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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
04 July 2024 (04.07.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/145441 A1

(51) International Patent Classification:
C12Q 1/6806 (2018.01) *C12Q 1/6844* (2018.01)

(21) International Application Number:
PCT/US2023/086181

(22) International Filing Date:
28 December 2023 (28.12.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/435,997 29 December 2022 (29.12.2022) US

(71) Applicant: **10X GENOMICS, INC.** [US/US]; 6230 Stoneridge Mall Road, Pleasanton, California 94588-3260 (US).

(72) Inventors: **KATIRAEE, Layla**; 6230 Stoneridge Mall Road, Pleasanton, California 94588-3260 (US). **HAYES, Ashley**; 6230 Stoneridge Mall Road, Pleasanton, California 94588-3260 (US).

(74) Agent: **PADILLA, Garett M.** et al.; Fish & Richardson P.C., PO Box 1022, Minneapolis, Minnesota 55440 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT,

(54) Title: METHODS, COMPOSITIONS, AND KITS FOR DETERMINING A LOCATION OF A TARGET NUCLEIC ACID IN A FIXED BIOLOGICAL SAMPLE

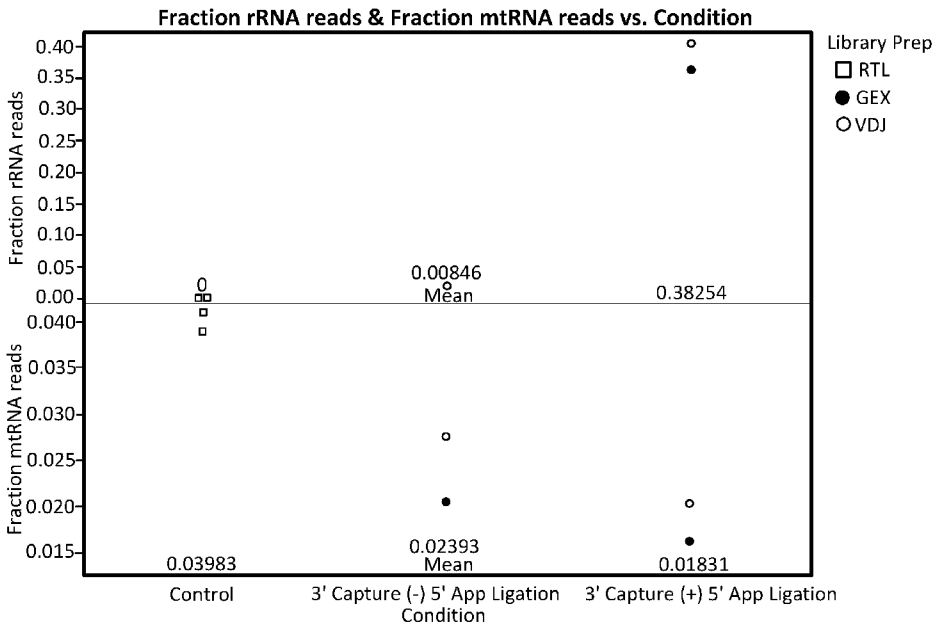


FIG. 13

(57) Abstract: The present disclosure relates to methods, compositions, and kits for determining the location of target nucleic acids and/or processing target nucleic acids from biological samples. The methods typically include incorporating an analyte capture sequence at an end of a nucleic acid analyte in a biological sample *in situ* to facilitate subsequent analyte capture on a substrate including an array of capture probes. In some embodiments, the analyte capture sequence includes a homopolymeric nucleotide sequence, such as a poly(A) sequence.

[Continued on next page]

WO 2024/145441 A1

LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

METHODS, COMPOSITIONS, AND KITS FOR DETERMINING A LOCATION OF A TARGET NUCLEIC ACID IN A FIXED BIOLOGICAL SAMPLE

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims the benefit of priority to U.S. Provisional Patent Application No. 63/435,997, filed on December 29, 2022, the contents of which are incorporated herein by reference in its entirety.

BACKGROUND

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

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

Fixed biological samples generally provide the most convenient method to archive biological samples, such as biological samples obtained from patients. Analytes (e.g., target nucleic acids) can be successfully retrieved and analyzed from fixed biological samples after a substantial amount of time (e.g., years), however, improved methods of retrieving analytes are still needed.

SUMMARY

The preservation of biological samples is useful since preservation can maintain cellular morphology and details of biological samples for extended periods of time. Preserving biological samples also creates an archived library that can be further examined or interrogated retrospectively when more information is known about a particular disease, for example. Different types of preservation methods are known, including, formalin-fixed

paraffin-embedded (FFPE) fixation, paraformaldehyde-fixation (PFA), acetone-fixation, methanol-fixation, and ethanol-fixation.

Methods, compositions, and kits for spatially capturing analytes from fixed biological samples have been previously described. However, there remains a need for improved methods of spatially capturing analytes from preserved biological samples, such as fixed biological samples where there is an increased likelihood of nucleic acid degradation (e.g., degradation of the poly(A) tail). The present disclosure features methods, compositions, and kits for the spatial capture of analytes from a biological sample by incorporating an analyte capture sequence to analytes in a biological sample *in situ* to facilitate subsequent analyte capture on a substrate (e.g., a slide) containing an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes a capture domain and a spatial barcode.

Thus provided herein are methods for determining a location of a target nucleic acid in a biological sample, where the target nucleic acid is a non-polyadenylated target nucleic acid or a truncated polyadenylated target nucleic acid, the method including: (a) providing an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) incorporating an analyte capture sequence to a 3' end of the target nucleic acid in the biological sample, where the biological sample is disposed on a substrate, and where the analyte capture sequence includes a pre-adenylated 5' end and a blocked 3' end, thereby generating a product; (c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe; (d) extending the capture probe using the product as an extension template, thereby generating an extended capture probe; and (e) determining (i) the sequence of the spatial barcode or a complement thereof, and (ii) all or a portion of the sequence of the target nucleic acid or a complement thereof and using the determined sequences of (i) and (ii) to determine the location of the target nucleic acid in the biological sample.

In some embodiments, the target nucleic acid is an RNA. In some embodiments, the RNA is an mRNA, rRNA, tRNA, miRNA, lincRNA, antisense RNA, viral RNA, siRNA, snoRNA, or piRNA. In some embodiments, the target nucleic acid is DNA. In some embodiments, the DNA is genomic DNA.

In some embodiments, the substrate includes the array. In some embodiments, the substrate is a glass slide. In some embodiments, the method includes aligning the substrate with the array, such that at least a portion of the biological sample is aligned with at least a portion of the array, optionally where the array is included in a second substrate.

In some embodiments, incorporating the analyte capture sequence includes ligating the analyte capture sequence to the 3' end of the target nucleic acid. In some embodiments, the ligating includes the use of a ligase. In some embodiments, the ligase includes an RNA ligase, preferably T4 RNA ligase, more preferably T4 RNA ligase 2.

