sequence of the analyte. In some embodiments, the first sequence of the second probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the second target sequence in the analyte. In some embodiments, the first sequence of the second probe that is substantially complementary to a second target sequence can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein.

In some embodiments where a second probe includes a linker sequence, the second probe includes a second sequence that is substantially complementary to a third target sequence of the analyte. In some embodiments, the second sequence of the second probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the third target sequence in the analyte. In some embodiments, the

second sequence of the second probe that is substantially complementary to a third target sequence can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein.

5 In some embodiments where a second probe includes a linker sequence, a second target sequence is not adjacent to a third target sequence in the mRNA molecule. For example, the second target sequence and third target sequence are located on different exons of the same mRNA molecule. In another example, the second target sequence and the third target sequence are located on the same exon of the same mRNA molecule but are not adjacent.

10 In some embodiments where a second probe includes a linker sequence, a first target sequence is directly adjacent to a second target sequence.

In some embodiments, a second probe that includes a linker sequence includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some
15 embodiments, the second probe that includes a linker sequence includes deoxyribonucleotides. In some embodiments, the second probe that includes a linker sequence includes deoxyribonucleotides and ribonucleotides. In some embodiments where the second probe that includes a linker sequence includes deoxyribonucleotides, hybridization of the second probe to the mRNA molecule results in a DNA:RNA hybrid. In some
20 embodiments, the second probe that includes a linker sequence includes only deoxyribonucleotides and upon hybridization of the second probe to the mRNA molecule results in a DNA:RNA hybrid.

(vii) Probe Combination with Linkers on the First Probe and the Second Probe

25 Also provided herein are methods for identifying a location of an analyte in a biological sample where the method includes a first probe that includes a linker sequence and a second probe that includes a linker sequence. Using a pair of probes where the first probe and second probe each include a linker sequence enables greater flexibility in designing RTL probes, primarily by increasing the sequences within the analyte that can be used as optional
30 target sequences.

(c) Probe Combinations including a First Probe, a Second Probe and a Spanning Probe

Also provided herein are methods for identifying a location of an analyte in a biological sample where the method includes a first probe, a spanning probe, and a second probe. Using a spanning probe enables greater flexibility in designing RTL probes, primarily by increasing the sequences within the analyte that can be used as optional target sequences.

- 5 In some cases, using a spanning probe can also be used to interrogate the variants (e.g., splice variants) that span greater distances that can be interrogated using a first or second probe with a linker sequence.

- A non-limiting example of a method for identifying a location of an analyte in a biological sample where the method includes a first probe, a spanning probe, and a second probe, includes: (a) contacting the biological sample with a substrate including a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes a capture domain and a spatial barcode; (b) contacting the biological sample with a first probe, a second probe, and one or more spanning probes, wherein the first probe is substantially complementary to a first portion of the analyte, wherein the second probe is substantially complementary to a second portion of the analyte and further includes a capture probe binding domain, and wherein the spanning probe includes: (i) a first sequence that is substantially complementary to a first target sequence of the analyte, and (ii) a second sequence that is substantially complementary to a second target sequence of the analyte; (c) hybridizing the first probe, the second probe, and the spanning probe to the analyte; (d) ligating the first probe, the one or more spanning probes, and the second probe, thereby creating a ligation product that is substantially complementary to the analyte; (e) releasing the ligation product from the analyte; (f) hybridizing the capture probe binding domain to a capture domain; and (g) determining (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample.

- A non-limiting example of a method for identifying a location of an analyte in a biological sample where the method includes a first probe, a second probe, and a spanning probe can include the components as shown **FIG. 10**. A first probe **1001** includes a functional sequence **1002**, a sequence **1003** that is substantially complementary to a first portion **1004** of the analyte. A spanning probe **1005** includes a first sequence **1006** that is substantially complementary to a first target sequence **1007** of the analyte, a linker sequence **1008**, and a second sequence **1009** that is substantially complementary to a second target sequence **1010** of the analyte. A second probe **1011** includes a sequence **1012** that is substantially

complementary to a second portion **1013** of the analyte and a capture probe capture domain **1014**.

5 (i) First Probe

In some embodiments where the method includes a spanning probe, a first probe includes a sequence that is substantially complementary to a first portion of the analyte. In some embodiments, a sequence of the first probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to a first portion of the analyte. In some embodiments, the sequence of the first probe that is substantially complementary to a first portion of the analyte can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein. In some embodiments, the first probe includes a functional sequence. In some embodiments, the functional sequence is a primer sequence.

In some embodiments where the method includes a spanning probe, a first probe includes at least two ribonucleic acid bases at the 3' end. In some embodiments, a first probe includes at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least ten ribonucleic acid bases at the 3' end.

20 In some embodiments where the method includes a spanning probe, a first probe includes from 5' to 3': a functional sequence, a sequence that is substantially complementary to a first portion of the analyte, and two or more ribonucleic acid bases.

In some embodiments where the method includes a spanning probe, a first probe includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some embodiments, the first probe includes deoxyribonucleotides. In some embodiments, the first probe includes deoxyribonucleotides and ribonucleotides. In some embodiments where the first probe includes deoxyribonucleotides, hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the first probe includes only deoxyribonucleotides and upon hybridization of the first probe to the mRNA molecule results in a DNA:RNA hybrid.

(ii) Spanning Probe

In some embodiments where the method includes a spanning probe, the spanning probe includes a first sequence that is substantially complementary to a first target sequence of the analyte, and a second sequence that is substantially complementary to a second target sequence of the analyte. In some embodiments, the spanning probe includes a first sequence that is substantially complementary to a first target sequence of the analyte, a functional sequence, and a second sequence that is substantially complementary to a second target sequence of the analyte. In some embodiments, a spanning probe includes from 5' to 3': a first sequence, a functional sequence, and a second sequence. In some embodiments, the functional sequence is a linker sequence. The linker sequence can include a total of about 10 nucleotides to about 100 nucleotides, or any of the subranges described herein.

In some embodiments, the functional sequence includes a barcode sequence. In some embodiments, the spanning probe can include a linker and a barcode sequence. In such cases, linker sequences can flank the barcode, the barcode can be 5' to a linker sequence, or the barcode can be 3' to a linker sequence. In some embodiments, a barcode sequence is flanked by a 5' linker sequence (e.g., any of the exemplary linker sequences described herein) and a 3' linker sequence (e.g., any of the exemplary linker sequences described herein).

In some embodiments, the spanning probe includes from 5' to 3': a first sequence, a 5' linker sequence, a barcode, a 3' linker sequence, and a second sequence.

In some embodiments, the spanning probes includes a sequence that is about 10 nucleotides to about 300 nucleotides or any of the subranges described herein.

In some embodiments, a first sequence of the spanning probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the first target sequence of the analyte. In some embodiments, a second sequence of the spanning probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to the second target sequence of the analyte.

In some embodiments, the first sequence of the spanning probe and the second sequence of the spanning probe are substantially complementary to sequences within the same exon.

In some embodiments, the first target sequence of the analyte and the second target of the analyte are located within the same exon. In such cases, the first target sequence and the second target sequence are not directly adjacent.

In some embodiments, the first sequence of the spanning probe and the second sequence of the spanning probe are substantially complementary to sequences within the different exons of the same gene. In some embodiments, the first target sequence of the analyte and the second target sequence of the analyte are located on different exons of the same gene.

In some embodiments, the spanning probe includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some embodiments, the spanning probe includes deoxyribonucleotides. In some embodiments, the spanning probe includes deoxyribonucleotides and ribonucleotides. In some embodiments where the spanning probe includes deoxyribonucleotides, hybridization of the spanning probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the spanning probe includes only deoxyribonucleotides and upon hybridization of the spanning probe to the mRNA molecule results in a DNA:RNA hybrid.

15

(iii) Second probe

In some embodiments where the method includes a spanning probe, a second probe includes a sequence that is substantially complementary to a second portion of the analyte and a capture probe capture domain (e.g., any of the exemplary capture probe capture domains described herein).

20

In some embodiments where the method includes a spanning probe, a sequence of the second probe is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to a second portion of the analyte. In some embodiments, the sequence of the second probe that is substantially complementary to a second portion of the analyte can include a sequence that is about 5 nucleotides to about 50 nucleotides, or any of the subranges described herein.

25

In some embodiments where the method includes a spanning probe, the first portion of the analyte is directly adjacent to the first target sequence, and/or the second portion of the analyte is directly adjacent to the second target sequence. In such cases, the sequence of the first probe is ligated to the first sequence of the spanning probe, and the sequence of the second probe is ligated to the second sequence of the spanning probe. In some embodiments, the spanning probe includes at least two ribonucleic acid based at the 3' end, the first probe includes at least two ribonucleic acids at the 3' end, or both. In some embodiments, the

30

spanning probe includes a phosphorylated nucleotide at the 5' end, the second probe includes a phosphorylated nucleotide at the 5' end, or both.

In some embodiments where the method includes a spanning probe, a second probe includes ribonucleotides, deoxyribonucleotides, and/or synthetic nucleotides that are capable of participating in Watson-Crick type or analogous base pair interactions. In some embodiments, the second probe includes deoxyribonucleotides. In some embodiments, the second probe includes deoxyribonucleotides and ribonucleotides. In some embodiments where the second probe includes deoxyribonucleotides, hybridization of the second probe to the mRNA molecule results in a DNA:RNA hybrid. In some embodiments, the second probe includes only deoxyribonucleotides and upon hybridization of the second probe to the mRNA molecule results in a DNA:RNA hybrid.

(iv) Probe Combinations including a First Probe, a Second Probe, and Multiple Spanning Probes

Also provided herein are methods for identifying a location of an analyte in a biological sample where the method includes a first probe, at least two spanning probes, and a second probe. Using two or more spanning probes enables greater flexibility in designing RTL probes, primarily by increasing the sequences within the analyte that can be used as optional target sequences. In some cases, using two or more spanning probe can also be used to interrogate the variants (e.g., splice variants) that span greater distances that can be interrogated using one spanning probe.

A non-limiting example of a method for identifying a location of an analyte in a biological sample where the method includes a first probe, two or more spanning probes, and a second probe, includes: (a) contacting the biological sample with a substrate including a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes a capture domain and a spatial barcode; (b) contacting the biological sample with a first probe, a second probe, and two spanning probes, wherein the first probe is substantially complementary to a first portion of the analyte, wherein the second probe is substantially complementary to a second portion of the analyte and further includes a capture probe binding domain, and wherein the first spanning probe includes: (i) a first sequence that is substantially complementary to a first target sequence of the analyte, and (ii) a second sequence that is substantially complementary to a second target sequence of the analyte; and the second spanning probe includes (i) a third sequence that is substantially complementary to a third target sequence of the analyte, and (ii) a fourth sequence that is substantially

complementary to a fourth target sequence of the analyte; (c) hybridizing the first probe, the second probe, and the spanning probe to the analyte; (d) ligating the first probe, the one or more spanning probes, and the second probe, thereby creating a ligation product that is substantially complementary to the analyte; (e) releasing the ligation product from the analyte; (f) hybridizing the capture probe binding domain to a capture domain; and (g) determining (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample.

10 In some embodiments, the methods that include one or more spanning probes include at least two, at least three, at least four, at least five, or more spanning probes. In such cases, the one or more spanning probes includes (i) a third sequence that is substantially complementary to a third target sequence of the analyte, and (ii) a fourth sequence that is substantially complementary to a fourth target sequence of the analyte.

15 In some embodiments where the method includes two (or more) spanning probes, the first target sequence is located in a first exon, the second target sequence is located in a second exon, and the third target sequence and the fourth target sequence are located in a third exon. In some embodiments where the method includes two (or more) spanning probes, the first target sequence is located in a first exon, the second target sequence is located in a second exon, and the third target sequence is located in a third exon, and the fourth target sequence is located in a fourth exon. In some embodiments where the method includes two (or more) spanning probes, the first target sequence and the second target sequences are located in a first exon, and the third target sequence and the fourth target sequence are located in a second exon. In some embodiments where the method includes two (or more) spanning probes, the first target sequence and the second target sequences are located in a first exon, and the third target sequence is located in a second exon, and the fourth target sequence is located in a third exon.

25 In some embodiments, where the methods include two (or more) spanning probes, the method includes ligating: the first probe to the spanning probe, the spanning probe to the one or more additional spanning probes, and the one or more additional spanning probes spanning oligonucleotide to the second probe, thereby creating a ligation product that includes one or more sequences that are substantially complementary to the analyte. In some embodiments, where the methods include two (or more) spanning probes, the method includes ligating: the first probe to the one or more additional spanning probes, the one or more additional

spanning probes to the spanning probe, and the spanning probe to the second probe, thereby creating a ligation product that includes one or more sequences that are substantially complementary to the analyte.

In some embodiments, each additional spanning probe can include a functional sequence (e.g., any of the functional sequence described herein). For example, each additional spanning probe can include a linker sequence (e.g., any of the exemplary linker sequences described herein). In another example, each additional spanning probe can include a barcode sequence (e.g., any of the exemplary barcode sequences described herein) and a linker sequence (e.g., any of the linker sequences described herein). In some embodiments where an additional spanning probe includes a barcode and a linker, a linker sequences can flank the barcode, the barcode can be 5' to a linker sequence, or the barcode can be 3' to a linker sequence. In some embodiments, a barcode sequence is flanked by a 5' linker sequence (e.g., any of the exemplary linker sequences described herein) and a 3' linker sequence (e.g., any of the exemplary linker sequences described herein). In some embodiments, an additional spanning probe can include from 5' to 3': a first sequence, a 5' linker sequence, a barcode, a 3' linker sequence, and a second sequence.

(d) Pre-Hybridization Methods

(i) Imaging and Staining

Prior to addition of the probes, in some instances, biological samples can be stained using a wide variety of stains and staining techniques. In some instances, the biological sample is a section on a slide (e.g., a 10 μm section). In some instances, the biological sample is dried after placement onto a glass slide. In some instances, the biological sample is dried at 42°C. In some instances, drying occurs for about 1 hour, about 2, hours, about 3 hours, or until the sections become transparent. In some instances, the biological sample can be dried overnight (e.g., in a desiccator at room temperature).

In some embodiments, a sample can be stained using any number of biological stains, including but not limited to, acridine orange, Bismarck brown, carmine, coomassie blue, cresyl violet, DAPI, eosin, ethidium bromide, acid fuchsin, hematoxylin, Hoechst stains, iodine, methyl green, methylene blue, neutral red, Nile blue, Nile red, osmium tetroxide, propidium iodide, rhodamine, or safranin. In some instances, the methods disclosed herein include imaging the biological sample. In some instances, imaging the sample occurs prior to deaminating the biological sample. In some instances, the sample can be stained using known staining techniques, including Can-Grunwald, Giemsa, hematoxylin and eosin (H&E),

Jenner's, Leishman, Masson's trichrome, Papanicolaou, Romanowsky, silver, Sudan, Wright's, and/or Periodic Acid Schiff (PAS) staining techniques. PAS staining is typically performed after formalin or acetone fixation. In some instances, the stain is an H&E stain.

In some embodiments, the biological sample can be stained using a detectable label (e.g., radioisotopes, fluorophores, chemiluminescent compounds, bioluminescent compounds, and dyes) as described elsewhere herein. In some embodiments, a biological sample is stained using only one type of stain or one technique. In some embodiments, staining includes biological staining techniques such as H&E staining. In some embodiments, staining includes identifying analytes using fluorescently-conjugated antibodies. In some
5
10
embodiments, a biological sample is stained using two or more different types of stains, or two or more different staining techniques. For example, a biological sample can be prepared by staining and imaging using one technique (e.g., H&E staining and brightfield imaging), followed by staining and imaging using another technique (e.g., IHC/IF staining and fluorescence microscopy) for the same biological sample.

In some embodiments, biological samples can be destained. Methods of destaining or discoloring a biological sample are known in the art, and generally depend on the nature of the stain(s) applied to the sample. For example, H&E staining can be destained by washing the sample in HCl, or any other acid (e.g., selenic acid, sulfuric acid, hydroiodic acid, benzoic acid, carbonic acid, malic acid, phosphoric acid, oxalic acid, succinic acid, salicylic acid,
15
20
tartaric acid, sulfurous acid, trichloroacetic acid, hydrobromic acid, hydrochloric acid, nitric acid, orthophosphoric acid, arsenic acid, selenous acid, chromic acid, citric acid, hydrofluoric acid, nitrous acid, isocyanic acid, formic acid, hydrogen selenide, molybdic acid, lactic acid, acetic acid, carbonic acid, hydrogen sulfide, or combinations thereof). In some embodiments, destaining can include 1, 2, 3, 4, 5, or more washes in an acid (e.g., HCl). In some
25
embodiments, destaining can include adding HCl to a downstream solution (e.g., permeabilization solution). In some embodiments, destaining can include dissolving an enzyme used in the disclosed methods (e.g., pepsin) in an acid (e.g., HCl) solution. In some embodiments, after destaining hematoxylin with an acid, other reagents can be added to the destaining solution to raise the pH for use in other applications. For example, SDS can be
30
added to an acid destaining solution in order to raise the pH as compared to the acid destaining solution alone. As another example, in some embodiments, one or more immunofluorescence stains are applied to the sample via antibody coupling. Such stains can be removed using techniques such as cleavage of disulfide linkages via treatment with a reducing agent and detergent washing, chaotropic salt treatment, treatment with antigen

retrieval solution, and treatment with an acidic glycine buffer. Methods for multiplexed staining and destaining are described, for example, in Bolognesi et al., J. Histochem. Cytochem. 2017; 65(8): 431-444, Lin et al., Nat Commun. 2015; 6:8390, Pirici et al., J. Histochem. Cytochem. 2009; 57:567-75, and Glass et al., J. Histochem. Cytochem. 2009; 57:899-905, the entire contents of each of which are incorporated herein by reference.

In some embodiments, immunofluorescence or immunohistochemistry protocols (direct and indirect staining techniques) can be performed as a part of, or in addition to, the exemplary spatial workflows presented herein. For example, tissue sections can be fixed according to methods described herein. The biological sample can be transferred to an array (e.g., capture probe array), wherein analytes (e.g., proteins) are probed using immunofluorescence protocols. For example, the sample can be rehydrated, blocked, and permeabilized (3X SSC, 2% BSA, 0.1% Triton X, 1 U/μl RNase inhibitor for 10 minutes at 4°C) before being stained with fluorescent primary antibodies (1:100 in 3XSSC, 2% BSA, 0.1% Triton X, 1 U/μl RNase inhibitor for 30 minutes at 4°C). The biological sample can be washed, coverslipped (in glycerol + 1 U/μl RNase inhibitor), imaged (e.g., using a confocal microscope or other apparatus capable of fluorescent detection), washed, and processed according to analyte capture or spatial workflows described herein.

In some instances, a glycerol solution and a cover slip can be added to the sample. In some instances, the glycerol solution can include a counterstain (e.g., DAPI).

As used herein, an antigen retrieval buffer can improve antibody capture in IF/IHC protocols. An exemplary protocol for antigen retrieval can be preheating the antigen retrieval buffer (e.g., to 95°C), immersing the biological sample in the heated antigen retrieval buffer for a predetermined time, and then removing the biological sample from the antigen retrieval buffer and washing the biological sample.

In some embodiments, optimizing permeabilization can be useful for identifying intracellular analytes. Permeabilization optimization can include selection of permeabilization agents, concentration of permeabilization agents, and permeabilization duration. Tissue permeabilization is discussed elsewhere herein.

In some embodiments, blocking an array and/or a biological sample in preparation of labeling the biological sample decreases nonspecific binding of the antibodies to the array and/or biological sample (decreases background). Some embodiments provide for blocking buffers/blocking solutions that can be applied before and/or during application of the label, wherein the blocking buffer can include a blocking agent, and optionally a surfactant and/or a salt solution. In some embodiments, a blocking agent can be bovine serum albumin (BSA),

serum, gelatin (e.g., fish gelatin), milk (e.g., non-fat dry milk), casein, polyethylene glycol (PEG), polyvinyl alcohol (PVA), or polyvinylpyrrolidone (PVP), biotin blocking reagent, a peroxidase blocking reagent, levamisole, Carnoy's solution, glycine, lysine, sodium borohydride, pontamine sky blue, Sudan Black, trypan blue, FITC blocking agent, and/or
 5 acetic acid. The blocking buffer/blocking solution can be applied to the array and/or biological sample prior to and/or during labeling (e.g., application of fluorophore-conjugated antibodies) to the biological sample.

(ii) Preparation of Sample for Application of Probes

10 In some instances, the biological sample is deparaffinized. Deparaffinization can be achieved using any method known in the art. For example, in some instances, the biological samples is treated with a series of washes that include xylene and various concentrations of ethanol. In some instances, methods of deparaffinization include treatment of xylene (e.g., three washes at 5 minutes each). In some instances, the methods further include treatment
 15 with ethanol (e.g., 100% ethanol, two washes 10 minutes each; 95% ethanol, two washes 10 minutes each; 70% ethanol, two washes 10 minutes each; 50% ethanol, two washes 10 minutes each). In some instances, after ethanol washes, the biological sample can be washed with deionized water (e.g., two washes for 5 minutes each). It is appreciated that one skilled in the art can adjust these methods to optimize deparaffinization.

20 In some instances, the biological sample is decrosslinked. In some instances, the biological sample is decrosslinked in a solution containing TE buffer (comprising Tris and EDTA). In some instances, the TE buffer is basic (e.g., at a pH of about 9). In some instances, decrosslinking occurs at about 50°C to about 80°C. In some instances, decrosslinking occurs at about 70°C. In some instances, decrosslinking occurs for about 1
 25 hour at 70°C. Just prior to decrosslinking, the biological sample can be treated with an acid (e.g., 0.1M HCl for about 1 minute). After the decrosslinking step, the biological sample can be washed (e.g., with 1x PBST).

In some instances, the methods of preparing a biological sample for probe application include permeabilizing the sample. In some instances, the biological sample is permeabilized
 30 using a phosphate buffer. In some instances, the phosphate buffer is PBS (e.g., 1x PBS). In some instances, the phosphate buffer is PBST (e.g., 1x PBST). In some instances, the permeabilization step is performed multiple times (e.g., 3 times at 5 minutes each).

In some instances, the methods of preparing a biological sample for probe application include steps of equilibrating and blocking the biological sample. In some instances,

equilibrating is performed using a pre-hybridization (pre-Hyb) buffer. In some instances, the pre-Hyb buffer is RNase-free. In some instances, the pre-Hyb buffer contains no bovine serum albumin (BSA), solutions like Denhardt's, or other potentially nuclease-contaminated biological materials.

5 In some instances, the equilibrating step is performed multiple times (e.g., 2 times at 5 minutes each; 3 times at 5 minutes each). In some instances, the biological sample is blocked with a blocking buffer. In some instances, the blocking buffer includes a carrier such as tRNA, for example yeast tRNA such as from brewer's yeast (e.g., at a final concentration of 10–20 µg/mL). In some instances, blocking can be performed for 5, 10, 15, 20, 25, or 30
10 minutes.

Any of the foregoing steps can be optimized for performance. For example, one can vary the temperature. In some instances, the pre-hybridization methods are performed at room temperature. In some instances, the pre-hybridization methods are performed at 4°C (in some instances, varying the timeframes provided herein).

15

(e) Hybridizing the Probes

In some embodiments, the methods of targeted RNA capture provided herein include hybridizing a first probe oligonucleotide and a second probe oligonucleotide (e.g., a probe pair). In some instances, the first and second probe oligonucleotides each include sequences
20 that are substantially complementary to one or more sequences (e.g., one or more target sequences) of an analyte of interest. In some embodiments, the first probe and the second probe bind to complementary sequences that are completely adjacent (i.e., no gap of nucleotides) to one another or are on the same transcript.

In some instances, the methods include hybridization of probe sets, wherein the probe
25 pairs are in a medium at a concentration of about 1 to about 100 nM. In some instances, the concentration of the probe pairs is about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, or 500 nM. In some instances, the concentration of the probe pairs is 5 nM. In some instances, the probe sets are diluted in a hybridization (Hyb) buffer. In some instances, the probe sets are at a concentration of 5 nM in Hyb buffer.

30 In some instances, probe hybridization occurs at about 50°C. In some instances, the temperature of probe hybridization ranges from about 30°C to about 75°C, from about 35°C to about 70°C, or from about 40°C to about 65°C. In some embodiments, the temperature is about 30°C, about 31°C, about 32°C, about 33°C, about 34°C, about 35°C, about 36°C, about 37°C, about 38°C, about 39°C, about 40°C, about 41°C, about 42°C, about 43°C, about 44°C,

about 45°C, about 46°C, about 47°C, about 48°C, about 49°C, about 50°C, about 51°C, about 52°C, about 53°C, about 54°C, about 55°C, about 56°C, about 57°C, about 58°C, about 59°C, about 60°C, about 61°C, about 62°C, about 63°C, about 64°C, about 65°C, about 66°C, about 67°C, about 68°C, about 69°C, or about 70°C. In some instances, probe hybridization occurs for about 30 minutes, about 1 hour, about 2 hours, about 2.5 hours, about 3 hours, or more. In some instances, probe hybridization occurs for about 2.5 hours at 50°C.

In some instances, the hybridization buffer includes SSC (e.g., 1x SSC) or SSPE. In some instances, the hybridization buffer includes formamide or ethylene carbonate. In some instances, the hybridization buffer includes one or more salts, like Mg salt for example MgCl₂, Na salt for example NaCl, Mn salt for example MnCl₂. In some instances, the hybridization buffer includes Denhardt's solution, dextran sulfate, ficoll, PEG or other hybridization rate accelerators. In some instances, the hybridization buffer includes a carrier such as yeast tRNA, salmon sperm DNA, and/or lambda phage DNA. In some instances, the hybridization buffer includes one or more blockers. In some instances, the hybridization buffer includes RNase inhibitor(s). In some instances, the hybridization buffer can include BSA, sequence specific blockers, non-specific blockers, EDTA, RNase inhibitor(s), betaine, TMAC, or DMSO. In some instances, a hybridization buffer can further include detergents such as Tween, Triton-X 100, sarkosyl, and SDS. In some instances, the hybridization buffer includes nuclease-free water, DEPC water.

In some embodiments, the complementary sequences to which the first probe oligonucleotide and the second probe oligonucleotide bind are 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, about 15, about 20, about 25, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 95, about 100, about 125, about 150, about 175, about 200, about 250, about 300, about 350, about 400, about 450, about 500, about 600, about 700, about 800, about 900, or about 1000 nucleotides away from each other. Gaps between the probe oligonucleotides may first be filled prior to ligation, using, for example, Mu polymerase, DNA polymerase, RNA polymerase, reverse transcriptase, VENT polymerase, Taq polymerase, and/or any combinations, derivatives, and variants (e.g., engineered mutants) thereof. In some embodiments, when the first and second probe oligonucleotides are separated from each other by one or more nucleotides, nucleotides are ligated between the first and second probe oligonucleotides. In some embodiments, when the first and second probe oligonucleotides are separated from each other by one or more nucleotides, deoxyribonucleotides are ligated between the first and second probe oligonucleotides.

In some instances, after hybridization, the biological sample is washed with a post-hybridization wash buffer. In some instances, the post-hybridization wash buffer includes one or more of SSC, yeast tRNA, formamide, ethylene carbonate, and nuclease-free water.

Additional embodiments regarding probe hybridization are further provided.

5

(i) Hybridizing Temperatures

In some embodiments, the method described utilizes oligonucleotides that include deoxyribonucleic acids (instead of strictly utilizing ribonucleotides) at the site of ligation. Utilizing deoxyribonucleic acids in the methods described herein create a more uniform efficiency that can be readily-controlled and flexible for various applications.

In a non-limiting example, the methods disclosed herein include contacting a biological sample with a plurality of oligonucleotides (e.g., probes) including a first oligonucleotide (e.g., a first probe) and a second oligonucleotide (e.g., a second probe), wherein the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) are complementary to a first sequence present in an analyte and a second sequence present in the analyte, respectively; hybridizing the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) to the analyte at a first temperature; hybridizing the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) to a third oligonucleotide (e.g., a splint oligonucleotide) at a second temperature such that the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) abut each other; ligating the first oligonucleotide (e.g., the first probe) to the second oligonucleotide (e.g., the second probe) to create a ligation product; contacting the biological sample with a substrate, wherein a capture probe is immobilized on the substrate, wherein the capture probe includes a spatial barcode and a capture domain; allowing the ligation product to specifically bind to the capture domain; and determining (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample; wherein the first oligonucleotide (e.g., the first probe), the second oligonucleotide (e.g., the second probe), and the third oligonucleotide are DNA oligonucleotides, and wherein the first temperature is a higher temperature than the second temperature.

A non-limiting example of this method is shown in **FIG. 11**. A biological sample including an analyte **1101** is contacted with a first probe **1102** and a second probe **1103**. The

first probe **1102** and the second probe **1103** hybridize to the analyte at a first target sequence **1104** and a second target sequence **1105**, respectively. The first probe and the second probe include free ends **1107-1110**. As shown in **FIG. 11**, the first and second target sequences are not directly adjacent in the analyte. After hybridization, unbound first and second probes are washed away. Then, a third oligonucleotide **1106** hybridizes to the first and the second probes at **1108** and **1109**, respectively. After hybridization, the first probe is extended **1112** and a ligation product is created that includes the first probe sequence and the second probe sequence. Alternatively, instead of extending the first probe, the third oligonucleotide is used to “bind” the first probe and the second probe together. In such cases, the first probe and the second probe bound together by the third oligonucleotide can be referred to as a ligation product. The ligation product then is contacted with a substrate **1111**, and the ligation product is bound to a capture probe **1113** of the substrate **1111** on the array at distinct spatial positions. In some embodiments, the biological sample is contacted with the substrate **1111** prior to being contacted with the first probe and the second probe.

In some embodiments, the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) hybridize to an analyte at a first temperature. In some embodiments, the first temperature ranges from about 50°C to about 75°C, from about 55°C to about 70°C, or from about 60°C to about 65°C. In some embodiments, the first temperature is about 55°C, about 56°C, about 57°C, about 58°C, about 59°C, about 60°C, about 61°C, about 62°C, about 63°C, about 64°C, about 65°C, about 66°C, about 67°C, about 68°C, about 69°C, or about 70°C.

In some embodiments, after the step of hybridizing the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) to the analyte, a wash step is performed to remove unbound oligonucleotides (e.g., probes). The wash step can be performed using any of the wash methods and solutions described herein.

In some embodiments, after the step of hybridizing the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) to the analyte, a third oligonucleotide (e.g., a splint oligonucleotide) is added to the analyte. In some embodiments, the third oligonucleotide is an oligonucleotide. In some embodiments, the third oligonucleotide is a DNA oligonucleotide.

In some embodiments, the third oligonucleotide includes a sequence that is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to a portion of the first probe oligonucleotide (e.g., a portion of the first

probe that is not hybridized to the analyte (e.g., an auxiliary sequence)). In some embodiments, the third oligonucleotide includes a sequence that is 100% complementary to a portion of the first oligonucleotide (e.g., the first probe). In some embodiments, the third oligonucleotide includes a sequence that is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary to a portion of the second probe oligonucleotide (e.g., a portion of the second probe that is not hybridized to the analyte (e.g., an auxiliary sequence)). In some embodiments, the third oligonucleotide includes a sequence that is 100% complementary to a portion of the second oligonucleotide (e.g., the second probe). In some embodiments, the third oligonucleotide hybridizes to the first oligonucleotide (e.g., the first probe) at the complementary portion. In some embodiments, the third oligonucleotide hybridizes to the second oligonucleotide (e.g., the second probe) at the complementary portion.

In some embodiments, the third oligonucleotide hybridizes to the first oligonucleotide (e.g., the first probe) and to the second oligonucleotide (e.g., the second probe) at a second temperature. In some embodiments, the second temperature is lower than the first temperature at which the first and second oligonucleotides (e.g., the first and second probes) bind the analyte. In some embodiments, the second temperature ranges from about 15°C to about 35°C, from about 20°C to about 30°C, or from about 25°C to about 30°C. In some embodiments, the first temperature is about 15°C, about 16°C, about 17°C, about 18°C, about 19°C, about 20°C, about 21°C, about 22°C, about 23°C, about 24°C, about 25°C, about 26°C, about 27°C, about 28°C, about 29°C, about 30°C, about 31°C, about 32°C, about 33°C, about 34°C, or about 35°C. Methods including a third, or splint, oligonucleotide have been described in U.S. Patent Pub. No. 2019/0055594A1, which is herein incorporated by reference in its entirety.