In some embodiments, the extending includes the use of a reverse transcriptase.

In some embodiments, the method includes the use of a template switch oligonucleotide. In some embodiments, the method includes generating a second strand complementary to the extended capture probe. In some embodiments, generating the second strand includes incorporating at least 3 non-templated nucleotides during the extension in step (d). In some embodiments, the method includes hybridizing the template switch oligonucleotide to the at least 3 non-templated nucleotides and extending the extended capture probe using the template switch oligonucleotide as a template.

In some embodiments, the method includes imaging the biological sample. In some embodiments, the method includes staining the biological sample. In some embodiments, the staining includes hematoxylin and/or eosin staining. In some embodiments, the staining includes the use of a detectable label selected from the group consisting of a radioisotope, a fluorophore, a chemiluminescent compound, a bioluminescent compound, or a combination thereof.

In some embodiments, the capture probe includes a cleavage domain, one or more functional domains, a unique molecular identifier, or a combination thereof.

In some embodiments, the method includes permeabilizing the biological sample. In some embodiments, the permeabilizing includes use of a protease. In some embodiments, the protease includes pepsin or proteinase K.

In some embodiments, the array includes one or more features selected from the group consisting of: a spot, an inkjet spot, a masked spot, a pit, a post, a well, a ridge, a divot, a hydrogel pad, and a bead.

In some embodiments, the method includes migrating the product from the biological sample to the array. In some embodiments, the migrating includes electrophoresis.

In some embodiments, the biological sample is a tissue sample. In some embodiments, the tissue sample is a fresh-frozen tissue sample. In some embodiments, the tissue sample is a fixed tissue sample, and optionally, where the fixed tissue sample is a formalin-fixed paraffin-embedded tissue sample, an acetone-fixed tissue sample, a methanol-fixed tissue sample, or a paraformaldehyde-fixed tissue sample. In some embodiments, the biological sample is a tissue section. In some embodiments, the tissue section is a fresh-

frozen tissue section. In some embodiments, the tissue section is a fixed tissue section. In some embodiments, the fixed tissue section is a formalin-fixed paraffin-embedded tissue section, an acetone-fixed tissue section, a methanol-fixed tissue section, or a paraformaldehyde-fixed tissue section.

In some embodiments, the method includes contacting the biological sample with a DNase.

In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more carbon atoms. In some embodiments, the 3' blocked end of the analyte capture sequence includes a biotin moiety. In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more inverted nucleotides.

In some embodiments, the analyte capture sequence includes a homopolymeric nucleotide sequence including DNA or RNA. In some embodiments, the homopolymeric nucleotide sequence includes from about 20 nucleotides to about 50 nucleotides. In some embodiments, the homopolymeric nucleotide sequence includes from about 25 nucleotides to about 35 nucleotides. In some embodiments, the homopolymeric nucleotide sequence includes a poly(A) sequence. In some embodiments, the capture domain of the capture probe includes a poly(T) sequence.

In some embodiments, the method includes contacting the biological sample with one or more ribosomal RNA depletion probes. In some embodiments, the one or more ribosomal RNA depletion probes includes nucleic acid probes complementary to ribosomal RNA and a binding moiety. In some embodiments, the one or more ribosomal RNA depletion probes hybridize to the ribosomal RNA, thereby generating a ribosomal depletion probe/ribosomal RNA complex. In some embodiments, the binding moiety is biotin.

In some embodiments, the method includes contacting the biological sample with one or more mitochondrial RNA depletion probes. In some embodiments, the one or more mitochondrial RNA depletion probes includes nucleic acid probes complementary to mitochondrial RNA and a binding moiety. In some embodiments, the one or more mitochondrial RNA depletion probes hybridize to the mitochondrial RNA, thereby generating a mitochondrial depletion probe/mitochondrial RNA complex. In some embodiments, the binding moiety is biotin.

In some embodiments, the ribosomal depletion probe/ribosomal RNA complex and/or the mitochondrial depletion probe/mitochondrial RNA complex are removed. In some embodiments, the removal includes the use of an RNase. In some embodiments, the RNase is

RNase H1, RNase H2, or a thermostable RNase H. In some embodiments, the removal includes the use of streptavidin.

In some embodiments, the method includes generating a polynucleotide including: (i) the spatial barcode or a complement thereof; (ii) the analyte capture sequence or a complement thereof; and (iii) all or a portion of the sequence of the target nucleic acid or a complement thereof.

In some embodiments, the polynucleotide is generated by amplification of the extended capture probe or a complement thereof. In some embodiments, the complement of the extended capture probe is removed from the array. In some embodiments, the removing includes use of KOH. In some embodiments, the removing includes use of heat.

Also provided herein are kits including: (a) an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) a plurality of analyte capture sequences, where an analyte capture sequence of the plurality of analyte capture sequences is capable of hybridizing to the capture domain and includes a pre-adenylated 5' end and a blocked 3' end; and (c) a ligase.

In some embodiments, the capture probe includes one or more functional domains, a cleavage domain, a unique molecular identifier, or a combination thereof. In some embodiments, the analyte capture sequence includes a homopolymeric nucleotide sequence. In some embodiments, the homopolymeric nucleotide sequence includes a poly(A) sequence.

In some embodiments, the ligase includes an RNA ligase, preferably a T4 RNA ligase, more preferably T4 RNA ligase 2.

In some embodiments, the kit includes one or more permeabilization reagents. In some embodiments, the one or more permeabilization reagents includes a protease, and optionally, where the protease includes proteinase K, pepsin, or collagenase.

In some embodiments, the kit includes a DNase.

In some embodiments, the kit includes one or more ribosomal RNA depletion probes. In some embodiments, the kit includes one or more mitochondrial RNA depletion probes.

In some embodiments, the kit includes instructions for performing any of the methods described herein.

In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more carbon atoms, a biotin moiety, and/or one or more inverted nucleotides.

Also provided herein are methods for enriching RNA from a pool of nucleic acids in a biological sample, where the RNA is a non-polyadenylated nucleic acid or a truncated polyadenylated nucleic acid, the method including: (a) providing an array including a

plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) incorporating an analyte capture sequence to a 3' end of the RNA in the biological sample, where the biological sample is disposed on a substrate, and the analyte capture sequence includes a pre-adenylated 5' end and a blocked 3' end, thereby generating a product; and (c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe, thereby enriching the RNA from a pool of nucleic acids in the biological sample.

In some embodiments, the RNA is an mRNA, rRNA, tRNA, miRNA, lincRNA, antisense RNA, viral RNA, siRNA, snoRNA, or piRNA.