In some embodiments, after the step of hybridizing the third oligonucleotide to the analyte, a wash step is performed to remove unbound third oligonucleotides. The wash step can be performed using any of the wash methods and solutions described herein. In some embodiments, after the washing step, the first and second oligonucleotides (e.g., the first and second probes) are bound to (e.g., hybridized to) the analyte, and the third oligonucleotide is bound to (e.g., hybridized to) the first and second oligonucleotides (e.g., at portions of the first and second probes that are not bound to the analyte).

In some embodiments, the first oligonucleotide (e.g., the first probe), the second oligonucleotide (e.g., the second probe), and the third oligonucleotide are added to the

biological sample at the same time. Then, in some embodiments, the temperature is adjusted to the first temperature to allow the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe) to hybridize to the analyte in the biological sample. Next, the temperature is adjusted to the second temperature to allow the third oligonucleotide to hybridize to the first oligonucleotide and the second oligonucleotide.

In some embodiments where a third oligonucleotide hybridizes to a first probe and a second probe that are hybridized to targets sequences that are not directly adjacent in the analyte, the third oligonucleotide is extended to fill the gap between the first probe and the second probe. In some instances, a polymerase (e.g., a DNA polymerase) can extend one of the probes (e.g., the first probe) prior to ligation. For example, as shown in **FIG. 11**, the first probe **1102** is extended **1112** to fill the gap **1114** between the first probe **1102** and the second probe **1103**.

In some embodiments, a ligation step is performed. Ligation can be performed using any of the methods described herein. In some embodiments, the step includes ligation of the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe), forming a ligation product. In some embodiments, the third oligonucleotide serves as an oligonucleotide splint to facilitate ligation of the first oligonucleotide (e.g., the first probe) and the second oligonucleotide (e.g., the second probe). In some embodiments, ligation is chemical ligation. In some embodiments, ligation is enzymatic ligation. In some embodiments, the ligase is a T4 RNA ligase (Rnl2), a splintR ligase, a single stranded DNA ligase, or a T4 DNA ligase.

(ii) Hybridization Buffer

In some embodiments, a first probe and a second probe are hybridized to the analyte in a hybridization buffer. In some instances, the hybridization buffer contains formamide. In other instances the hybridization buffer is formamide free. Formamide is not human friendly and it is a known health hazard. Chemically, it can oxidize over time, thereby impacting reagent shelf life and, most importantly, reagent efficacy. As such, the methods described herein can include formamide-free buffers, including formamide-free hybridization buffer.

In some embodiments, the formamide-free hybridization buffer is a saline-sodium citrate (SSC) hybridization buffer. In some embodiment, the SSC is present in the SSC hybridization buffer from about 1x SSC to about 6x SSC (e.g., about 1x SSC to about 5x SSC, about 1x SSC to about 4x SSC, about 1x SSC to about 3x SSC, about 1x SSC to about 2x SSC, about 2x SSC to about 6x SSC, about 2x SSC to about 5x SSC, about 2x SSC to

about 4x SSC, about 2x SSC to about 3x SSC, about 3x SSC to about 5x SSC, about 3x SSC to about 4x SSC, about 4x SSC to about 6x SSC, about 4x SSC to about 6x SSC, about 4x SSC to about 5x SSC, or about 5x SSC to about 6x SSC). In some embodiments, the SSC is present in the SSC hybridization buffer from about 2x SSC to about 4x SSC. In some
 5 embodiments, SSPE hybridization buffer can be used.

In some embodiments, the SSC hybridization buffer comprises a solvent. In some embodiments, the solvent comprises ethylene carbonate instead of formamide (2020, Kalinka et al., *Scientia Agricola* 78(4):e20190315). In some embodiments, ethylene carbonate is present in the SSC hybridization buffer from about 10% (w/v) to about 25% (w/v) (e.g., about
 10 10% (w/v) to about 20% (w/v), about 10% (w/v) to about 15% (w/v), about 15% (w/v) to about 25% (w/v), about 15% (w/v) to about 20% (w/v), or about 20% (w/v) to about 25% (w/v)). In some embodiments, ethylene carbonate is present in the SSC hybridization buffer from about 15% (w/v) to about 20% (w/v). In some embodiments, ethylene carbonate is present in the SSC hybridization buffer at about 10% (w/v), about 11% (w/v), about 12%
 15 (w/v), about 13% (w/v), about 14% (w/v), about 15% (w/v), about 16% (w/v), about 17% (w/v), about 18% (w/v), about 19% (w/v), about 20% (w/v), about 21% (w/v), about 22% (w/v), about 23% (w/v), about 24% (w/v), or about 25% (w/v). In some embodiments, ethylene carbonate is present in the SSC hybridization buffer at about 13% (w/v).

In some embodiments, the SSC hybridization buffer is at a temperature from about
 20 40°C to about 60°C (e.g., about 40°C to about 55°C, about 40°C to about 50°C, about 40°C to about 45°C, about 45°C to about 60°C, about 45°C to about 55°C, about 45°C to about 50°C, about 50°C to about 60°C, about 50°C to about 55°C, or about 55°C to about 60°C). In some embodiments, the SSC hybridization buffer is at temperature from about 45°C to about 55°C, or any of the subranges described herein. In some embodiments, the SSC hybridization
 25 buffer is at a temperature of about 40°C, about 41°C, about 42°C, about 43°C, about 44°C, about 45°C, about 46°C, about 47°C, about 48°C, about 49°C, about 50°C, about 51°C, about 52°C, about 53°C, about 54°C, about 55°C, about 56°C, about 57°C, about 58°C, about 59°C, or about 60°C. In some embodiments, the SSC hybridization buffer is at a temperature of about 50°C.

30 In some embodiments, the SSC hybridization buffer further comprises one or more of a carrier, a crowder, or an additive. Non-limiting examples of a carrier that can be included in the hybridization buffer include: yeast tRNA, salmon sperm DNA, lambda phage DNA, glycogen, and cholesterol. Non-limiting examples of a molecular crowder that can be included in the hybridization buffer include: Ficoll, dextran, Denhardt's solution, and PEG.

Non-limiting examples of additives that can be included in the hybridization buffer include: binding blockers, RNase inhibitors, T_m adjustors and adjuvants for relaxing secondary nucleic acid structures (e.g., betaine, TMAC, and DMSO). Further, a hybridization buffer can include detergents such as SDS, Tween, Triton-X 100, and sarkosyl (e.g., N-Lauroylsarcosine sodium salt). A skilled artisan would understand that a buffer for hybridization of nucleic acids could include many different compounds that could enhance the hybridization reaction.

(f) Washing

In some embodiments, the methods disclosed herein also include a wash step. The wash step removes any unbound probes. Wash steps could be performed between any of the steps in the methods disclosed herein. For example, a wash step can be performed after adding probes to the biological sample. As such, free/unbound probes are washed away, leaving only probes that have hybridized to an analyte. In some instances, multiple (i.e., at least 2, 3, 4, 5, or more) wash steps occur between the methods disclosed herein. Wash steps can be performed at times (e.g., 1, 2, 3, 4, or 5 minutes) and temperatures (e.g., room temperature; 4°C known in the art and determined by a person of skill in the art).

In some instances, wash steps are performed using a wash buffer. In some instances, the wash buffer includes SSC (e.g., 1x SSC). In some instances, the wash buffer includes PBS (e.g., 1x PBS). In some instances, the wash buffer includes PBST (e.g., 1x PBST). In some instances, the wash buffer can also include formamide or be formamide free.

Additional embodiments regarding wash steps are provided herein.

(i) Formamide Free Wash Buffer

In some embodiments, after ligating a first probe and a second probe, the one or more unhybridized first probes, one or more unhybridized second probes, or both, are removed from the array. In some embodiments, after ligating a first probe, one or more spanning probes, and a second probe, the one or more unhybridized first, second, and/or spanning probes, are removed from the array. In some embodiments, after ligating a first probe, a second probe, and a third oligonucleotide, the one or more unhybridized first probes, one or more unhybridized second probes, or one or more third oligonucleotides, or all the above, are removed from the array.

In some embodiments, a pre-hybridization buffer is used to wash the sample. In some embodiments, a phosphate buffer is used. In some embodiments, multiple wash steps are performed to remove unbound oligonucleotides.

In some embodiments, removing includes washing the one or more unhybridized probes (e.g., a first probe, a second probe, a spanning probe, additional spanning probes, and a third oligonucleotide) from the array in a formamide-free wash buffer.

In some embodiments, the formamide-free wash buffer is an SSC wash buffer. In some embodiments, SSC is present in the SSC wash buffer from about 0.01x SSC to about 1x SSC (e.g., about 0.01x SSC to about 0.5x SSC, 0.01x SSC to about 0.1x SSC, about 0.01x SSC to about 0.05x SSC, about 0.05x SSC to about 1x SSC, about 0.05x SSC to about 0.5x SSC, about 0.05x SSC to about 0.1x SSC, about 0.1x SSC to about 1x SSC, about 0.1x SSC to about 0.5x SSC, or about 0.5x SSC to about 1x SSC). In some embodiments, SSC is present in the SSC wash buffer at about 0.01x SSC, about 0.02x SSC, about 0.03x SSC, about 0.04x SSC, about 0.05x SSC, about 0.06x SSC, about 0.07x SSC, about 0.08x SSC, about 0.09x SSC, about 0.1x SSC, about 0.2x SSC, about 0.3x SSC, about 0.4x SSC, about 0.5x SSC, about 0.6x SSC, about 0.7x SSC, about 0.8x SSC, about 0.9x SSC, or about 1x SSC. In some embodiments, SSC is present in the SSC wash buffer at about 0.1x SSC.

In some embodiments, the SSC wash buffer comprises a detergent. In some embodiments, the detergent comprises sodium dodecyl sulfate (SDS). In some embodiments, SDS is present in the SSC wash buffer from about 0.01% (v/v) to about 0.5% (v/v) (e.g., about 0.01% (v/v) to about 0.4% (v/v), about 0.01% (v/v) to about 0.3% (v/v), about 0.01% (v/v) to about 0.2% (v/v), about 0.01% (v/v) to about 0.1% (v/v), about 0.05% (v/v) to about 0.5% (v/v), about 0.05% (v/v) to about 0.4% (v/v), about 0.05% (v/v) to about 0.3% (v/v), about 0.05% (v/v) to about 0.2% (v/v), about 0.05% (v/v) to about 0.1% (v/v), about 0.1% (v/v) to about 0.5% (v/v), about 0.1% (v/v) to about 0.4% (v/v), about 0.1% (v/v) to about 0.3% (v/v), about 0.1% (v/v) to about 0.2% (v/v), about 0.2% (v/v) to about 0.5% (v/v), about 0.2% (v/v) to about 0.4% (v/v), about 0.2% (v/v) to about 0.3% (v/v), about 0.3% (v/v) to about 0.5% (v/v), about 0.3% (v/v) to about 0.4% (v/v), or about 0.4% (v/v) to about 0.5% (v/v)). In some embodiments, the SDS is present the SSC wash buffer at about 0.01% (v/v), about 0.02% (v/v), about 0.03% (v/v), about 0.04% (v/v), about 0.05% (v/v), about 0.06% (v/v), about 0.07% (v/v), about 0.08% (v/v), about 0.09% (v/v), about 0.10% (v/v), about 0.2% (v/v), about 0.3% (v/v), about 0.4% (v/v), or about 0.5% (v/v). In some embodiments, the SDS is present in the SSC wash buffer at about 0.1% (v/v). In some embodiments, sarkosyl may be present in the SSC wash buffer.

In some embodiments, the SSC wash buffer comprises a solvent. In some embodiments, the solvent comprises ethylene carbonate. In some embodiments, ethylene carbonate is present in the SSC wash buffer from about 10% (w/v) to about 25% (w/v), or

any of the subranges described herein. In some embodiments, ethylene carbonate is present in the SSC wash buffer from about 15% (w/v) to about 20% (w/v). In some embodiments, ethylene carbonate is present in the SSC wash buffer at about 16% (w/v).

In some embodiments, the SSC wash buffer is at a temperature from about 50°C to about 70°C (e.g., about 50°C to about 65°C, about 50°C to about 60°C, about 50°C to about 55°C, about 55°C to about 70°C, about 55°C to about 65°C, about 55°C to about 60°C, about 60°C to about 70°C, about 60°C to about 65°C, or about 65°C to about 70°C). In some embodiments, the SSC wash buffer is at a temperature from about 55°C to about 65°C. In some embodiments, the SSC wash buffer is at a temperature about 50°C, about 51°C, about 52°C, about 53°C, about 54°C, about 55°C, about 56°C, about 57°C, about 58°C, about 59°C, about 60°C, about 61°C, about 62°C, about 63°C, about 64°C, about 65°C, about 66°C, about 67°C, about 68°C, about 69°C, or about 70°C. In some embodiments, the SSC wash buffer is at a temperature of about 60°C.

In some embodiments, the method includes releasing the ligation product, where releasing is performed after the array is washed to remove the one or more unhybridized first and second probes.

(g) Ligation

In some embodiments, after hybridization of the probe oligonucleotides (e.g., a first probe, a second probe, a spanning probe, additional spanning probes, and/or a third oligonucleotide) to the analyte, the probe (e.g., a first probe, a second probe, a spanning probe, additional spanning probes, and/or a third oligonucleotide) can be ligated together, creating a single ligation product that includes one or more sequences that are complementary to the analyte. Ligation can be performed enzymatically or chemically, as described herein.

In some instances, the ligation is an enzymatic ligation reaction, using a ligase (e.g., T4 RNA ligase (Rnl2), a SplintR ligase, a single stranded DNA ligase, or a T4 DNA ligase). See, e.g., Zhang *et al.*; *RNA Biol.* 2017; 14(1): 36–44, which is incorporated by reference in its entirety, for a description of KOD ligase. Following the enzymatic ligation reaction, the probes (e.g., a first probe, a second probe, a spanning probe, additional spanning probes, and/or a third oligonucleotide) may be considered ligated.

In some embodiments, a polymerase catalyzes synthesis of a complementary strand of the ligation product, creating a double-stranded ligation product. In some instances, the polymerase is DNA polymerase. In some embodiments, the polymerase has 5' to 3' polymerase activity. In some embodiments, the polymerase has 3' to 5' exonuclease activity

for proofreading. In some embodiments, the polymerase has 5' to 3' polymerase activity and 3' to 5' exonuclease activity for proofreading.

In some embodiments, the probe (e.g., a first probe, a second probe, a spanning probe, additional spanning probes, and/or a third oligonucleotide) may each comprise a reactive moiety such that, upon hybridization to the target and exposure to appropriate ligation conditions, the probe oligonucleotides may ligate to one another. In some embodiments, probe oligonucleotides that include a reactive moiety are ligated chemically. For example, a first probe capable of hybridizing to a first target region (e.g., a first target sequence or a first portion) of a nucleic acid molecule may comprise a first reactive moiety, and a second probe oligonucleotide capable of hybridizing to a second target region (e.g., a second target sequence or a second portion) of the nucleic acid molecule may comprise a second reactive moiety. When the first and second probes are hybridized to the first and second target regions (e.g., first and second target sequences) of the nucleic acid molecule, the first and second reactive moieties may be adjacent to one another. A reactive moiety of a probe may be selected from the non-limiting group consisting of azides, alkynes, nitrones (e.g., 1,3-nitrones), strained alkenes (e.g., trans-cycloalkenes such as cyclooctenes or oxanorbornadiene), tetrazines, tetrazoles, iodides, thioates (e.g., phosphorothioate), acids, amines, and phosphates. For example, the first reactive moiety of a first probe may comprise an azide moiety, and a second reactive moiety of a second probe may comprise an alkyne moiety. The first and second reactive moieties may react to form a linking moiety. A reaction between the first and second reactive moieties may be, for example, a cycloaddition reaction such as a strain-promoted azide-alkyne cycloaddition, a copper-catalyzed azide-alkyne cycloaddition, a strain-promoted alkyne-nitrone cycloaddition, a Diels-Alder reaction, a [3+2] cycloaddition, a [4+2] cycloaddition, or a [4+1] cycloaddition; a thiol-ene reaction; a nucleophilic substitution reaction; or another reaction. In some cases, reaction between the first and second reactive moieties may yield a triazole moiety or an isoxazoline moiety. A reaction between the first and second reactive moieties may involve subjecting the reactive moieties to suitable conditions such as a suitable temperature, pH, or pressure and providing one or more reagents or catalysts for the reaction. For example, a reaction between the first and second reactive moieties may be catalyzed by a copper catalyst, a ruthenium catalyst, or a strained species such as a difluorooctyne, dibenzylcyclooctyne, or biarylazacyclooctynone. Reaction between a first reactive moiety of a first probe hybridized to a first target region (e.g., a first target sequence or first portion) of the nucleic acid molecule and a second reactive moiety of a third probe oligonucleotide hybridized to a second target region (e.g., a

first target sequence or a first portion) of the nucleic acid molecule may link the first probe and the second probe to provide a ligated probe. Upon linking, the first and second probe may be considered ligated. Accordingly, reaction of the first and second reactive moieties may comprise a chemical ligation reaction such as a copper-catalyzed 5' azide to 3' alkyne "click" chemistry reaction to form a triazole linkage between two probe oligonucleotides. In other non-limiting examples, an iodide moiety may be chemically ligated to a phosphorothioate moiety to form a phosphorothioate bond, an acid may be ligated to an amine to form an amide bond, and/or a phosphate and amine may be ligated to form a phosphoramidate bond.