In some embodiments, the substrate is the array. In some embodiments, the substrate is a glass slide. In some embodiments, the method includes aligning the substrate with the array, such that at least a portion of the biological sample is aligned with at least a portion of the array, optionally where the array is included in a second substrate.

In some embodiments, incorporating the analyte capture sequence includes ligating the analyte capture sequence to the 3' end of the RNA. In some embodiments, the ligating includes the use of a ligase. In some embodiments, the ligase includes an RNA ligase, preferably a T4 RNA ligase, more preferably T4 RNA ligase 2.

In some embodiments, the method includes extending the capture probe using the product as an extension template, thereby generating an extended capture probe. In some embodiments, the extending includes the use of a reverse transcriptase.

In some embodiments, the method includes the use of a template switch oligonucleotide. In some embodiments, the method includes generating a second strand complementary to the extended capture probe. In some embodiments, generating the second strand includes adding at least 3 non-templated nucleotides during the extension. In some embodiments, the method includes hybridizing the template switch oligonucleotide to the at least 3 non-templated nucleotides and extending the extended capture probe using the capture probe as a template.

In some embodiments, the method includes imaging the biological sample. In some embodiments, the method includes staining the biological sample. In some embodiments, the staining includes hematoxylin and/or eosin staining. In some embodiments, the staining includes the use of a detectable label selected from the group consisting of a radioisotope, a fluorophore, a chemiluminescent compound, a bioluminescent compound, or a combination thereof.

In some embodiments, the capture probe includes a cleavage domain, one or more functional domains, a unique molecular identifier, and combinations thereof.

In some embodiments, the method includes permeabilizing the biological sample. In some embodiments, the permeabilizing includes the use of a protease. In some embodiments, the protease includes pepsin, collagenase, or proteinase K.

In some embodiments, the array includes one or more features selected from the group consisting of: a spot, an inkjet spot, a masked spot, a pit, a post, a well, a ridge, a divot, a hydrogel pad, and a bead.

In some embodiments, the method includes migrating the product from the biological sample to the array. In some embodiments, the migrating includes electrophoresis.

In some embodiments, the biological sample is a tissue sample. In some embodiments, the tissue sample is a fresh-frozen tissue sample. In some embodiments, the tissue sample is a fixed tissue sample, and optionally, where the fixed tissue sample is a formalin-fixed paraffin-embedded tissue sample, an acetone-fixed tissue sample, a methanol-fixed tissue sample, or a paraformaldehyde-fixed tissue sample. In some embodiments, the biological sample is a tissue section. In some embodiments, the tissue section is a fresh-frozen tissue section. In some embodiments, the tissue section is a fixed tissue section. In some embodiments, the fixed tissue section is a formalin-fixed paraffin-embedded tissue section, an acetone-fixed tissue section, a methanol-fixed tissue section, or a paraformaldehyde-fixed tissue section.

In some embodiments, the method includes contacting the biological sample with a DNase.

In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more carbon atoms. In some embodiments, the 3' blocked end of the analyte capture sequence includes a biotin moiety. In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more inverted nucleotides.

In some embodiments, the analyte capture sequence includes a homopolymeric nucleotide sequence including DNA or RNA. In some embodiments, the homopolymeric nucleotide sequence includes from about 20 nucleotides to about 50 nucleotides. In some embodiments, the homopolymeric nucleotide sequence includes from about 25 nucleotides to about 35 nucleotides. In some embodiments, the homopolymeric nucleotide sequence includes a poly(A) sequence.

In some embodiments, the method includes contacting the biological sample with one or more ribosomal RNA depletion probes. In some embodiments, the one or more ribosomal

RNA depletion probes includes nucleic acid probes complementary to ribosomal RNA and a binding moiety. In some embodiments, the one or more ribosomal RNA depletion probes hybridize to the ribosomal RNA, thereby generating a ribosomal depletion probe/ribosomal RNA complex. In some embodiments, the binding moiety is biotin.

In some embodiments, the method includes contacting the biological sample with one or more mitochondrial RNA depletion probes. In some embodiments, the one or more mitochondrial RNA depletion probes includes nucleic acid probes complementary to mitochondrial RNA and a binding moiety. In some embodiments, the one or more mitochondrial RNA depletion probes hybridize to the mitochondrial RNA, thereby generating a mitochondrial depletion probe/mitochondrial RNA complex. In some embodiments, the binding moiety is biotin.

In some embodiments, the ribosomal depletion probe/ribosomal RNA complex and/or the mitochondrial depletion probe/mitochondrial RNA complex are removed. In some embodiments, the removal includes the use of an RNase. In some embodiments, the RNase is RNase H1, RNase H2, or a thermostable RNase H. In some embodiments, the removal includes the use of streptavidin.

Also provided herein are compositions including: (a) a target nucleic acid, where the target nucleic acid is a non-polyadenylated nucleic acid or a truncated polyadenylated nucleic acid; (b) an analyte capture sequence including a 5' pre-adenylated end and a 3' blocked end, where the analyte capture sequence is capable of hybridizing to a capture domain of a capture probe; and (c) a ligase.

In some embodiments, the ligase includes RNA ligase 2.

In some embodiments, the analyte capture sequence is covalently attached to a 3' end of the target nucleic acid.

In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more carbon atoms. In some embodiments, the 3' blocked end of the analyte capture sequence includes a biotin moiety. In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more inverted nucleotides.

In some embodiments, the analyte capture sequence includes a homopolymeric nucleotide sequence including DNA or RNA. In some embodiments, the homopolymeric nucleotide sequence includes from about 20 nucleotides to about 50 nucleotides. In some embodiments, the homopolymeric nucleotide sequence includes about 30 nucleotides. In some embodiments, the homopolymeric nucleotide sequence includes a poly(A) sequence.

In some embodiments, the composition includes a DNase.

In some embodiments, the composition includes one or more ribosomal RNA depletion probes. In some embodiments, the composition includes one or more mitochondrial RNA depletion probes.

In some embodiments, the composition includes an array including the capture probe, where the capture probe includes a spatial barcode. In some embodiments, the capture probe includes a cleavage domain, one or more functional domains, a unique molecular identifier, and combinations thereof.

In some embodiments, the analyte capture sequence is hybridized to the capture domain of the capture probe.

In some embodiments, the capture probe has been extended using the analyte capture sequence covalently attached to the 3' end of the target nucleic acid as an extension template.

In some embodiments, the target nucleic acid is an RNA. In some embodiments, the RNA is mRNA, rRNA, tRNA, miRNA, viral RNA, siRNA, snoRNA, or piRNA.

Also provided herein are methods for processing a target nucleic acid in a biological sample, the method including: (a) providing an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) incorporating an analyte capture sequence to a 3' end of the target nucleic acid in the biological sample, where the biological sample is disposed on a substrate, and where the analyte capture sequence includes a pre-adenylated 5' end and a blocked 3' end, thereby generating a product; (c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe; and (d) extending the capture probe using the product as an extension template, thereby generating an extended capture probe.