FIGs. 12A-12E illustrates examples of representative reactions. **FIG. 12A** shows a chemical ligation reaction of an alkyne moiety **1202** and an azide moiety **1204** reacting under copper-mediated cycloaddition to form a triazole linkage **1206**. **FIG. 12B** shows a chemical ligation reaction of a phosphorothioate group **1208** with an iodide group **1210** to form a phosphorothioate linkage **1212**. **FIG. 12C** shows a chemical ligation reaction of an acid **1214** and amine **1216** to form an amide linkage **1218**. **FIG. 12D** shows a chemical ligation reaction of a phosphate moiety **1220** and an amine moiety **1222** to form a phosphoramidate linkage **1224**. **FIG. 12E** shows a conjugation reaction of two species **1226** and **1228**.

In some instances, ligation is performed in a ligation buffer. In instances where probe ligation is performed on diribo-containing probes, the ligation buffer can include T4 RNA Ligase Buffer 2, enzyme (e.g., RNL2 ligase), and nuclease free water. In instances where probe ligation is performed on DNA probes, the ligation buffer can include Tris-HCl pH7.5, MnCl₂, ATP, DTT, surrogate fluid (e.g., glycerol), enzyme (e.g., SplintR ligase), and nuclease-free water.

In some embodiments, the ligation buffer includes additional reagents. In some instances, the ligation buffer includes adenosine triphosphate (ATP) is added during the ligation reaction. DNA ligase-catalyzed sealing of nicked DNA substrates is first activated through ATP hydrolysis, resulting in covalent addition of an AMP group to the enzyme. After binding to a nicked site in a DNA duplex, the ligase transfers the AMP to the phosphorylated 5'-end at the nick, forming a 5'-5' pyrophosphate bond. Finally, the ligase catalyzes an attack on this pyrophosphate bond by the OH group at the 3'-end of the nick, thereby sealing it, whereafter ligase and AMP are released. If the ligase detaches from the substrate before the 3' attack, e.g. because of premature AMP reloading of the enzyme, then the 5' AMP is left at the 5'-end, blocking further ligation attempts. In some instances, ATP is added at a concentration of about 1 μM, about 10 μM, about 100 μM, about 1000 μM, or about 10000 μM during the ligation reaction.

In some embodiments, cofactors that aid in joining of the probe oligonucleotides are added during the ligation process. In some instances, the cofactors include magnesium ions (Mg^{2+}). In some instances, the cofactors include manganese ions (Mn^{2+}). In some instances, Mg^{2+} is added in the form of $MgCl_2$. In some instances, Mn^{2+} is added in the form of $MnCl_2$.

- 5 In some instances, the concentration of $MgCl_2$ is at about 1 mM, at about 10 mM, at about 100 mM, or at about 1000 mM. In some instances, the concentration of $MnCl_2$ is at about 1 mM, at about 10 mM, at about 100 mM, or at about 1000 mM.

- 10 In some embodiments, the ligation product includes a capture probe capture domain, which can bind to a capture probe (e.g., a capture probe immobilized, directly or indirectly, on a substrate). In some embodiments, methods provided herein include contacting a biological sample with a substrate, wherein the capture probe is affixed to the substrate (e.g., immobilized to the substrate, directly or indirectly). In some embodiments, the capture probe capture domain of the ligated probe specifically binds to the capture domain.

- 15 After ligation, in some instances, the biological sample is washed with a post-ligation wash buffer. In some instances, the post-ligation wash buffer includes one or more of SSC (e.g., 1x SSC), ethylene carbonate or formamide, and nuclease free water. In some instances, the biological sample is washed at this stage at about 50°C to about 70°C. In some instances, the biological sample is washed at about 60°C.

20 (i) Ligation Including Pre-Adenylated 5' Phosphate on Second Probe

- Provided herein are methods for determining a location of a target nucleic acid in a biological sample that include: (a) contacting the biological sample with a substrate comprising a plurality of capture probes, where a capture probe of the plurality of capture probes comprises a capture domain and a spatial barcode; (b) hybridizing a target nucleic acid
- 25 in the biological sample with a first probe and a second probe, where the first probe comprises, from 3' to 5', a sequence substantially complementary to the capture domain and a sequence that is substantially complementary to a first sequence in the target nucleic acid and has a pre-adenylated phosphate group at its 5' end; the second probe comprises a sequence substantially complementary to a second sequence in the target nucleic acid; (c)
- 30 generating a ligation product by ligating a 3' end of the second probe to the 5' end of the first probe using a ligase that does not require adenosine triphosphate for ligase activity; (d) releasing the ligation product from the target nucleic acid and binding the capture domain of the ligation product specifically to the capture domain of capture probe; and (e) determining
- (i) all or a part of a sequence corresponding to the ligation product, or a complement thereof,

and (ii) all or a part of a sequence corresponding to the spatial barcode, or a complement thereof, and using the determined sequences of (i) and (ii) to identify the location of the target nucleic acid in the biological sample

In some instances, the ligase that does not require adenosine triphosphate for ligase activity (e.g., thermostable 5' AppDNA/RNA Ligase, truncated T4 RNA Ligase 2 (trRnl2), truncated T4 RNA Ligase 2 K227Q, truncated T4 RNA Ligase 2 KQ, Chlorella Virus PBCV-1 DNA Ligase, and combinations thereof). See, e.g., Nichols et al., "RNA Ligases," *Curr. Protocol. Molec. Biol.* 84(1):3.15.1–.4 (2008); Viollet et al., "T4 RNA Ligase 2 Truncated Active Site Mutants: Improved Tools for RNA Analysis," *BMC Biotechnol.* 11: 72 (2011); and Ho et al., "Bacteriophage T4 RNA Ligase 2 (gp24.1) Exemplifies a Family of RNA Ligases Found in All Phylogenetic Domains," *PNAS* 99(20):12709–14 (2002), which are hereby incorporated by reference in their entirety for a description of T4 RNA Ligases and truncated T4 RNA Ligases. Thermostable 5' AppDNA/RNA Ligase is an enzyme belonging to the Ligase family that catalyzes the ligation of the 3' end of ssRNA or ssDNA to a 5'-adenylated ssDNA or 5'-adenylated ssRNA. Truncated T4 RNA Ligase 2 is an enzyme belonging to the Ligase family that catalyzes the ligation of dsRNA nicks and ssRNA to ssRNA. It can also ligate the 3' end of RNA or DNA to a 5'-pDNA when annealed to an RNA complement, and the 3' end of RNA to a 5'-pRNA when annealed to a DNA complement, with reduced efficiency. Truncated T4 RNA Ligase 2 K227Q is an enzyme belonging to the Ligase family that catalyzes the ligation of the 3' end of ssRNA to 5' adenylated ssDNA and 5' adenylated ssRNA. It has a reduction of side products as compared to truncated T4 RNA Ligase 2. Truncated T4 RNA Ligase 2 KQ is an enzyme belonging to the Ligase family that catalyzes the ligation of the 3' end of ssRNA to 5' adenylated ssDNA and 5' adenylated ssRNA. It is a preferred choice for ligation of ssRNA to preadenylated adapters and has a reduction of side products as compared to truncated T4 RNA Ligase 2.

In some embodiments, the T4 RNA Ligase comprises a K227Q mutation. See Viollet et al., "T4 RNA Ligase 2 Truncated Active Site Mutants: Improved Tools for RNA Analysis," *BMC Biotechnol.* 11, which is hereby incorporated by reference in its entirety.

In some instances, cofactors that aid in ligation of the first and second probe are added during ligation. In some instances, the cofactors include magnesium ions (Mg^{2+}). In some instances, the cofactors include manganese ions (Mn^{2+}). In some instances, Mg^{2+} is added in the form of $MgCl_2$. In some instances, Mn^{2+} is added in the form of $MnCl_2$. In some instances, the concentration of $MgCl_2$ is at about 1 mM to about 10 mM. In some instances, the concentration of $MnCl_2$ is at about 1 mM to about 10 mM.

In some instances, the ligation occurs at a pH in the range of about 6.5 to about 9.0, about 6.5 to about 8.0, or about 7.5 to about 8.0.

In some embodiments, the ligation buffer includes an enzyme storage buffer. In some embodiments, the enzymes storage buffer includes glycerol. In some embodiments, the ligation buffer is supplemented with glycerol. In some embodiments, the glycerol is present in the ligation buffer at a total volume of 15% v/v.

(h) Permeabilization and Releasing the Ligation Product

In some embodiments, the methods provided herein include a permeabilizing step. In some embodiments, permeabilization occurs using a protease. In some embodiments, the protease is an endopeptidase. Endopeptidases that can be used include but are not limited to trypsin, chymotrypsin, elastase, thermolysin, pepsin, clostripain, glutamyl endopeptidase (GluC), ArgC, peptidyl-asp endopeptidase (ApsN), endopeptidase LysC and endopeptidase LysN. In some embodiments, the endopeptidase is pepsin. In some embodiments, after creating a ligation product (e.g., by ligating a first probe and a second probe that are hybridized to adjacent sequences in the analyte), the biological sample is permeabilized. In some embodiments, the biological sample is permeabilized contemporaneously with or prior to contacting the biological sample with a first probe and a second probe, hybridizing the first probe and the second probe to the analyte, generating a ligation product by ligating the first probe and the second probe, and releasing the ligated product from the analyte.

In some embodiments, methods provided herein include permeabilization of the biological sample such that the capture probe can more easily bind to the captured ligated probe (i.e., compared to no permeabilization). In some embodiments, reverse transcription (RT) reagents can be added to permeabilized biological samples. Incubation with the RT reagents can produce spatially-barcoded full-length cDNA from the captured analytes (e.g., polyadenylated mRNA). Second strand reagents (e.g., second strand primers, enzymes) can be added to the biological sample on the slide to initiate second strand synthesis.

In some instances, the permeabilization step includes application of a permeabilization buffer to the biological sample. In some instances, the permeabilization buffer includes a buffer (e.g., Tris pH 7.5), MgCl₂, sarkosyl detergent (e.g., sodium lauroyl sarcosinate), enzyme (e.g., proteinase K, and nuclease free water. In some instances, the permeabilization step is performed at 37°C. In some instances, the permeabilization step is performed for about 20 minutes to 2 hours (e.g., about 20 minutes, about 30 minutes, about

40 minutes, about 50 minutes, about 1 hour, about 1.5 hours, or about 2 hours). In some instances, the releasing step is performed for about 40 minutes.

In some embodiments, after generating a ligation product, the ligation product is released from the analyte. In some embodiments, a ligation product is released from the analyte using an endoribonuclease. In some embodiments, the endoribonuclease is RNase H, RNase A, RNase C, or RNase I. In some embodiments, the endoribonuclease is RNase H. RNase H is an endoribonuclease that specifically hydrolyzes the phosphodiester bonds of RNA, when hybridized to DNA. RNase H is part of a conserved family of ribonucleases which are present in many different organisms. There are two primary classes of RNase H: RNase H1 and RNase H2. Retroviral RNase H enzymes are similar to the prokaryotic RNase H1. All of these enzymes share the characteristic that they are able to cleave the RNA component of an RNA:DNA heteroduplex. In some embodiments, the RNase H is RNase H1, RNase H2, or RNase H1, or RNase H2. In some embodiments, the RNase H includes but is not limited to RNase HII from *Pyrococcus furiosus*, RNase HII from *Pyrococcus horikoshi*, RNase HI from *Thermococcus litoralis*, RNase HI from *Thermus thermophilus*, RNase HI from *E. coli*, or RNase HII from *E. coli*.

In some instances, the releasing step is performed using a releasing buffer. In some instances, the release buffer includes one or more of a buffer (e.g., Tris pH 7.5), enzyme (e.g., RNase H) and nuclease-free water. In some instances, the releasing step is performed at 37°C. In some instances, the releasing step is performed for about 20 minutes to 2 hours (e.g., about 20 minutes, about 30 minutes, about 40 minutes, about 50 minutes, about 1 hour, about 1.5 hours, or about 2 hours). In some instances, the releasing step is performed for about 30 minutes.

In some instances, the releasing step occurs before the permeabilization step. In some instances, the releasing step occurs after the permeabilization step. In some instances, the releasing step occurs at the same time as the permeabilization step (e.g., in the same buffer).

(i) Blocking probes

In some embodiments, a capture probe capture domain is blocked prior to adding a second probe oligonucleotide to a biological sample. This prevents the capture probe capture domain from prematurely hybridizing to the capture domain.

In some embodiments, a blocking probe is used to block or modify the free 3' end of the capture probe capture domain. In some embodiments, a blocking probe can be hybridized to the capture probe capture domain of the second probe to mask the free 3' end of the

capture probe capture domain. In some embodiments, a blocking probe can be a hairpin probe or partially double stranded probe. In some embodiments, the free 3' end of the capture probe capture domain of the second probe can be blocked by chemical modification, e.g., addition of an azidomethyl group as a chemically reversible capping moiety such that the capture probes do not include a free 3' end. Blocking or modifying the capture probe capture domain, particularly at the free 3' end of the capture probe capture domain, prior to contacting second probe with the substrate, prevents hybridization of the second probe to the capture domain (e.g., prevents the capture of a poly(A) of a capture probe capture domain to a poly(T) capture domain). In some embodiments, a blocking probe can be referred to as a capture probe capture domain blocking moiety.

In some embodiments, the blocking probes can be reversibly removed. For example, blocking probes can be applied to block the free 3' end of either or both the capture probe capture domain and/or the capture probes. Blocking interaction between the capture probe capture domain and the capture probe on the substrate can reduce non-specific capture to the capture probes. After the second probe hybridizes to the analyte and is ligated to a first probe, one or more spanning probes, or a third oligonucleotide, the blocking probes can be removed from the 3' end of the capture probe capture domain and/or the capture probe, and the ligation product can migrate to and become bound by the capture probes on the substrate. In some embodiments, the removal includes denaturing the blocking probe from capture probe capture domain and/or capture probe. In some embodiments, the removal includes removing a chemically reversible capping moiety. In some embodiments, the removal includes digesting the blocking probe with an RNase (e.g., RNase H).

In some embodiments, the blocking probes are oligo (dT) blocking probes. In some embodiments, the oligo (dT) blocking probes can have a length of 15-30 nucleotides. In some embodiments, the oligo (dT) blocking probes can have a length of 10-50 nucleotides, e.g., 10-50, 10-45, 10-40, 10-35, 10-30, 10-25, 10-20, 10-15, 15-50, 15-45, 15-40, 15-35, 15-30, 15-25, 15-20, 20-50, 20-45, 20-40, 20-35, 20-30, 20-25, 25-50, 25-45, 25-40, 25-35, 25-30, 30-50, 30-45, 30-40, 30-35, 35-50, 35-45, 35-40, 40-50, 40-45, or 45-50 nucleotides. In some embodiments, the analyte capture agents can be blocked at different temperatures (e.g., 4°C and 37°C).

(j) Biological Samples

Methods disclosed herein can be performed on any type of sample. In some embodiments, the sample is a fresh tissue. In some embodiments, the sample is a frozen

sample. In some embodiments, the sample was previously frozen. In some embodiments, the sample is a formalin-fixed, paraffin embedded (FFPE) sample.

Subjects from which biological samples can be obtained can be healthy or asymptomatic individuals, individuals that have or are suspected of having a disease (e.g., cancer) or a pre-disposition to a disease, and/or individuals that are in need of therapy or suspected of needing therapy. In some instances, the biological sample can include one or more diseased cells. A diseased cell can have altered metabolic properties, gene expression, protein expression, and/or morphologic features. Examples of diseases include inflammatory disorders, metabolic disorders, nervous system disorders, and cancer. In some instances, the biological sample includes cancer or tumor cells. Cancer cells can be derived from solid tumors, hematological malignancies, cell lines, or obtained as circulating tumor cells. In some instances, the biological sample is a heterogenous sample. In some instances, the biological sample is a heterogenous sample that includes tumor or cancer cells and/or stromal cells,

In some instances, the cancer is breast cancer. In some instances, the breast cancer is triple positive breast cancer (TPBC). In some instances, the breast cancer is triple negative breast cancer (TNBC).

In some instances, the cancer is colorectal cancer. In some instances, the cancer is ovarian cancer. In certain embodiments, the cancer is squamous cell cancer, small-cell lung cancer, non-small cell lung cancer, gastrointestinal cancer, Hodgkin's or non-Hodgkin's lymphoma, pancreatic cancer, glioblastoma, glioma, cervical cancer, ovarian cancer, liver cancer, bladder cancer, breast cancer, colon cancer, colorectal cancer, endometrial carcinoma, myeloma, salivary gland carcinoma, kidney cancer, basal cell carcinoma, melanoma, prostate cancer, vulval cancer, thyroid cancer, testicular cancer, esophageal cancer, or a type of head or neck cancer. In certain embodiments, the cancer treated is desmoplastic melanoma, inflammatory breast cancer, thymoma, rectal cancer, anal cancer, or surgically treatable or non-surgically treatable brain stem glioma. In some embodiments, the subject is a human.

FFPE samples generally are heavily cross-linked and fragmented, and therefore this type of sample allows for limited RNA recovery using conventional detection techniques. In certain embodiments, methods of targeted RNA capture provided herein are less affected by RNA degradation associated with FFPE fixation than other methods (e.g., methods that take advantage of oligo-dT capture and reverse transcription of mRNA). In certain embodiments, methods provided herein enable sensitive measurement of specific genes of interest that otherwise might be missed with a whole transcriptomic approach.

In some instances, FFPE samples are stained (e.g., using H&E). The methods disclosed herein are compatible with H&E will allow for morphological context overlaid with transcriptomic analysis. However, depending on the need some samples may be stained with only a nuclear stain, such as staining a sample with only hematoxylin and not eosin, when
 5 location of a cell nucleus is needed.

In some embodiments, a biological sample (e.g. tissue section) can be fixed with methanol, stained with hematoxylin and eosin, and imaged. In some embodiments, fixing, staining, and imaging occurs before one or more probes are hybridized to the sample. Some embodiments of any of the workflows described herein can further include a destaining step
 10 (e.g., a hematoxylin and eosin destaining step), after imaging of the sample and prior to permeabilizing the sample. For example, destaining can be performed by performing one or more (e.g., one, two, three, four, or five) washing steps (e.g., one or more (e.g., one, two, three, four, or five) washing steps performed using a buffer including HCl). The images can be used to map spatial gene expression patterns back to the biological sample. A
 15 permeabilization enzyme can be used to permeabilize the biological sample directly on the slide.

In some embodiments, the FFPE sample is deparaffinized, permeabilized, equilibrated, and blocked before target probe oligonucleotides are added. In some embodiments, deparaffinization using xylenes. In some embodiments, deparaffinization
 20 includes multiple washes with xylenes. In some embodiments, deparaffinization includes multiple washes with xylenes followed by removal of xylenes using multiple rounds of graded alcohol followed by washing the sample with water. In some aspects, the water is deionized water. In some embodiments, equilibrating and blocking includes incubating the sample in a pre-Hyb buffer. In some embodiments, the pre-Hyb buffer includes yeast tRNA.
 25 In some embodiments, permeabilizing a sample includes washing the sample with a phosphate buffer. In some embodiments, the buffer is PBS. In some embodiments, the buffer is PBST.

(k) Determining the Sequence of the Ligation Product

30 After a ligation product from the sample has hybridized or otherwise been associated with a capture probe according to any of the methods described above in connection with the general spatial cell-based analytical methodology, the barcoded constructs that result from hybridization/association are analyzed.

In some embodiments, after contacting a biological sample with a substrate that includes capture probes, a removal step can optionally be performed to remove all or a portion of the biological sample from the substrate. In some embodiments, the removal step includes enzymatic and/or chemical degradation of cells of the biological sample. For example, the removal step can include treating the biological sample with an enzyme (e.g., a proteinase, e.g., proteinase K) to remove at least a portion of the biological sample from the substrate. In some embodiments, the removal step can include ablation of the tissue (e.g., laser ablation).

In some embodiments, provided herein are methods for spatially detecting an analyte (e.g., detecting the location of an analyte, e.g., a biological analyte) from a biological sample (e.g., present in a biological sample), the method comprising: (a) optionally staining and/or imaging a biological sample on a substrate; (b) permeabilizing (e.g., providing a solution comprising a permeabilization reagent to) the biological sample on the substrate; (c) contacting the biological sample with an array comprising a plurality of capture probes, wherein a capture probe of the plurality captures the biological analyte; and (d) analyzing the captured biological analyte, thereby spatially detecting the biological analyte; wherein the biological sample is fully or partially removed from the substrate.

In some embodiments, a biological sample is not removed from the substrate. For example, the biological sample is not removed from the substrate prior to releasing a capture probe (e.g., a capture probe bound to an analyte) from the substrate. In some embodiments, such releasing comprises cleavage of the capture probe from the substrate (e.g., via a cleavage domain). In some embodiments, such releasing does not comprise releasing the capture probe from the substrate (e.g., a copy of the capture probe bound to an analyte can be made and the copy can be released from the substrate, e.g., via denaturation). In some embodiments, the biological sample is not removed from the substrate prior to analysis of an analyte bound to a capture probe after it is released from the substrate. In some embodiments, the biological sample remains on the substrate during removal of a capture probe from the substrate and/or analysis of an analyte bound to the capture probe after it is released from the substrate. In some embodiments, the biological sample remains on the substrate during removal (e.g., via denaturation) of a copy of the capture probe (e.g., complement). In some embodiments, analysis of an analyte bound to capture probe from the substrate can be performed without subjecting the biological sample to enzymatic and/or chemical degradation of the cells (e.g., permeabilized cells) or ablation of the tissue (e.g., laser ablation).

In some embodiments, at least a portion of the biological sample is not removed from the substrate. For example, a portion of the biological sample can remain on the substrate prior to releasing a capture probe (e.g., a capture probe bound to an analyte) from the substrate and/or analyzing an analyte bound to a capture probe released from the substrate. In
5 some embodiments, at least a portion of the biological sample is not subjected to enzymatic and/or chemical degradation of the cells (e.g., permeabilized cells) or ablation of the tissue (e.g., laser ablation) prior to analysis of an analyte bound to a capture probe from the substrate.

In some embodiments, provided herein are methods for spatially detecting an analyte
10 (e.g., detecting the location of an analyte, e.g., a biological analyte) from a biological sample (e.g., present in a biological sample) that include: (a) optionally staining and/or imaging a biological sample on a substrate; (b) permeabilizing (e.g., providing a solution comprising a permeabilization reagent to) the biological sample on the substrate; (c) contacting the biological sample with an array comprising a plurality of capture probes, wherein a capture
15 probe of the plurality captures the biological analyte; and (d) analyzing the captured biological analyte, thereby spatially detecting the biological analyte; where the biological sample is not removed from the substrate.

In some embodiments, provided herein are methods for spatially detecting a biological analyte of interest from a biological sample that include: (a) staining and imaging a
20 biological sample on a substrate; (b) providing a solution comprising a permeabilization reagent to the biological sample on the substrate; (c) contacting the biological sample with an array on a substrate, wherein the array comprises one or more capture probe pluralities thereby allowing the one or more pluralities of capture probes to capture the biological analyte of interest; and (d) analyzing the captured biological analyte, thereby spatially
25 detecting the biological analyte of interest; where the biological sample is not removed from the substrate.

In some embodiments, the method further includes subjecting a region of interest in the biological sample to spatial transcriptomic analysis. In some embodiments, one or more of the capture probes includes a capture domain. In some embodiments, one or more of the
30 capture probes comprises a unique molecular identifier (UMI). In some embodiments, one or more of the capture probes comprises a cleavage domain. In some embodiments, the cleavage domain comprises a sequence recognized and cleaved by a uracil-DNA glycosylase, apurinic/apyrimidinic (AP) endonuclease (APE1), U uracil-specific excision reagent (USER),

and/or an endonuclease VIII. In some embodiments, one or more capture probes do not comprise a cleavage domain and is not cleaved from the array.

In some embodiments, a capture probe can be extended (an “extended capture probe,” e.g., as described herein). For example, extending a capture probe can include generating cDNA from a captured (hybridized) RNA. This process involves synthesis of a complementary strand of the hybridized nucleic acid, e.g., generating cDNA based on the captured RNA template (the RNA hybridized to the capture domain of the capture probe). Thus, in an initial step of extending a capture probe, e.g., the cDNA generation, the captured (hybridized) nucleic acid, e.g., RNA, acts as a template for the extension, e.g., reverse transcription, step.

In some embodiments, the capture probe is extended using reverse transcription. For example, reverse transcription includes synthesizing cDNA (complementary or copy DNA) from RNA, e.g., (messenger RNA), using a reverse transcriptase. In some embodiments, reverse transcription is performed while the tissue is still in place, generating an analyte library, where the analyte library includes the spatial barcodes from the adjacent capture probes. In some embodiments, the capture probe is extended using one or more DNA polymerases.

In some embodiments, a capture domain of a capture probe includes a primer for producing the complementary strand of a nucleic acid hybridized to the capture probe, e.g., a primer for DNA polymerase and/or reverse transcription. The nucleic acid, e.g., DNA and/or cDNA, molecules generated by the extension reaction incorporate the sequence of the capture probe. The extension of the capture probe, e.g., a DNA polymerase and/or reverse transcription reaction, can be performed using a variety of suitable enzymes and protocols.

In some embodiments, a full-length DNA (e.g., cDNA) molecule is generated. In some embodiments, a “full-length” DNA molecule refers to the whole of the captured nucleic acid molecule. However, if a nucleic acid (e.g., RNA) was partially degraded in the tissue sample, then the captured nucleic acid molecules will not be the same length as the initial RNA in the tissue sample. In some embodiments, the 3' end of the extended probes, e.g., first strand cDNA molecules, is modified. For example, a linker or adaptor can be ligated to the 3' end of the extended probes. This can be achieved using single stranded ligation enzymes such as T4 RNA ligase or CircLigase™ (available from Lucigen, Middleton, WI). In some embodiments, template switching oligonucleotides are used to extend cDNA in order to generate a full-length cDNA (or as close to a full-length cDNA as possible). In some embodiments, a second strand synthesis helper probe (a partially double stranded DNA

molecule capable of hybridizing to the 3' end of the extended capture probe), can be ligated to the 3' end of the extended probe, e.g., first strand cDNA, molecule using a double stranded ligation enzyme such as T4 DNA ligase. Other enzymes appropriate for the ligation step are known in the art and include, e.g., Tth DNA ligase, Taq DNA ligase, *Thermococcus* sp.

5 (strain 9°N) DNA ligase (9°NTM DNA ligase, New England Biolabs), AmpligaseTM (available from Lucigen, Middleton, WI), and SplintR (available from New England Biolabs, Ipswich, MA). In some embodiments, a polynucleotide tail, e.g., a poly(A) tail, is incorporated at the 3' end of the extended probe molecules. In some embodiments, the polynucleotide tail is incorporated using a terminal transferase active enzyme.

10 In some embodiments, double-stranded extended capture probes are treated to remove any unextended capture probes prior to amplification and/or analysis, e.g., sequence analysis. This can be achieved by a variety of methods, e.g., using an enzyme to degrade the unextended probes, such as an exonuclease enzyme, or purification columns.

In some embodiments, extended capture probes are amplified to yield quantities that
15 are sufficient for analysis, e.g., via DNA sequencing. In some embodiments, the first strand of the extended capture probes (e.g., DNA and/or cDNA molecules) acts as a template for the amplification reaction (e.g., a polymerase chain reaction).

In some embodiments, the amplification reaction incorporates an affinity group onto the extended capture probe (e.g., RNA-cDNA hybrid) using a primer including the affinity
20 group. In some embodiments, the primer includes an affinity group and the extended capture probes includes the affinity group. The affinity group can correspond to any of the affinity groups described previously.

In some embodiments, the extended capture probes including the affinity group can be coupled to a substrate specific for the affinity group. In some embodiments, the substrate
25 can include an antibody or antibody fragment. In some embodiments, the substrate includes avidin or streptavidin and the affinity group includes biotin. In some embodiments, the substrate includes maltose and the affinity group includes maltose-binding protein. In some embodiments, the substrate includes maltose-binding protein and the affinity group includes maltose. In some embodiments, amplifying the extended capture probes can function to
30 release the extended probes from the surface of the substrate, insofar as copies of the extended probes are not immobilized on the substrate.

In some embodiments, the extended capture probe or complement or amplicon thereof is released. The step of releasing the extended capture probe or complement or amplicon thereof from the surface of the substrate can be achieved in a number of ways. In some

embodiments, an extended capture probe or a complement thereof is released from the array by nucleic acid cleavage and/or by denaturation (e.g., by heating to denature a double-stranded molecule).

In some embodiments, the extended capture probe or complement or amplicon thereof is released from the surface of the substrate (e.g., array) by physical means. For example, where the extended capture probe is indirectly immobilized on the array substrate, e.g., via hybridization to a surface probe, it can be sufficient to disrupt the interaction between the extended capture probe and the surface probe. Methods for disrupting the interaction between nucleic acid molecules include denaturing double stranded nucleic acid molecules are known in the art. A straightforward method for releasing the DNA molecules (i.e., of stripping the array of extended probes) is to use a solution that interferes with the hydrogen bonds of the double stranded molecules. In some embodiments, the extended capture probe is released by an applying heated solution, such as water or buffer, of at least 85°C, e.g., at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99°C. In some embodiments, a solution including salts, surfactants, etc. that can further destabilize the interaction between the nucleic acid molecules is added to release the extended capture probe from the substrate.

In some embodiments, where the extended capture probe includes a cleavage domain, the extended capture probe is released from the surface of the substrate by cleavage. For example, the cleavage domain of the extended capture probe can be cleaved by any of the methods described herein. In some embodiments, the extended capture probe is released from the surface of the substrate, e.g., via cleavage of a cleavage domain in the extended capture probe, prior to the step of amplifying the extended capture probe.

In some embodiments, probes complementary to the extended capture probe can be contacted with the substrate. In some embodiments, the biological sample can be in contact with the substrate when the probes are contacted with the substrate. In some embodiments, the biological sample can be removed from the substrate prior to contacting the substrate with probes. In some embodiments, the probes can be labeled with a detectable label (e.g., any of the detectable labels described herein). In some embodiments, probes that do not specially bind (e.g., hybridize) to an extended capture probe can be washed away. In some embodiments, probes complementary to the extended capture probe can be detected on the substrate (e.g., imaging, any of the detection methods described herein).

In some embodiments, probes complementary to an extended capture probe can be about 4 nucleotides to about 100 nucleotides long. In some embodiments, probes (e.g., detectable probes) complementary to an extended capture probe can be about 10 nucleotides

to about 90 nucleotides long. In some embodiments, probes (e.g., detectable probes) complementary to an extended capture probe can be about 20 nucleotides to about 80 nucleotides long. In some embodiments, probes (e.g., detectable probes) complementary to an extended capture probe can be about 30 nucleotides to about 60 nucleotides long. In some
5 embodiments, probes (e.g., detectable probes) complementary to an extended capture probe can be about 40 nucleotides to about 50 nucleotides long. In some embodiments, probes (e.g., detectable probes) complementary to an extended capture probe can be about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25,
10 about 26, about 27, about 28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36, about 37, about 38, about 39, about 40, about 41, about 42, about 43, about 44, about 45, about 46, about 47, about 48, about 49, about 50, about 51, about 52, about 53, about 54, about 55, about 56, about 57, about 58, about 59, about 60, about 61, about 62, about 63, about 64, about 65, about 66, about 67, about 68, about 69, about 70,
15 about 71, about 72, about 73, about 74, about 75, about 76, about 77, about 78, about 79, about 80, about 81, about 82, about 83, about 84, about 85, about 86, about 87, about 88, about 89, about 90, about 91, about 92, about 93, about 94, about 95, about 96, about 97, about 98, and about 99 nucleotides long.