In some embodiments, the target nucleic acid is a non-polyadenylated target nucleic acid or a truncated polyadenylated target nucleic acid.

In some embodiments, the substrate is the array. In some embodiments, the substrate is a glass slide. In some embodiments, the method includes aligning the substrate with the array, such that at least a portion of the biological sample is aligned with at least a portion of the array, optionally where the array is included in a second substrate.

In some embodiments, incorporating the analyte capture sequence includes ligating the analyte capture sequence to the 3' end of the target nucleic acid. In some embodiments, the ligating includes the use of a ligase. In some embodiments, the ligase includes an RNA ligase, preferably a T4 RNA ligase, more preferably T4 RNA ligase 2.

In some embodiments, the extending includes the use of a reverse transcriptase. In some embodiments, the method includes the use of a template switch oligonucleotide. In

some embodiments, the method includes generating a second strand complementary to the extended capture probe. In some embodiments, generating the second strand includes adding at least 3 non-templated nucleotides during the extension in step (d). In some embodiments, the method includes hybridizing the template switch oligonucleotide to the at least 3 non-templated nucleotides and extending the extended capture probe using the extended capture probe as a template.

In some embodiments, the analyte capture sequence includes a homopolymeric nucleotide sequence. In some embodiments, the homopolymeric nucleotide sequence includes a poly(A) sequence, optionally from about 20 nucleotides to about 50 nucleotides.

Also provided herein are methods for determining a location of a target nucleic acid in a biological sample, where the target nucleic acid includes a non-native poly(A) sequence acid, the method including: (a) providing an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) hybridizing the non-native poly(A) sequence of the target nucleic acid to the capture domain of the capture probe; (c) extending the capture probe using the target nucleic as an extension template, thereby generating an extended capture probe; and (d) determining (i) the sequence of the spatial barcode or a complement thereof, and (ii) all or a portion of the sequence of the target nucleic acid or a complement thereof and using the determined sequences of (i) and (ii) to determine the location of the target nucleic acid in the biological sample.

In some embodiments, the non-native poly(A) sequence was previously incorporated at a 3' end of the target nucleic acid in the biological sample (*in situ*), wherein the biological sample is disposed on a substrate, and optionally wherein the non-native poly(A) sequence includes a blocked 3' end.

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.

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.

DESCRIPTION OF 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. 1A shows an exemplary sandwiching process where a first substrate (e.g., a slide), including a biological sample, and a second substrate (e.g., array slide) are brought into proximity with one another.

FIG. 1B shows a fully formed sandwich configuration creating a chamber formed from the one or more spacers, the first substrate, and the second substrate.

FIG. 2A shows a perspective view of an exemplary sample handling apparatus in a closed position.

FIG. 2B shows a perspective view of an exemplary sample handling apparatus in an open position.

FIG. 3A shows the first substrate angled over (superior to) the second substrate.

FIG. 3B shows that as the first substrate lowers, and/or as the second substrate rises, the dropped side of the first substrate may contact a drop of reagent medium.

FIG. 3C shows a full closure of the sandwich between the first substrate and the second substrate with one or more spacers contacting both the first substrate and the second substrate.

FIG. 4A shows a side view of the angled closure workflow.

FIG. 4B shows a top view of the angled closure workflow.

FIG. 5 is a schematic diagram showing an example of a barcoded capture probe, as described herein.

FIG. 6 shows a schematic illustrating a cleavable capture probe.

FIG. 7 shows exemplary capture domains on capture probes.

FIG. 8 shows an exemplary arrangement of barcoded features within an array.

FIG. 9A shows and exemplary workflow for performing a templated capture and producing a ligation product, and **FIG. 9B** shows an exemplary workflow for capturing a ligation product from **FIG. 9A** on a substrate.

FIG. 10 is a schematic diagram of an exemplary analyte capture agent.

FIG. 11 is a schematic diagram depicting an exemplary interaction between a feature-immobilized capture probe **1124** and an analyte capture agent **1126**.

FIGs. 12A-B show quantitative PCR (qPCR) results from a control experiment without 5' app ligation of a pre-adenylated oligonucleotide (**FIG. 12A**) and an experiment with 5' app ligation of a pre-adenylated oligonucleotide (**FIG. 12B**).

FIG. 13 shows ribosomal RNA (rRNA) and mitochondrial RNA (mtRNA) sequencing data under various conditions.

DETAILED DESCRIPTION

A. *Spatial Analysis Methods*

Spatial analysis methodologies 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 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. 11,447,807, 11,352,667, 11,168,350, 11,104,936, 11,008,608,

10,995,361, 10,913,975, 10,774,374, 10,724,078, 10,640,816, 10,494,662, 10,480,022, 10,364,457, 10,317,321, 10,059,990, 10,041,949, 10,030,261, 10,002,316, 9,879,313, 9,783,841, 9,727,810, 9,593,365, 8,951,726, 8,604,182, and 7,709,198; U.S. Patent Application Publication Nos. 2020/0239946, 2020/0080136, 2020/0277663, 2019/0330617, 2020/0256867, 2020/0224244, 2019/0085383, and 2013/0171621; PCT Publication Nos. WO2018/091676, WO2020/176788, WO2017/144338, and WO2016/057552; Non-patent literature references 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 F, dated January 2022); and/or the Visium Spatial Gene Expression Reagent Kits - Tissue Optimization User Guide (e.g., Rev E, dated February 2022), both of which are available at the 10x Genomics Support Documentation website, and can be used herein in any combination, and each of which is incorporated herein by reference in their entireties. 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 Section (I)(b) of PCT Publication No. WO2020/176788 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.

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 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, 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 Section (I)(c) of PCT Publication No. WO2020/176788 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 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 subject. In some embodiments, the biological sample is a tissue sample. In some embodiments, the biological sample (e.g., tissue sample) is a tissue microarray (TMA). A tissue microarray contains multiple representative tissue samples – which can be from different tissues or organisms – assembled on a single histologic slide. The TMA can therefore allow for high throughput analysis of multiple specimens at the same time. Tissue microarrays are paraffin blocks produced by extracting cylindrical tissue cores from different paraffin donor blocks and re-embedding these into a single recipient (microarray) block at defined array coordinates.

The biological sample as used herein can be any suitable biological sample described herein or known in the art. In some embodiments, the biological sample is a tissue. In some embodiments, the tissue sample is a solid tissue sample. In some embodiments, the biological sample is a tissue section. In some embodiments, the tissue is flash-frozen and sectioned. Any suitable method described herein or known in the art can be used to flash-freeze and section the tissue sample. In some embodiments, the biological sample, e.g., the tissue, is flash-frozen using liquid nitrogen before sectioning. In some embodiments, the biological sample, e.g., a tissue sample, is flash-frozen using nitrogen (e.g., liquid nitrogen), isopentane, or hexane.