In some embodiments, about 1 to about 100 probes can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe. In some embodiments, about 1 to about 10 probes can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe. In some embodiments, about 10 to about 100 probes can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe. In some embodiments, about 20 to about 90 probes can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe. In some embodiments, about 30 to about 80 probes (e.g., detectable probes) can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe. In some embodiments, about 40 to about 70 probes can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe. In some embodiments, about 50 to about 60 probes can be
25 contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe. In some embodiments, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36,
30 about 37, about 38, about 39, about 40, about 41, about 42, about 43, about 44, about 45, about 46, about 47, about 48, about 49, about 50, about 51, about 52, about 53, about 54, about 55, about 56, about 57, about 58, about 59, about 60, about 61, about 62, about 63, about 64, about 65, about 66, about 67, about 68, about 69, about 70, about 71, about 72, about 73, about 74, about 75, about 76, about 77, about 78, about 79, about 80, about 81, about 82, about 83, about 84, about 85, about 86, about 87, about 88, about 89, about 90, about 91, about 92, about 93, about 94, about 95, about 96, about 97, about 98, and about 99 probes can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe.

about 37, about 38, about 39, about 40, about 41, about 42, about 43, about 44, about 45, about 46, about 47, about 48, about 49, about 50, about 51, about 52, about 53, about 54, about 55, about 56, about 57, about 58, about 59, about 60, about 61, about 62, about 63, about 64, about 65, about 66, about 67, about 68, about 69, about 70, about 71, about 72, about 73, about 74, about 75, about 76, about 77, about 78, about 79, about 80, about 81, about 82, about 83, about 84, about 85, about 86, about 87, about 88, about 89, about 90, about 91, about 92, about 93, about 94, about 95, about 96, about 97, about 98, and about 99 probes can be contacted to the substrate and specifically bind (e.g., hybridize) to an extended capture probe.

10 In some embodiments, the probes can be complementary to a single analyte (e.g., a single gene). In some embodiments, the probes can be complementary to one or more analytes (e.g., analytes in a family of genes). In some embodiments, the probes (e.g., detectable probes) can be for a panel of genes associated with a disease (e.g., cancer, Alzheimer's disease, Parkinson's disease).

15 In some instances, the ligated probe and capture probe can be amplified or copied, creating a plurality of cDNA molecules. In some embodiments, cDNA can be denatured from the capture probe template and transferred (e.g., to a clean tube) for amplification, and/or library construction. The spatially-barcoded cDNA can be amplified via PCR prior to library construction. The cDNA can then be enzymatically fragmented and size-selected in order to
20 optimize for cDNA amplicon size. P5 and P7 sequences directed to capturing the amplicons on a sequencing flowcell (Illumina sequencing instruments) can be appended to the amplicons, i7, and i5 can be used as sample indexes, and TruSeq Read 2 can be added via End Repair, A-tailing, Adaptor Ligation, and PCR. The cDNA fragments can then be
25 sequenced using paired-end sequencing using TruSeq Read 1 and TruSeq Read 2 as sequencing primer sites. The additional sequences are directed toward Illumina sequencing instruments or sequencing instruments that utilize those sequences; however a skilled artisan will understand that additional or alternative sequences used by other sequencing instruments or technologies are also equally applicable for use in the aforementioned methods.

In some embodiments, where a sample is barcoded directly via hybridization with
30 capture probes or analyte capture agents hybridized, bound, or associated with either the cell surface, or introduced into the cell, as described above, sequencing can be performed on the intact sample.

A wide variety of different sequencing methods can be used to analyze barcoded analyte (e.g., the ligation product). In general, sequenced polynucleotides can be, for

example, nucleic acid molecules such as deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), including variants or derivatives thereof (e.g., single stranded DNA or DNA/RNA hybrids, and nucleic acid molecules with a nucleotide analog).

Sequencing of polynucleotides can be performed by various systems. More generally, sequencing can be performed using nucleic acid amplification, polymerase chain reaction (PCR) (e.g., digital PCR and droplet digital PCR (ddPCR), quantitative PCR, real time PCR, multiplex PCR, PCR-based single plex methods, emulsion PCR), and/or isothermal amplification. Non-limiting examples of methods for sequencing genetic material include, but are not limited to, DNA hybridization methods (e.g., Southern blotting), restriction enzyme digestion methods, Sanger sequencing methods, next-generation sequencing methods (e.g., single-molecule real-time sequencing, nanopore sequencing, and Polony sequencing), ligation methods, and microarray methods.

(l) Kits

In some embodiments, also provided herein are kits that include one or more reagents to detect one or more analytes described herein. In some instances, the kit includes a substrate comprising a plurality of capture probes comprising a spatial barcode and the capture domain. In some instances, the kit includes a plurality of probes (e.g., a first probe, a second probe, one or more spanning probes, and/or a third oligonucleotide).

A non-limiting example of a kit used to perform any of the methods described herein includes: (a) a substrate comprising a plurality of capture probes comprising a spatial barcode and a capture domain; (b) a system comprising: a plurality of first probes and second probes, wherein a first probe and a second probe each comprises sequences that are substantially complementary to an analyte, and wherein the second probe comprises a capture binding domain; and (c) instructions for performing the method of any one of the preceding claims.

Another non-limiting example of a kit used to perform any of the methods described herein includes: (a) an array comprising a plurality of capture probes; (b) a plurality of probes comprising a first probe and a second, wherein the first probe and the second probe are substantially complementary to adjacent sequences of an analyte, wherein the second probe comprises (i) a capture probe binding domain that is capable of binding to a capture domain of the capture probe and (ii) a linker sequence; (c) a plurality of enzymes comprising a ribonuclease and a ligase; and (d) instructions for performing the method of any one of the preceding claims.

Another non-limiting example of a kit used to perform any of the methods described herein includes: (a) an array comprising a plurality of capture probes; (b) a plurality of probes comprising a first probe and a second probe, wherein the first probe and the second probe are substantially complementary to adjacent sequences of an analyte, wherein the first probe
 5 includes a linker sequence, wherein the second probe comprises a capture probe binding domain that is capable of binding to a capture domain of the capture probe; (c) a plurality of enzymes comprising a ribonuclease and a ligase; and (d) instructions for performing the method of any one of the preceding claims.

In some embodiments of any of the kits described herein, the kit includes a second
 10 probe that includes a preadenylated phosphate group at its 5' end and a first probe comprising at least two ribonucleic acid bases at the 3' end.

EXAMPLES

Example 1. Spatial gene expression analysis of FFPE-fixed samples using RNA- 15 templated ligation

Others have demonstrated in situ ligation. See Credle *et al.*, *Nucleic Acids Research*, Volume 45, Issue 14, 21 August 2017, Page e128, <https://doi.org/10.1093/nar/gkx471> (2017). However, the previous approaches have utilized hybridization using poly(A) tails which led to off-target binding. Here, the poly(A) tail on one probe oligonucleotide was switched to
 20 another sequence.

As an overview, a non-limiting example of RNA-templated ligation on an FFPE-fixed sample was performed as described in **FIG. 13**. FFPE-fixed samples were deparaffinized, stained (e.g., H&E stain), and imaged **1301**. Samples were destained (e.g., using HCl) and decrosslinked **1302**. Following decrosslinking, samples treated with pre-hybridization buffer
 25 (e.g., hybridization buffer without the first and second probes), probes were added to the sample, probes hybridized, and samples were washed **1303**. Ligase was added to the samples to ligate hybridized probes to generate a ligation product and samples were then washed **1304**. Probes were released from the analyte by contacting the biological sample with RNase H **1305**. Samples were then permeabilized to facilitate capture of the ligation product by the
 30 capture probes on the substrate **1306**. Ligation products that hybridized to the capture probes were then extended **1307**. The extended capture probes were denatured **1308**. Denatured, extended capture probes were indexed and the amplified libraries were subjected to quality control **1309** before being sequenced.

FFPE sectioned mouse brain tissue slides were deparaffinized, permeabilized with PBST, and equilibrated with a pre-Hyb buffer twice for 5 minutes each. Decrosslinking the sample was performed using either a TE decrosslinking reagent or a PBS-tween decrosslinking reagent. For TE decrosslinking, 100 μ l of TE buffer (pH 9.0) (Genemed 10-0046) was added per sample. Samples treated with TE were subjected to a thermal cycler protocol according to Table 1.

Table 1. Decrosslinking Thermal Cycler Protocol

Decrosslinking: Lid Temp: 70°C, Volume: 100uL		
Step	Temp	Time
Pre-heat	70 °C	∞
Decrosslinking	70 °C	60 minutes
Hold	22 °C	∞

Following the thermal cycler protocol, slides were placed into a metal cassette and subject to the methods described in Table 2.

Table 2. Decrosslinking with HCl*

Step	Timing
1. Wash 1: Added 100 μ l 0.1N HCl	1 minute
2. Wash 2: Removed the HCl and add 100 μ l HCl	1 minute
3. Wash 3: Removed the HCl and add 100 μ l HCl	1 minute
4. Buffer exchange: Removed HCl and add 100 μ l TE buffer (pH 9.0)	
5. Removed TE and add 100 μ l TE (pH 9.0).	
6. Incubated at 70°C	1 hour
7. Removed TE and add 100 μ l 1x PBS-Tween (0.05%)	
8. Incubated at room temperature	15 minutes

* Note: All liquid was removed at each step.

RTL probes were designed to hybridize to adjacent sequences of each analyte (e.g., mRNA sequence) of interest in the genome, including estrogen receptor, progesterone receptor, and ERBB2, also known as HER2. Here, 20,056 probe pairs (e.g., RHS and LHS probes) were added to each tissue sample to capture 19,490 different genes. Two RTL probes (a left-hand side (LHS) probe and a right-hand side (RHS) probe (see e.g., **FIG. 6**)) for each analyte were added simultaneously and hybridized at adjacent sequences of the target mRNA, forming RNA:DNA duplex structures. Following decrosslinking, the PBS-Tween was removed and the DNA probes (1 nM of each probe) were added to the tissue samples in a hybridization buffer for hybridizing the DNA probes to their respective mRNA targets. The

hybridization buffer included SSC, formamide or an equivalent, yeast tRNA as carrier and the RHS and LHS DNA probes.

- One probe oligonucleotide (e.g., the RHS probe or the 3' probe) comprises a non-target functional sequence at its 5' end while the other probe oligonucleotide (e.g., the LHS probe or the 5' probe) comprises a non-target polyA sequence at their 3' ends. The DNA probes were added to the tissue samples and incubated at 50 °C for 2 hours and 30 minutes according to thermal cycler protocol described in Table 3.

Table 3. Hybridization Protocol

Hybridization Thermal Cycler Protocol		
Hyb: Lid Temp: 50C, Volume: 100uL		
Step	Temp	Time
Pre-heat	50 °C	∞
Hyb	50 °C	2 hours and 30 minutes
Hold for post hyb washes	50 °C	∞
Hold	22 °C	∞

10

Following incubation at 50 °C for 2 ½ hours, the hybridization buffer was removed and the tissue was washed twice with each wash including 5 minutes at 50 °C with post-hybridization wash buffer pre-heated to 50 °C. Post-hybridization wash buffer (Post-hyb buffer) included: SSC, yeast tRNA and nuclease free water.

- 15 To ligate two diribo probe oligonucleotides that adjacently hybridized on the target mRNA as described above, each sample was incubated with RNL2 T4 DNA ligase reaction mixture (ligation mix). The ligase reaction mixture included: T4 RNA Ligase Buffer (NEB B0293S), RNL2 ligase and nuclease free water.

- 20 To ligate two DNA probe oligonucleotides that adjacently hybridized on the target mRNAs as described above, SplintR (NEB) in a ligation buffer was added to each sample. The ligase reaction mixture included: Tris-HCl (pH 7.5) MnCl₂ ATP, DTT, surrogate fluid, SplintR ligase (NEB), and nuclease free water per sample. The ligation and wash steps were performed as described in Table 4.

- 25 Following DNA probe ligation the tissue samples were washed twice for 5 minutes at 60°C in a SSC/formamide post ligation wash buffer.

Table 4. Ligation Protocol

Probe Ligation

Step	Time
Removed ALL Post-Hyb Buffer before adding Ligase Mix Added Ligase Mix and Incubated at 37 °C	1 h
Wash 1: Removed Ligase Mix, added post-ligation wash buffer*, and incubated at 60 °C for 5 minutes Remaining wash buffer at heated to 60 °C	5min
Wash 2: Removed post-ligation wash buffer, added pre-heated post-ligation buffer and incubated at 60 °C for 5 minutes.	5min
Removed post-ligation wash buffer and added 2X SSC**. Repeated 2X SSC wash.	
The cassette was cooled down to room temperature before adding RNase	

* Post ligation wash buffer included (per sample): SSC, formamide or an equivalent, and nuclease free water.

RNase H was added to digest the RNA strand of the hybridized RNA:DNA duplex.

- 5 Briefly, the RNA of the DNA:RNA hybrids was digested by incubating the samples with RNase H mix for 30 minutes at 37°C, where the RNase H mix included: RNase H buffer and RNaseH. Following the incubation and while the same remained at 37°C, the biological sample was permeabilized to release the ligated RTL probes using 1.25mg/mL Proteinase K. In particular, the Proteinase K solution included (per sample): Tris (pH 7.5), MgCl₂,
- 10 Sarkosyl, or SDS , Proteinase K (Enzyme), and nuclease free water. The sample was incubated at 37°C for at least 5 minutes in the Proteinase K solution. The samples were then washed three times with 2X SSC.

- The released, ligated DNA probes that served as a proxy for the target mRNA were allowed to hybridize to the capture domain on the capture probe immobilized on the spatial
- 15 array via the polyA tail on the 3' end of the RHS probe. The captured ligated probes were copied, using the capture probe as a template and the extension product was released from the spatial array. Briefly, the tissues were incubated with a second strand extension mix comprising Kapa Hifi DNA polymerase for 25 minutes at 53°C. Following incubation, the second strand extension mix was removed from the tissues and the tissues were washed with
- 20 2X SSC. A solution of KOH was added to each of the tissue wells, the tissues were incubated

at room temperature for 10 minutes to release the extension product from the spatial array and the supernatant from each tissue well was transferred for quantification, and library preparation. Sample quantification was performed using qPCR and KAPA SYBR FAST qPCR master mix according to the manufacturer's instructions. Briefly, KAPA SYBR master mix was prepared by adding qPCR primer cDNA_F and qPCR primer sRNA_R2. The thermal cycler protocol included: 3 minutes at 98°C; 30 cycles with 5 seconds at 98°C, and 30 seconds at 63°C. For library preparation, samples were indexed using an Amp Mix that included dual indexing primers and an Amp Mix. Nucleic acids were then sequenced and analyzed.

10 As controls, concurrent experiments were run in which no ligase was added, pepsin was added before RNase H treatment, or pepsin was added after RNase H treatment.

As shown in **FIGs. 14-19**, spatial expression of ligated probe pairs highlights underlying mouse brain tissue physiology. **FIG. 14** shows PCR results demonstrating that a ligation product was captured and amplifiable whether the permeabilization occurred before or after RNase H treatment (positive control=standard, negative control=minus ligase). Arrow 1 points to a PCR band that represents desired ligation products, whereas arrows 2, 3 and 4 represent non-ligation products. **FIG. 15** shows non-specific probe detected as a fraction of total reads for each condition. **FIGs. 16A-B** shows most probe combinations are specific. The arrows indicate RHS probes that had increased background and therefore include some non-specific hybridization. Looking at total counts in **FIG. 17**, gene-specific LHS+RHS probes overlay with the tissue footprint (see **FIGs. 17A-C** with black circle indicating tissue footprint).

Further, probes specific for various genes (e.g., Grp88, Penk, Plp1, Nptrx, and Mpb2) were added to the sample in Hybridization buffer first for 30 minutes at 60°C, and then for 2 hours at 45°C (See, e.g., **FIG. 18** and **FIG. 19**). When target-specific UMIs were counted, a more detailed expression pattern was observed. See, e.g., **FIGs. 19B-19F**. This analysis revealed that longer hybridization led to higher sensitivity. Referring to **FIGs. 18A-18E**, white circles indicate spots reporting expression for the indicated gene.

An overnight hybridization was performed comparative to a 2 hour hybridization as discussed. Analysis on mouse brain samples using 21,833 probe pairs (e.g., RHS and LHS probes) added to each tissue sample to capture 21,604 different genes revealed that overnight hybridization, thus longer hybridization, led to higher sensitivity compared to a positive control (see **FIGs. 20A-20B**). The mouse probes were designed from targets derived from Appris (see Rodriguez et al., Nucleic Acids Research, 46: D213-217, doi:

10.1093/nar/gkx997 (2018), which is herein incorporated by reference in its entirety) and GENCODE. All probe pairs were non-overlapping and include generally about 1 probe pair per gene.

5 **Example 2. Spatial gene expression analysis of Triple Positive Breast Cancer (TPBC) using RNA templated ligation (RTL)**

This example demonstrates that RTL can be performed on a sample in order to identify analyte abundance and spatial location in an unbiased manner.

A two-year old triple positive ("TPBC," HER2, estrogen receptor, and progesterone
10 receptor-positive) breast cancer sample preserved by FFPE processing was examined for analyte abundance and spatial location. The TPBC samples were queried with DNA probes via RNA-templated ligation methods. Before the ligation step, the TPBC tissue samples were deparaffinized and stained per established protocols. For example, FFPE TPBC tissue samples were prewarmed in a water bath (40°C), sectioned (10µm), dried at 42°C for several
15 hours and placed in a desiccator at room temperature overnight. The dry, sectioned tissues were deparaffinized by baking at 60°C, moved through a series of xylene and EtOH washes, rinsed in water several times. Following rinsing, the deparaffinized tissues were stained with hematoxylin per established protocols. The stained tissues were imaged, identifying regions of tumor and stroma. See **FIG. 21A**.

20 The tissues were decrosslinked to remove formaldehyde crosslinks within the sample thereby releasing the analytes for RNA templated ligation. Briefly, the tissue samples were incubated with an HCl solution for 1 minute, repeated twice for a total of 3 minutes. Following HCl incubations, the tissue sections were incubated at 70°C for 1 hour in TE pH 9.0. TE was removed and the tissues were incubation in 1x PBS-Tween for 15 minutes.

25 RTL probes were designed to hybridize to adjacent sequences of each analyte (e.g., mRNA sequence) of interest in the genome, including estrogen receptor, progesterone receptor, and ERBB2, also known as HER2.

Probes were designed from targets derived from Appris (see, Rodriguez et al., *Nucleic Acids Research*, 46: D213-217, doi: 10.1093/nar/gkx997 (2018), which is herein incorporated
30 by reference in its entirety) and GENCODE. All probe pairs were non-overlapping and include generally about 1 probe pair per gene.

Here, 20,056 probe pairs (e.g., RHS and LHS probes) were added to each tissue sample to capture 19,490 different genes, including ESR1, PGR and HER2, in the human genome. Two RTL probes (a left-hand side (LHS) probe and a right-hand side (RHS) probe

(see, e.g., **FIG. 6**) for each analyte were added simultaneously and hybridized at adjacent sequences of the target mRNA, forming RNA:DNA duplex structures.

Following decrosslinking, the DNA probes (1nm of each probe) were added to the tissue samples in a hybridization buffer for hybridizing the DNA probes to their respective mRNA targets. One probe oligonucleotide (e.g., the RHS probe or the 3' probe) comprises a non-target functional sequence at its 5' end while the other probe oligonucleotide (e.g., the LHS probe or the 5' probe) comprise a non-target polyA sequence at their 3' ends. Briefly, hybridization buffer with the DNA probes was added to the tissue samples and the tissues were incubated at 50°C of approximately 2 ½ hrs. The hybridization/DNA probe buffer was removed and the tissues washed by addition of a post hybridization buffer without DNA probes and incubation at 50°C for 5 minutes, for a total of 3 post hybridization washes.

To ligate the two DNA probe oligonucleotides that adjacently hybridized on the target mRNAs as described above, SplintR (NEB) in a ligation buffer was added to each tissue sample and the tissues were incubated at 37°C for 60 minutes. Following DNA probe ligation the tissue samples were washed twice for 5 minutes at 60°C in a SSC/formamide post ligation wash buffer.

Next, RNase H was added to digest the RNA strand of the hybridized RNA:DNA duplex. Briefly, the RNA of the DNA:RNA hybrids was digested by incubating the tissues with RNase H for 30 minutes at 37°C. The biological sample then was permeabilized to release the ligated RTL probes and contacted with a plurality of capture probes attached to a slide. In particular, after 30 minutes, the tissues were washed and permeabilized by adding 1.25mg/ml Proteinase K, incubated at 37°C for at least 5 minutes and then washed to remove the protease.

The released, ligated DNA probes that served as a proxy for the target mRNA were allowed to hybridize to the capture domain on the capture probe (capture probes immobilized on the spatial array) via the polyA tail on the 3' end of the RHS probe. The captured ligated probes were copied, using the capture probe as a template and the extension product was released from the spatial array. Briefly, the tissues were incubated with a second strand extension mix comprising Kapa Hifi DNA polymerase (Roche) for 25 minutes at 53°C. Following incubation, the extension mix was removed from the tissues and the tissues were washed with SSC. A solution of KOH was added to each of the tissue wells, the tissues were incubated at room temperature for 10 minutes to release the extension product from the spatial array and the supernatant from each tissue well was transferred for quantitation, library preparation and sequencing on the Illumina NextSeq sequencing instrument.

Assessing the whole transcriptome helped to better understand TPBC heterogeneity. For example, it was observed that using 20,056 probe pairs (e.g., RHS and LHS probes) to capture 19,490 different genes revealed comparable results to a positive control in terms of number of UMIs per cell versus number of genes per cell (**FIG. 21B**). Analysis of spatial location and abundance of each RTL ligated probe that hybridized to the array revealed eight (8) different clusters of expression, demonstrating that differential gene expression and location of the expressed genes can be determined using RTL probes. See **FIGs. 21C-21D**. Referring to **FIG. 21C**, each point on the plot is a spot on the array. Each spot is assigned to a cluster, which are indicated by black circles. In some cases, spots assigned to a cluster are outside the indicated circle on the plot. Referring to **FIG. 21D**, each spot on the array is assigned to a cluster and each cluster is indicated, in part, by circles. In some cases, spots are assigned to a cluster but are not within the indicated circles on the array. Finally, individual analyte expression was determined. A TPBC sample routinely exhibited elevated levels of estrogen receptor, progesterone receptor, and ERBB2 (HER2). As shown in **FIGs. 21E-21GA**.

Embodiment A1. A method for determining a location of an analyte in a biological sample comprising:

- (a) contacting the biological sample with an array comprising a plurality of capture probes, wherein a capture probe of the plurality comprises: (i) a spatial barcode and (ii) a capture domain;
- (b) contacting the biological sample with a first probe oligonucleotide and a second probe oligonucleotide, wherein the first probe oligonucleotide and the second probe oligonucleotide each comprise a sequence that is substantially complementary to adjacent sequences of the analyte, and wherein the second probe oligonucleotide comprises a capture probe capture domain;
- (c) hybridizing the first probe oligonucleotide and the second probe oligonucleotide to the analyte in a formamide-free hybridization buffer;
- (d) ligating the first probe oligonucleotide and the second probe oligonucleotide, thereby generating a ligation product;
- (e) releasing the ligated product from the analyte and hybridizing the ligation product to the capture domain; and
- (f) determining (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of

the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample.

Embodiment A2. The method of Embodiment A1, wherein the capture probe further comprises one or more functional domains, a unique molecular identifier, a cleavage domain, and combinations thereof.

Embodiment A3. The method of Embodiment A1 or A2, wherein the array comprises one or more features on a substrate.

Embodiment A4. The method of Embodiment A3, wherein the one or more features comprises a bead.

Embodiment A5. The method of Embodiment A3, wherein the substrate comprises a slide.

Embodiment A6. The method of any one of Embodiments A1-A5, wherein the formamide-free hybridization buffer is a saline-sodium citrate (SSC) hybridization buffer.

Embodiment A7. The method of Embodiment A6, wherein SSC is present in the SSC hybridization buffer from about 1x SSC to about 6x SSC.

Embodiment A8. The method of Embodiments A6 or A7, wherein SSC is present in the SSC hybridization buffer from about 2x SSC to about 4x SSC.

Embodiment A9. The method of any one of Embodiments A6-A8, wherein the SSC hybridization buffer comprises a solvent.

Embodiment A10. The method of Embodiment A9, wherein the solvent comprises ethylene carbonate.

Embodiment A11. The method of any one of Embodiments A6-A10, wherein ethylene carbonate is present in the SSC hybridization buffer from about 10% (w/v) to about 25% (w/v).

Embodiment A12. The method of any one of Embodiments A6-A11, wherein ethylene carbonate is present in the SSC hybridization buffer from about 15% (w/v) to about 20% (w/v).

Embodiment A13. The method of any one of Embodiments A6-A12, wherein ethylene carbonate is present in the SSC hybridization buffer at about 13% (w/v).

Embodiment A14. The method of any one of Embodiments A6-A13, wherein the SSC hybridization buffer is at a temperature from about 40°C to about 60°C.

Embodiment A15. The method of any one of Embodiments A6-A14, wherein the SSC hybridization buffer is at temperature from about 45°C to about 55°C.

Embodiment A16. The method of any one Embodiments A6-A15, where in the SSC hybridization buffer is at a temperature of about 50°C.

Embodiment A17. The method of any one of Embodiments A6-A16, wherein the SSC hybridization buffer further comprises one or more of a yeast tRNA, a crowder, or an additive.

Embodiment A18. The method of any one of Embodiments A1-A17, wherein ligating in step (d) comprises a ligase.

Embodiment A19. The method of Embodiment A18, wherein the ligase is one or more of a T4 RNA ligase (Rnl2), a SplintR ligase, a single stranded DNA ligase, or a T4 DNA ligase.

Embodiment A20. The method of Embodiment A19, wherein the ligase is a T4 DNA ligase.

Embodiment A21. The method of any one of Embodiments A1-A20, further comprising removing one or more unhybridized first probe oligonucleotides, one or more unhybridized second probe oligonucleotides, or both, from the array.

Embodiment A22. The method of Embodiment A21, wherein the removing comprises washing the one or more unhybridized first probe oligonucleotides, the one or more unhybridized second probe oligonucleotides, or both, from the array in a formamide-free wash buffer.

Embodiment A23. The method of Embodiment A21 or A22, wherein the formamide-free wash buffer is an SSC wash buffer.

Embodiment A24. The method of any one of Embodiments A21-A23, wherein SSC is present in the SSC wash buffer from about .01x SSC to about 1x SSC.

Embodiment A25. The method of any one of Embodiments A21-A24, wherein SSC is present in the SSC wash buffer at about 0.1x SSC.

Embodiment A26. The method of any one of Embodiments A21-A25, wherein the SSC wash buffer comprises a detergent.

Embodiment A27. The method of Embodiment A26, wherein the detergent comprises sodium dodecyl sulfate (SDS).

Embodiment A28. The method of any one of Embodiments A21-A26, wherein SDS is present in the SSC wash buffer from about 0.01% (v/v) to about 0.5% (v/v).

Embodiment A29. The method of any one of Embodiments A21-A28, wherein the SDS is present in the SSC wash buffer at about 0.1% (v/v).

Embodiment A30. The method of any one of Embodiments A21-A29, wherein the SSC wash buffer comprises a solvent.

Embodiment A31. The method of any one of Embodiments A21-A30, wherein the solvent comprises ethylene carbonate.

5 Embodiment A32. The method of any one of Embodiments A21-A31, wherein ethylene carbonate is present in the SSC wash buffer from about 10% (w/v) to about 25% (w/v).

Embodiment A33. The method of any one of Embodiments A21-A32, wherein ethylene carbonate is present in the SSC wash buffer from about 15% (w/v) to about 20%
10 (w/v).

Embodiment A34. The method of any one of Embodiments A21-A33, wherein ethylene carbonate is present in the SSC wash buffer at about 16% (w/v).

Embodiment A35. The method of any one of Embodiments A21-A34, wherein the SSC wash buffer is at a temperature from about 50°C to about 70°C.

15 Embodiment A36. The method of any one of Embodiments A21-A35, wherein the SSC wash buffer is at temperature from about 55°C to about 65°C.

Embodiment A37. The method of any one Embodiments A21-A36, where in the SSC wash buffer is at a temperature of about 60°C.

Embodiment A38. The method of any one of Embodiments A1-A37, wherein
20 releasing in step (e) comprises contacting the ligation product with an endoribonuclease.

Embodiment A39. The method of Embodiment A38, wherein the endoribonuclease is one or more of RNase H, RNase A, RNase C, or RNase I.

Embodiment A40. The method of Embodiment A38 or A39, wherein the endoribonuclease is RNase H.

25 Embodiment A41. The method of Embodiment A40, wherein the RNase H comprises RNase H1, RNase H2, or both.

Embodiment A42. The method of any one Embodiments A1-A41, further comprising extending a 3' end of the capture probe using the ligation product as a template for an extension reaction.

30 Embodiment A43. The method of Embodiment A42, wherein extending the 3' end of the capture probe comprises reverse transcribing the analyte, thereby generating a sequence complementary to the analyte.

Embodiment A44. The method of Embodiment A43, wherein reverse transcribing the analyte comprises a reverse transcriptase.

- Embodiment A45. The method of any of one of Embodiments A1-A44, wherein the analyte is RNA.
- Embodiment A46. The method of any of one of Embodiments A1-A45, wherein the RNA is an mRNA.
- 5 Embodiment A47. The method of any one of Embodiments A1-A46, wherein the biological sample is a tissue sample.
- Embodiment A48. The method of Embodiment A47, wherein the tissue sample is a tissue section.
- Embodiment A49. The method of any one of Embodiments A1-A46, wherein the
10 biological sample a fresh frozen biological sample.
- Embodiment A50. The method of any one of Embodiments A1-A46, wherein the biological sample is a fixed biological sample.
- Embodiment A51. The method of Embodiment A41, wherein the fixed biological sample is a formalin-fixed paraffin-embedded sample.
- 15 Embodiment A52. The method of any one of Embodiments A1-A51, wherein the method further comprises permeabilizing the biological sample.
- Embodiment A53. The method of Embodiment A52, wherein permeabilizing the biological sample occurs before releasing the ligation product from the analyte.
- Embodiment A54. The method of Embodiment A52 or A53, wherein
20 permeabilizing the biological sample comprises an endopeptidase.
- Embodiment A55. The method of any one of Embodiments A1-A54, wherein the method further comprises amplifying the ligation product prior to contacting the biological sample with the array.
- Embodiment A56. The method of any one of Embodiments A1-A55, wherein
25 determining in step (f) comprises sequencing.
- Embodiment A57. The method of any one of Embodiments A1-A56, wherein the method further comprises a capture probe capture domain blocking moiety that specifically binds the capture probe capture domain.
- Embodiment A58. The method of any one of Embodiments A1-A57, further
30 comprising releasing the capture probe capture domain blocking moiety from the capture probe capture domain prior to contacting the biological sample with the array.
- Embodiment A59. The method of any one of Embodiments A1-A58, wherein the capture probe capture domain comprises a homopolymeric sequence.

Embodiment A60. The method of Embodiment A59, wherein the capture probe capture domain comprises a poly(A) sequence.

Embodiment B

5 Embodiment B1. A method for identifying a location of an analyte in a biological sample, the method comprising:

(a) contacting the biological sample with a substrate comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises a capture domain and a spatial barcode;

10 (b) contacting the biological sample with a first probe and a second probe, wherein a portion of the first probe and a portion of the second probe are substantially complementary to adjacent sequences of the analyte,

wherein the first probe comprises a sequence that is substantially complementary to a first target sequence of the analyte,

15 wherein the second probe comprises:

(i) a first sequence that is substantially complementary to a second target sequence of the analyte;

(ii) a linker sequence;

(iii) a second sequence that is substantially complementary to a third target sequence of the analyte; and

20 (iv) a capture probe capture domain that is capable of binding to a capture domain of a capture probe;

(c) hybridizing the first probe and the second probe to the analyte;

(d) ligating the first probe and the second probe, thereby creating a ligation

25 product;

(e) releasing the ligation product from the analyte;

(f) hybridizing the capture probe capture domain to a capture domain; and

(g) determining (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample.