In some embodiments, the biological sample, e.g., the tissue, is embedded in a matrix e.g., optimal cutting temperature (OCT) compound to facilitate sectioning. OCT compound is a formulation of clear, water-soluble glycols and resins, providing a solid matrix to encapsulate biological (e.g., tissue) specimens. In some embodiments, the sectioning is performed by cryosectioning, for example using a microtome . In some embodiments, the methods further comprise a thawing step, after the cryosectioning.

The biological sample can be from a mammal. In some instances, the biological sample is from a human, mouse, or rat. In addition to the subjects described above, the biological sample can be obtained from non-mammalian organisms (e.g., a plants, an insect, an arachnid, a nematode (e.g., *Caenorhabditis elegans*), a fungi, an amphibian, or a fish (e.g., zebrafish)). A biological sample can be obtained from a prokaryote such as a bacterium, e.g., *Escherichia coli*, *Staphylococci* or *Mycoplasma pneumoniae*; an archaea; a virus such as Hepatitis C virus or human immunodeficiency virus; or a viroid. A biological sample can be obtained from a eukaryote, such as a patient derived organoid (PDO) or patient derived xenograft (PDX). The biological sample can include organoids, a miniaturized and simplified version of an organ produced in vitro in three dimensions that shows realistic micro-anatomy. Organoids can be generated from one or more cells from a tissue, embryonic stem cells, and/or induced pluripotent stem cells, which can self-organize in three-dimensional culture owing to their self-renewal and differentiation capacities. In some embodiments, an organoid is a cerebral organoid, an intestinal organoid, a stomach organoid, a lingual organoid, a thyroid organoid, a thymic organoid, a testicular organoid, a hepatic organoid, a pancreatic organoid, an epithelial organoid, a lung organoid, a kidney organoid, a gastruloid, a cardiac organoid, or a retinal organoid. 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.

Biological samples can be derived from a homogeneous culture or population of the subjects or organisms mentioned herein or alternatively from a collection of several different organisms, for example, in a community or ecosystem.

Biological samples 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. Cancer cells can be derived from solid tumors, hematological malignancies, cell lines, or obtained as circulating tumor cells.

In some embodiments, the biological sample, e.g., the tissue sample, is fixed in a fixative including alcohol, for example methanol. In some embodiments, instead of methanol, acetone, or an acetone-methanol mixture can be used. In some embodiments, the fixation is performed after sectioning. In some instances, the biological sample is not fixed with paraformaldehyde (PFA). In some instances, when the biological sample is fixed with a fixative including an alcohol (e.g., methanol or acetone-methanol mixture), it is not

decrosslinked afterward. In some preferred embodiments, the biological sample is fixed with a fixative including an alcohol (e.g., methanol or an acetone-methanol mixture) after freezing and/or sectioning. In some instances, the biological sample is flash-frozen, and then the biological sample is sectioned and fixed (e.g., using methanol, acetone, or an acetone-methanol mixture). In some instances when methanol, acetone, or an acetone-methanol mixture is used to fix the biological sample, the sample is not decrosslinked at a later step. In instances when the biological sample is frozen (e.g., flash frozen using liquid nitrogen and embedded in OCT) followed by sectioning and alcohol (e.g., methanol, acetone-methanol) fixation or acetone fixation, the biological sample is referred to as "fresh frozen". In some embodiments, fixation of the biological sample e.g., using acetone and/or alcohol (e.g., methanol, acetone-methanol) is performed while the sample is mounted on a substrate (e.g., glass slide, such as a positively charged glass slide).

In some embodiments, the biological sample, e.g., the tissue sample, is fixed e.g., immediately after being harvested from a subject. In such embodiments, the fixative is preferably an aldehyde fixative, such as paraformaldehyde (PFA) or formalin. In some embodiments, the fixative induces crosslinks within the biological sample. In some embodiments, after fixing e.g., by formalin or PFA, the biological sample is dehydrated via sucrose gradient. In some instances, the fixed biological sample is treated with a sucrose gradient and then embedded in a matrix e.g., OCT compound. In some instances, the fixed biological sample is not treated with a sucrose gradient, but rather is embedded in a matrix e.g., OCT compound after fixation. In some embodiments when a fixed frozen tissue sample is treated with a sucrose gradient, it can be rehydrated with an ethanol gradient. In some embodiments, the PFA or formalin fixed biological sample, which can be optionally dehydrated via sucrose gradient and/or embedded in OCT compound, is then frozen e.g., for storage or shipment. In such instances, the biological sample is referred to as "fixed frozen". In preferred embodiments, a fixed frozen biological sample is not treated with methanol. In preferred embodiments, a fixed frozen biological sample is not paraffin embedded. Thus, in preferred embodiments, a fixed frozen biological sample is not deparaffinized. In some embodiments, a fixed frozen biological sample is rehydrated in an ethanol gradient.

In some instances, the biological sample (e.g., a fixed frozen tissue sample) is treated with a citrate buffer. Citrate buffer can be used for antigen retrieval to decrosslink antigens and fixation medium in the biological sample. Thus, any suitable decrosslinking agent can be used in addition to or alternatively to citrate buffer. In some embodiments, for example, the biological sample (e.g., a fixed frozen tissue sample) is decrosslinked with TE buffer.

In any of the foregoing, the biological sample can further be stained, imaged, and/or destained. For example, in some embodiments, a fresh frozen tissue sample or fixed frozen tissue sample is stained (e.g., via eosin and/or hematoxylin), imaged, destained (e.g., via HCl), or a combination thereof. In some embodiments, when a fresh frozen tissue sample is fixed in methanol, it is treated with isopropanol prior to being stained (e.g., via eosin and/or hematoxylin), imaged, destained (e.g., via HCl), or a combination thereof. In some embodiments when a fixed frozen tissue sample is treated with a sucrose gradient, it can be rehydrated with an ethanol gradient before being stained, (e.g., via eosin and/or hematoxylin), imaged, destained (e.g., via HCl), decrosslinked (e.g., via TE buffer or citrate buffer), or a combination thereof. In some embodiments, the biological sample can undergo further fixation (e.g., while mounted on a substrate), stained, imaged, and/or destained. For example, a fixed frozen biological sample may be subject to an additional fixing step (e.g., using PFA) before optional ethanol rehydration, staining, imaging, and/or destaining.

In any of the foregoing, the biological sample can be fixed using PAXgene. For example, the biological sample can be fixed using PAXgene in addition, or alternatively to, a fixative disclosed herein or known in the art (e.g., alcohol, acetone, acetone-alcohol, formalin, paraformaldehyde). PAXgene is a non-cross-linking mixture of different alcohols, acid and a soluble organic compound that preserves morphology and bio-molecules. It is a two-reagent fixative system in which tissue is firstly fixed in a solution containing methanol and acetic acid then stabilized in a solution containing ethanol. See, Ergin B. et al., J Proteome Res. 2010 Oct 1;9(10):5188-96; Kap M. et al., PLoS One.; 6(11):e27704 (2011); and Mathieson W. et al., Am J Clin Pathol.; 146(1):25-40 (2016), each of which are hereby incorporated by reference in their entirety, for a description and evaluation of PAXgene for tissue fixation. Thus, in some embodiments, when the biological sample, e.g., the tissue sample, is fixed in a fixative including alcohol, the fixative is PAXgene. In some embodiments, a fresh frozen tissue sample is fixed with PAXgene. In some embodiments, a fixed frozen tissue sample is fixed with PAXgene.

In some embodiments, the biological sample, e.g., the tissue sample is fixed, for example in methanol, acetone, acetone-methanol, PFA, PAXgene or is formalin-fixed and paraffin-embedded (FFPE). In some embodiments, the biological sample comprises intact cells. In some embodiments, the biological sample is a cell pellet, e.g., a fixed cell pellet, e.g., an FFPE cell pellet. FFPE samples are used in some instances in the RTL methods disclosed herein. A limitation of direct RNA capture for fixed samples is that the RNA integrity of fixed (e.g., FFPE) samples can be lower than a fresh sample, thereby making it

more difficult to capture RNA directly, e.g., by capture of a common sequence such as a poly(A) tail of an mRNA molecule. However, by utilizing RTL probes that hybridize to RNA target sequences in the transcriptome, one can avoid a requirement for RNA analytes to have both a poly(A) tail and target sequences intact. Accordingly, RTL probes can be utilized to beneficially improve capture and spatial analysis of fixed samples. The biological sample, e.g., tissue sample, can be stained, and imaged prior, during, and/or after each step of the methods described herein. Any of the methods described herein or known in the art can be used to stain and/or image the biological sample. In some embodiments, the imaging occurs prior to destaining the sample. In some embodiments, the biological sample is stained using an H&E staining method. In some embodiments, the tissue sample is stained and imaged for about 10 minutes to about 2 hours (or any of the subranges of this range described herein). Additional time may be needed for staining and imaging of different types of biological samples.

The tissue sample can be obtained from any suitable location in a tissue or organ of a subject, e.g., a human subject. In some instances, the sample is a mouse sample. In some instances, the sample is a human sample. In some embodiments, the sample can be derived from skin, brain, breast, lung, liver, kidney, prostate, tonsil, thymus, testes, bone, lymph node, ovary, eye, heart, or spleen. In some instances, the sample is a human or mouse breast tissue sample. In some instances, the sample is a human or mouse brain tissue sample. In some instances, the sample is a human or mouse lung tissue sample. In some instances, the sample is a human or mouse tonsil tissue sample. In some instances, the sample is a human or mouse liver tissue sample. In some instances, the sample is a human or mouse bone, skin, kidney, thymus, testes, or prostate tissue sample. In some embodiments, the tissue sample is derived from normal or diseased tissue. In some embodiments, the sample is an embryo sample. The embryo sample can be a non-human embryo sample. In some instances, the sample is a mouse embryo sample.

Non-limiting examples of stains include histological stains (e.g., hematoxylin and/or eosin) and immunological stains (e.g., fluorescent stains). The biological sample can be stained using 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. In some instances, PAS staining is performed after formalin or acetone fixation. In some embodiments, a biological sample (e.g., a fixed and/or stained biological sample) can be imaged. Biological samples are also described in Section

(I)(d) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

The following embodiments can be used with any of the methods described herein. In some embodiments, the biological sample (e.g., a fixed and/or stained biological sample) is imaged. In some embodiments, the biological sample is visualized or imaged using bright field microscopy. In some embodiments, the biological sample is visualized or imaged using fluorescence microscopy. Additional methods of visualization and imaging are known in the art. Non-limiting examples of visualization and imaging include expansion microscopy, bright field microscopy, dark field microscopy, phase contrast microscopy, electron microscopy, fluorescence microscopy, reflection microscopy, interference microscopy and confocal microscopy. In some embodiments, the sample is stained and imaged prior to adding reagents for analyzing captured analytes as disclosed herein to the biological sample.

In some embodiments, the method includes staining the biological sample. In some embodiments, the staining includes the use of hematoxylin and/or eosin. 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 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 fuchsine, 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 biological 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 embodiments, the staining includes the use of a detectable label selected from the group consisting of a radioisotope, a fluorophore, a chemiluminescent compound, a bioluminescent compound, or a combination thereof.

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 Section (I)(d)(ii)(13) or the Exemplary Embodiments Section of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. Briefly, in any of the methods described herein, the method includes a step of permeabilizing the

biological sample. For example, the biological sample can be permeabilized to facilitate transfer of the extension products to the capture probes on the array. In some embodiments, the permeabilizing includes the use of an organic solvent (e.g., acetone, ethanol, and methanol), a detergent (e.g., saponin, Triton X-100™, Tween-20™, or sodium dodecyl sulfate (SDS)), an enzyme (an endopeptidase, an exopeptidase, a protease), or combinations thereof. In some embodiments, the permeabilizing includes the use of an endopeptidase, a protease, SDS, polyethylene glycol tert-octylphenyl ether, polysorbate 80, and polysorbate 20, N-lauroylsarcosine sodium salt solution, saponin, Triton X-100™, Tween-20™, or combinations thereof. In some embodiments, the endopeptidase is pepsin. In some embodiments, the endopeptidase is Proteinase K. Additional methods for sample permeabilization are described, for example, in Jamur et al., *Method Mol. Biol.* 588:63-66, 2010, the entire contents of which are incorporated herein by reference.

Array-based spatial analysis methods can involve the transfer of one or more analytes or derivatives thereof 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 instances, the capture probe includes a homopolymer sequence, such as a poly(T) sequence. 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., Section (II)(b) (e.g., subsections (i)-(vi)) of PCT Publication No. WO2020/176788 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 Section (II)(d)(ii) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

In some instances, a capture probe and a nucleic acid analyte (or any other nucleic acid to nucleic acid interaction) occurs because the sequences of the two nucleic acids are substantially complementary to one another. By "substantial," "substantially" and the like,

two nucleic acid sequences can be complementary when at least 60% of the nucleotide residues of one nucleic acid sequence are complementary to nucleotide residues in the other nucleic acid sequence. The complementary residues within a particular complementary nucleic acid sequence need not always be contiguous with each other and can be interrupted by one or more non-complementary residues within the complementary nucleic acid sequence. In some embodiments, at least 60%, but less than 100%, of the residues of one of the two complementary nucleic acid sequences are complementary to residues in the other nucleic acid sequence. In some embodiments, at least 70%, 80%, 90%, 95% or 99% of the residues of one nucleic acid sequence are complementary to residues in the other nucleic acid sequence. Sequences are said to be "substantially complementary" when at least 60% (e.g., at least 70%, at least 80%, or at least 90%) of the residues of one nucleic acid sequence are complementary to residues in the other nucleic acid sequence. In some embodiments, the biological sample is mounted on a first substrate and the substrate comprising the array of capture probes is a second substrate. During this process, one or more analytes or analyte derivatives (e.g., intermediate agents; e.g., ligation products) are released from the biological sample and migrate to the second substrate comprising an array of capture probes. In some embodiments, the release and migration of the analytes or analyte derivatives to the second substrate comprising the array of capture probes occurs in a manner that preserves the original spatial context of the analytes in the biological sample. This method can be referred to as a sandwiching process, which is described e.g., in U.S. Patent Application Pub. No. 2021/0189475 and PCT Pub. Nos. WO 2021/252747 A1, WO 2022/061152 A2, and WO 2022/140028 A1.