Embodiment B2. The method of Embodiment B1, wherein the second probe comprises from 5' to 3': a first sequence, a linker sequence, a second sequence, and a capture probe capture domain.

Embodiment B3. The method of any one of Embodiments B1-B2, wherein the first target sequence of the analyte is directly adjacent to the second target sequence of the analyte.

Embodiment B4. The method of any one of Embodiments B1-B3, wherein the second target sequence is not directly adjacent to the third target sequence on the analyte.

Embodiment B5. The method of any one of Embodiments B1-B4, wherein the second target sequence and the third target sequence are on different exons of the analyte.

Embodiment B6. The method of any one of Embodiments B1-B4, wherein the second target sequence and the third target sequence are within the same exon of the analyte but are not directly adjacent.

Embodiment B7. The method of any one of Embodiments B1-B6, wherein the linker sequence comprises a total of about 1 nucleotide to about 100 nucleotides.

Embodiment B8. The method of Embodiment B7, wherein the linker further comprises a barcode sequence that serves as a proxy for identifying the analyte.

Embodiment B9. A method for identifying a location of an analyte in a biological sample, the method comprising:

(a) contacting the biological sample with a substrate comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises a capture domain and a spatial barcode;

(b) contacting the biological sample with a first probe and a second probe, wherein a portion of the first probe and a portion of the second probe are substantially complementary to adjacent sequences of the analyte,

wherein the first probe comprises:

(i) a first sequence that is substantially complementary to a first target sequence of the analyte;

(ii) a linker sequence;

(iii) a second sequence that is substantially complementary to a second target sequence of the analyte; and

wherein the second probe comprises a sequence that is substantially complementary to a third target sequence of the analyte and a capture probe capture domain that is capable of binding to a capture domain of a capture probe;

(c) hybridizing the first probe and the second probe to the analyte;

(d) ligating the first probe and the second probe, thereby creating a ligation product;

- (e) releasing the ligation product from the analyte;
- (f) hybridizing the capture probe capture domain to a capture domain; and
- (g) determining (i) all or a part of the sequence of the ligation product specifically bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof, and using the determined sequence of (i) and (ii) to identify the location of the analyte in the biological sample.

Embodiment B10. The method of Embodiment F9, wherein the second target sequence is directly adjacent to the third target sequence.

- Embodiment B11. The method of any one of Embodiments B9-B10, wherein the first probe comprises from 5' to 3': a first sequence, a linker sequence, and a second sequence.

Embodiment B12. The method of any one of Embodiments B9-B11, wherein the first probe further comprises a functional sequence.

- Embodiment B13. The method of Embodiment B13, wherein the functional sequence is a primer sequence.

Embodiment B14. The method of Embodiment B12 or B13, wherein the first probe comprises from 5' to 3': a functional sequence, a first sequence, a linker sequence, and a second sequence.

- Embodiment B15. The method of any one of Embodiments B9-B14, wherein the first target sequence is not directly adjacent to the second target sequence on the analyte.

Embodiment B16. The method of Embodiment B15, wherein the first target sequence and second target sequence of are on different exons.

- Embodiment B17. The method of Embodiment B15, wherein the first target sequence and the second target sequence are within the same exon but are not directly adjacent.

Embodiment B18. The method of any one of Embodiments B9-B17, wherein the linker sequence comprises a total of about 1 nucleotide to about 100 nucleotides.

Embodiment B19. The method of Embodiment B18, wherein the linker further comprises a barcode sequence that serves as a proxy for identifying the analyte.

- Embodiment B20. A method for identifying a location of an analyte in a biological sample, the method comprising:

(a) contacting the biological sample with a substrate comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises a capture domain and a spatial barcode;

- (b) contacting the biological sample with a first probe, a second probe, and one or more spanning probes,
wherein the first probe is substantially complementary to a first portion of the analyte,
wherein the second probe is substantially complementary to a second portion of the
5 analyte and further comprises a capture probe capture domain, and
wherein the spanning probe comprises:
(i) a first sequence that is substantially complementary to a first target
sequence of the analyte, and
(ii) a second sequence that is substantially complementary to a second
10 target sequence of the analyte;
(c) hybridizing the first probe, the second probe, and the spanning probe to the
analyte;
(d) ligating the first probe, the one or more spanning probes, and the second
probe, thereby creating a ligation product that is substantially complementary to the analyte;
15 (e) releasing the ligation product from the analyte;
(f) hybridizing the capture probe capture domain to a capture domain; and
(g) determining (i) all or a part of the sequence of the ligation product specifically
bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of
the spatial barcode, or a complement thereof, and using the determined sequence of (i) and
20 (ii) to identify the location of the analyte in the biological sample.
- Embodiment B21. The method of Embodiment B20, wherein the spanning
oligonucleotide further comprises a functional sequence.
- Embodiment B22. The method of Embodiment B20 or B, wherein the spanning
oligonucleotides comprises from 5' to 3': a first sequence, a functional sequence, and a
25 second sequence.
- Embodiment B23. The method of any one of Embodiments B20-B22, wherein the
functional sequences comprises a linker sequence.
- Embodiment B24. The method of Embodiment B23, wherein the linker sequence
comprises a total of about 1 nucleotides to about 100 nucleotides.
- 30 Embodiment B25. The method of any one of Embodiments B20-B24, wherein the
functional sequences comprises a barcode sequence.
- Embodiment B26. The method of Embodiment B25, wherein the barcode
sequence comprises a sequence that serves as a proxy for identifying the analyte.

Embodiment B27. The method of any one of Embodiments B20-B26, wherein the functional sequence comprises one or more linker sequences and a barcode sequence.

Embodiment B28. The method of Embodiment B27, wherein the functional sequence comprises a barcode sequence flanked by linker sequences.

5 Embodiment B29. The method of any one of Embodiments B20-B28, wherein the linker sequence comprises a total of about 1 nucleotides to about 100 nucleotides.

Embodiment B30. The method of any one of Embodiments B20-B29, wherein the first sequence of the spanning probe and the second sequence of the spanning probe are substantially complementary to sequences within the same exon.

10 Embodiment B31. The method of Embodiment B30, wherein the first target sequence of the analyte and the second target of the analyte are located within the same exon.

Embodiment B32. The method of any one of Embodiments B20-B29, wherein the first sequence of the spanning probe and the second sequence of the spanning probe are substantially complementary to sequences within the different exons of the same gene.

15 Embodiment B33. The method of Embodiment B32, wherein the first target sequence of the analyte and the second target sequence of the analyte are located on different exons of the same gene.

Embodiment B34. The method of any one of Embodiments B20-33, wherein the first portion of the analyte is directly adjacent to the first target sequence, and/or wherein the
20 second portion of the analyte is directly adjacent to the second target sequence.

Embodiment B35. The method of any one of Embodiments B20-B34, wherein the spanning probe comprises at least two ribonucleic acid based at the 3' end.

Embodiment B36. The method of any one of Embodiments B20-B35, wherein the spanning probe comprises a phosphorylated nucleotide at the 5' end.

25 Embodiment B37. The method of any one of Embodiments B20-B36, wherein the one or more spanning probes comprises one spanning probe.

Embodiment B38. The method of any one of Embodiments B20-B36, wherein the one or more spanning probes comprises at least two, at least three, at least four, at least five, or more spanning probes.

30 Embodiment B39. The method of Embodiment B38, wherein the one or more spanning probes comprise:

- (i) a third sequence that is substantially complementary to a third target sequence of the analyte, and

(ii) a fourth sequence that is substantially complementary to a fourth target sequence of the analyte.

Embodiment B40. The method of Embodiment B39, wherein the first target sequence is located in a first exon, the second target sequence is located in a second exon,
5 and the third target sequence and the fourth target sequence are located in a third exon.

Embodiment B41. The method of Embodiments B39, wherein the first target sequence is located in a first exon, the second target sequence is located in a second exon, and the third target sequence is located in a third exon, and the fourth target sequence is located in a fourth exon.

10 Embodiment B42. The method of any one of Embodiments B38-B41, wherein the method comprises ligating:

the first probe to the spanning probe,
the spanning probe to one or more additional spanning probes, and
the one or more additional spanning probes spanning oligonucleotide to the
15 second probe, thereby creating a ligation product that is substantially complementary to the analyte.

Embodiment B43. The method of any one of Embodiments B38-B42, wherein the one or more additional spanning probes oligonucleotide further comprises a functional sequence.

20 Embodiment B44. The method of Embodiment B43, wherein the functional sequences comprises (i) a linker sequence, (ii) a barcode sequence, or (ii) one or more linkers and a barcode sequence.

Embodiment B45. The method of any one of Embodiments B38-B44, wherein the one or more additional spanning probes comprises at least two ribonucleic acid based at the
25 3' end.

Embodiment B46. The method of any one of Embodiments B38-B45, wherein the one or more additional spanning probes comprises a phosphorylated nucleotide at the 5' end.

Embodiment B47. The method of any one of Embodiments B20-B46, wherein the first probe further comprises a functional sequence.

30 Embodiment B48. The method of Embodiment B47, wherein the functional sequence is a primer sequence.

Embodiment B49. The method of any one of the preceding Embodiments, wherein the first probe comprises at least two ribonucleic acid bases at the 3' end.

Embodiment B50. The method of any one of the preceding Embodiments, wherein the second probe comprises a phosphorylated nucleotide at the 5' end.

Embodiment B51. The method of any one of the preceding Embodiments, wherein the method further comprises providing a capture probe capture domain blocking moiety that
5 interacts with the capture probe capture domain.

Embodiment B52. The method of Embodiment B51, wherein the method further comprises releasing the capture probe capture domain blocking moiety from the capture probe capture domain prior to step (f).

Embodiment B53. The method of any one of the preceding Embodiments, wherein
10 the capture probe capture domain comprises a poly-adenylated (poly(A)) sequence or a complement thereof.

Embodiment B54. The method of Embodiment B53, wherein the capture probe capture domain blocking moiety comprises a poly-uridine sequence, a poly-thymidine sequence, or both.

Embodiment B55. The method of Embodiment B52, wherein releasing the poly-uridine sequence from the poly(A) sequence comprises denaturing the ligation product or contacting the ligation product with an endonuclease or exonuclease.

Embodiment B56. The method of any one of the preceding Embodiments, wherein the capture probe capture domain comprises a sequence that is complementary to all
20 or a portion of the capture domain of the capture probe.

Embodiment B57. The method of any one of the preceding Embodiments, wherein the capture probe capture domain comprises a degenerate sequence.

Embodiment B58. The method of any one of the preceding Embodiments, wherein the ligation step comprises using enzymatic ligation or chemical ligation.

Embodiment B59. The method of Embodiment B58, wherein the enzymatic
25 ligation utilizes a ligase.

Embodiment B60. The method of Embodiment B59, wherein the ligase is one or more of a T4 RNA ligase (Rnl2), a splintR ligase, a single stranded DNA ligase, or a T4 DNA ligase.

Embodiment B61. The method of Embodiment B60, wherein the ligase is a T4
30 RNA ligase 2 (Rnl2) ligase.

Embodiment B62. The method of any one of the preceding Embodiments, wherein the first probe, the second probe, and the one or more spanning probes are DNA probes.

Embodiment B63. The method of Embodiment B62, wherein the steps (b) and (c) each creates a RNA: DNA hybrid.

Embodiment B64. The method of any one of the preceding Embodiments, wherein steps (b) and (c) are performed at substantially the same time.

5 Embodiment B65. The method of any one of the preceding Embodiments, wherein the biological sample comprises a FFPE sample.

Embodiment B66. The method of Embodiment B65, wherein the tissue sample is the FFPE tissue sample, and the tissue sample is decrosslinked.

10 Embodiment B67. The method of any one of the preceding Embodiments, wherein the biological sample comprises a tissue section.

Embodiment B68. The method of any one of the preceding Embodiments, wherein the biological sample comprises a fresh frozen sample.

Embodiment B69. The method of any one of the preceding Embodiments, wherein the biological sample comprises live cells.

15 Embodiment B70. The method of any one of the preceding Embodiments, wherein the analyte comprises RNA and/or DNA.

Embodiment B71. The method of any one of the preceding Embodiments, wherein the analyte is RNA.

20 Embodiment B72. The method of Embodiment B71, wherein the RNA is an mRNA.

Embodiment B73. The method of any one of the preceding Embodiments, wherein the biological sample was previously stained.

Embodiment B74. The method of Embodiment B73, wherein the biological sample was previously stained using hematoxylin and eosin (H&E).

25 Embodiment B75. The method of Embodiment B73 or B74, wherein the biological sample was previously stained using immunofluorescence or immunohistochemistry.

Embodiment B76. The method of any one of the preceding Embodiments, wherein the method further comprises contacting the biological sample with a permeabilization agent.

30 Embodiment B77. The method of any one of the preceding Embodiments, wherein the releasing step comprises removing the ligated probe from the analyte.

Embodiment B78. The method of Embodiment B77, wherein the releasing step comprises contacting the ligated probe with an endoribonuclease.

Embodiment B79. The method of Embodiment B78, wherein the endoribonuclease is one or more of RNase H, RNase A, RNase C, or RNase I.

Embodiment B80. The method of Embodiment B79, wherein the RNase H comprises RNase H1, RNase H2, or

5 RNase H1 and RNase H2.

Embodiment B81. The method of any one of the preceding Embodiments, wherein the determining step comprises amplifying all or part of the ligation product specifically bound to the capture domain.

Embodiment B82. The method of Embodiment B81, wherein the amplifying is
10 isothermal.

Embodiment B83. The method of Embodiment B81, wherein the amplifying is not isothermal.

Embodiment B84. The method of any one of Embodiments B81-B83, wherein an amplifying product comprises (i) all or part of sequence of the ligation product specifically
15 bound to the capture domain, or a complement thereof, and (ii) all or a part of the sequence of the spatial barcode, or a complement thereof.

Embodiment B85. The method of any one of the preceding Embodiments, wherein the determining step comprises sequencing.

Embodiment B86. The method of Embodiment B85, wherein the sequencing step
20 comprises in situ sequencing.

Embodiment B87. A kit comprising

(a) an array comprising a plurality of capture probes;

(b) a plurality of probes comprising a first probe and a second oligonucleotide, wherein the first probe and the second probe are substantially complementary to adjacent
25 sequences of an analyte, wherein the second probe comprises (i) a capture probe capture domain that is capable of binding to a capture domain of the capture probe and (ii) a linker sequence;

(c) a plurality of enzymes comprising a ribonuclease and a ligase; and

(d) an instruction for using the kit.

30 Embodiment B88. A kit comprising

(a) an array comprising a plurality of capture probes;

(b) a plurality of probes comprising a first probe and a second oligonucleotide, wherein the first probe and the second probe are substantially complementary to adjacent sequences of an analyte, wherein the first probe includes a linker sequence, wherein the

second probe comprises a capture probe capture domain that is capable of binding to a capture domain of the capture probe;

- (c) a plurality of enzymes comprising a ribonuclease and a ligase; and
- (d) an instruction for using the kit.

5 Embodiment B89. A kit comprising:

- (a) an array comprising a plurality of capture probes;
- (b) a plurality of probes comprising a first probe and a second oligonucleotide,

wherein the second probe comprises a capture probe capture domain that is capable of binding to a capture domain of the capture probe;

10 (c) a plurality of spanning probes, wherein a spanning probe of the plurality of spanning probes comprises a first sequence, a linker sequence, and a second sequence, wherein the first sequence of the spanning probe and the first probe are substantially complementary to adjacent sequences of an analyte, wherein the second sequence of the spanning probe and the second probe are substantially complementary to adjacent sequences
15 of the analyte;

- (d) a plurality of enzymes comprising a ribonuclease and a ligase; and
- (e) an instruction for using the kit.

Embodiment B90. The kit of any one of Embodiments B87-B89, wherein the ribonuclease is RNase H.

20 Embodiment B91. The kit of any one of Embodiments B87-B90, wherein the ligase is one or more of a T4 RNA ligase (Rnl2), a splintR ligase, a single stranded DNA ligase, or a T4 DNA ligase.

Embodiment B92. The kit of Embodiment B91, wherein the ligase is a T4 RNA ligase 2 (Rnl2) ligase.