FIG. 1A shows an exemplary sandwiching process **100** where a first substrate (e.g., slide **103**), including a biological sample **102** (e.g., a parasitic organism), and a second substrate (e.g., array slide **104** including an array having spatially barcoded capture probes **106**) are brought into proximity with one another. As shown in **FIG. 1A** a liquid reagent drop (e.g., permeabilization solution **105**) is introduced on the second substrate in proximity to the capture probes 106 and in between the biological sample **102** and the second substrate (e.g., slide **104** including an array having spatially barcoded capture probes **106**). The permeabilization solution **105** may release analytes or analyte derivatives (e.g., intermediate agents; e.g., ligation products) that can be captured by the capture probes of the array **106**.

During the exemplary sandwiching process, the first substrate is aligned with the second substrate, such that at least a portion of the biological sample is aligned with at least a portion of the capture probes (e.g., aligned in a sandwich configuration). As shown, the

second substrate (e.g., array slide **104**) is in an inferior position to the first substrate (e.g., slide **103**). In some embodiments, the first substrate (e.g., slide **103**) may be positioned superior to the second substrate (e.g., slide **104**). A reagent medium **105** within a gap between the first substrate (e.g., slide **103**) and the second substrate (e.g., slide **104**) creates a liquid interface between the two substrates. The reagent medium may be a permeabilization solution which permeabilizes and/or digests the biological sample **102**. In some embodiments wherein the biological sample **102** has been pre-permeabilized, the reagent medium is not a permeabilization solution. In some embodiments, analytes (e.g., mRNA transcripts) and/or analyte derivatives (e.g., intermediate agents; e.g., ligation products) of the biological sample **102** may release from the biological sample, and actively or passively migrate (e.g., diffuse) across the gap toward the capture probes on the array **106**. Alternatively, in certain embodiments, migration of the analyte or analyte derivative (e.g., intermediate agent; e.g., ligation product) from the biological sample is performed actively (e.g., electrophoretic, by applying an electric field to promote migration). Exemplary methods of electrophoretic migration are described in WO 2020/176788, and US. Patent Application Pub. No. 2021/0189475, each of which is hereby incorporated by reference.

As further shown, one or more spacers **110** may be positioned between the first substrate (e.g., slide **103**) and the second substrate (e.g., array slide **104** including spatially barcoded capture probes **106**). The one or more spacers **110** may be configured to maintain a separation distance between the first substrate and the second substrate. While the one or more spacers **110** is shown as disposed on the second substrate, the spacer may additionally or alternatively be disposed on the first substrate.

In some embodiments, the one or more spacers **110** is configured to maintain a separation distance between first and second substrates that is between about 2 microns and 1 mm (e.g., between about 2 microns and 800 microns, between about 2 microns and 700 microns, between about 2 microns and 600 microns, between about 2 microns and 500 microns, between about 2 microns and 400 microns, between about 2 microns and 300 microns, between about 2 microns and 200 microns, between about 2 microns and 100 microns, between about 2 microns and 25 microns, or between about 2 microns and 10 microns), measured in a direction orthogonal to the surface of first substrate that supports the biological sample. In some instances, the separation distance is about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 microns. In some embodiments, the separation distance is less than 50 microns. In some embodiments, the separation distance

is less than 25 microns. In some embodiments, the separation distance is less than 20 microns. The separation distance may include a distance of at least 2 μm.

FIG. 1B shows a fully formed sandwich configuration **125** creating a chamber **150** formed from the one or more spacers **110**, the first substrate (e.g., the slide **103**), and the second substrate (e.g., the slide **104** including an array **106** having spatially barcoded capture probes) in accordance with some example implementations. In the example of **FIG. 1B**, the liquid reagent (e.g., the permeabilization solution **105**) fills the volume of the chamber **150** and may create a permeabilization buffer that allows analytes (e.g., mRNA transcripts and/or other molecules) or analyte derivatives (e.g., intermediate agents; e.g., ligation products) to diffuse from the biological sample **102** toward the capture probes of the second substrate (e.g., slide **104**). In some aspects, flow of the permeabilization buffer may deflect transcripts and/or molecules from the biological sample **102** and may affect diffusive transfer of analytes or analyte derivatives (e.g., intermediate agents; e.g., ligation products) for spatial analysis. A partially or fully sealed chamber **150** resulting from the one or more spacers **110**, the first substrate, and the second substrate may reduce or prevent flow from undesirable convective movement of transcripts and/or molecules over the diffusive transfer from the biological sample **102** to the capture probes.

The sandwiching process methods described above can be implemented using a variety of hardware components. For example, the sandwiching process methods can be implemented using a sample holder (also referred to herein as a support device, a sample handling apparatus, and an array alignment device). Further details on support devices, sample holders, sample handling apparatuses, or systems for implementing a sandwiching process are described in, e.g., US. Patent Application Pub. No. 2021/0189475, and PCT Publ. No. WO 2022/061152 A2, each of which are incorporated by reference in their entirety.

In some embodiments of a sample holder, the sample holder can include a first member including a first retaining mechanism configured to retain a first substrate comprising a biological sample. The first retaining mechanism can be configured to retain the first substrate disposed in a first plane. The sample holder can further include a second member including a second retaining mechanism configured to retain a second substrate disposed in a second plane. The sample holder can further include an alignment mechanism connected to one or both of the first member and the second member. The alignment mechanism can be configured to align the first and second members along the first plane and/or the second plane such that the sample contacts at least a portion of the reagent medium when the first and second members are aligned and within a threshold distance along an axis

orthogonal to the second plane. The adjustment mechanism may be configured to move the second member along the axis orthogonal to the second plane and/or move the first member along an axis orthogonal to the first plane.

In some embodiments, the adjustment mechanism includes a linear actuator. In some embodiments, the linear actuator is configured to move the second member along an axis orthogonal to the plane of the first member and/or the second member. In some embodiments, the linear actuator is configured to move the first member along an axis orthogonal to the plane of the first member and/or the second member. In some embodiments, the linear actuator is configured to move the first member, the second member, or both the first member and the second member at a velocity of at least 0.1 mm/sec. In some embodiments, the linear actuator is configured to move the first member, the second member, or both the first member and the second member with an amount of force of at least 0.1 lbs.

FIG. 2A is a perspective view of an example sample handling apparatus **200** in a closed position in accordance with some example implementations. As shown, the sample handling apparatus **200** includes a first member **204**, a second member **210**, optionally an image capture device **220**, a first substrate **206**, optionally a hinge **215**, and optionally a mirror **216**. The hinge **215** may be configured to allow the first member **204** to be positioned in an open or closed configuration by opening and/or closing the first member **204** in a clamshell manner along the hinge **215**.

FIG. 2B is a perspective view of the example sample handling apparatus **200** in an open position in accordance with some example implementations. As shown, the sample handling apparatus **200** includes one or more first retaining mechanisms **208** configured to retain one or more first substrates **206**. In the example of **FIG. 2B**, the first member **204** is configured to retain two first substrates **206**, however the first member **204** may be configured to retain more or fewer first substrates **206**.

In some aspects, when the sample handling apparatus **200** is in an open position (e.g., in **FIG. 2B**), the first substrate **206** and/or the second substrate **212** may be loaded and positioned within the sample handling apparatus **200** such as within the first member **204** and the second member **210**, respectively. As noted, the hinge **215** may allow the first member **204** to close over the second member **210** and form a sandwich configuration.

In some aspects, after the first member **204** closes over the second member **210**, an adjustment mechanism of the sample handling apparatus **200** may actuate the first member **204** and/or the second member **210** to form the sandwich configuration for the permeabilization step (e.g., bringing the first substrate **206** and the second substrate **212**

closer to each other and within a threshold distance for the sandwich configuration). The adjustment mechanism may be configured to control a speed, an angle, a force, or the like of the sandwich configuration.

In some embodiments, the biological sample (e.g., sample **102** from **FIG. 1A**) may be aligned within the first member **204** (e.g., via the first retaining mechanism 208) prior to closing the first member **204** such that a desired region of interest of the sample is aligned with the barcoded array of the second substrate (e.g., the slide **104** from **FIG. 1A**), e.g., when the first and second substrates are aligned in the sandwich configuration. Such alignment may be accomplished manually (e.g., by a user) or automatically (e.g., via an automated alignment mechanism). After or before alignment, spacers may be applied to the first substrate **206** and/or the second substrate **212** to maintain a minimum spacing between the first substrate **206** and the second substrate **212** during sandwiching. In some aspects, the permeabilization solution (e.g., permeabilization solution **305**) may be applied to the first substrate **206** and/or the second substrate **212**. The first member **204** may then close over the second member **210** and form the sandwich configuration. Analytes or analyte derivatives (e.g., intermediate agents; e.g., ligation products) may be captured by the capture probes of the array and may be processed for spatial analysis.

In some embodiments, during the permeabilization step, the image capture device **220** may capture images of the overlap area between the biological sample and the capture probes on the array **106**. If more than one first substrates **206** and/or second substrates **212** are present within the sample handling apparatus **200**, the image capture device **220** may be configured to capture one or more images of one or more overlap areas.

Provided herein are methods for delivering a fluid to a biological sample disposed on an area of a first substrate and an array disposed on a second substrate. **FIGs. 3A-3C** depict a side view and a top view of an exemplary angled closure workflow **300** for sandwiching a first substrate (e.g., slide **303**) having a biological sample **302** and a second substrate (e.g., slide **304** having capture probes **306**) in accordance with some exemplary implementations.

FIG. 3A depicts the first substrate (e.g., the slide **303** including a biological sample **302**) angled over (superior to) the second substrate (e.g., slide **304**). As shown, reagent medium (e.g., permeabilization solution) **305** is located on the spacer **310** toward the right-hand side of the side view in **FIG. 3A**. While **FIG. 3A** depicts the reagent medium on the right hand side of side view, it should be understood that such depiction is not meant to be limiting as to the location of the reagent medium on the spacer.

FIG. 3B shows that as the first substrate lowers, and/or as the second substrate rises, the dropped side of the first substrate (e.g., a side of the slide **303** angled toward the second substrate) may contact the reagent medium **305**. The dropped side of the first substrate may urge the reagent medium **305** toward the opposite direction (e.g., towards an opposite side of the spacer **310**, towards an opposite side of the first substrate relative to the dropped side). For example, in the side view of **FIG. 3B** the reagent medium **305** may be urged from right to left as the sandwich is formed.

In some embodiments, the first substrate and/or the second substrate are further moved to achieve an approximately parallel arrangement of the first substrate and the second substrate.

FIG. 3C depicts a full closure of the sandwich between the first substrate and the second substrate with the spacer **310** contacting both the first substrate and the second substrate and maintaining a separation distance and optionally the approximately parallel arrangement between the two substrates. As shown in the top view of **FIG. 3C**, the spacer **310** fully encloses and surrounds the biological sample **302** and the capture probes **306**, and the spacer **310** form the sides of chamber **350** which holds a volume of the reagent medium **305**.

While **FIG. 3C** depicts the first substrate (e.g., the slide **303** including biological sample **302**) angled over (superior to) the second substrate (e.g., slide **304**) and the second substrate comprising the spacer **310**, it should be understood that an exemplary angled closure workflow can include the second substrate angled over (superior to) the first substrate and the first substrate comprising the spacer **310**.

It may be desirable that the reagent medium be free from air bubbles between the substrates to facilitate transfer of target analytes with spatial information. Additionally, air bubbles present between the substrates may obscure at least a portion of an image capture of a desired region of interest. Accordingly, it may be desirable to ensure or encourage suppression and/or elimination of air bubbles between the two substrates (e.g., slide **303** and slide **304**) during a permeabilization step (e.g., step **104**). In some aspects, it may be possible to reduce or eliminate bubble formation between the substrates using a variety of filling methods and/or closing methods. In some instances, the first substrate and the second substrate are arranged in an angled sandwich assembly as described herein. For example, during the sandwiching of the two substrates (e.g., the slide **303** and the slide **304**), an angled closure workflow may be used to suppress or eliminate bubble formation.

FIG. 4A is a side view of the angled closure workflow **400** in accordance with some exemplary implementations. **FIG. 4B** is a top view of the angled closure workflow **400** in accordance with some exemplary implementations. As shown at **405**, reagent medium **401** is positioned to the side of the substrate **402**.

At step **410**, the dropped side of the angled substrate **406** contacts the reagent medium **401** first. The contact of the substrate **406** with the reagent medium **401** may form a linear or low curvature flow front that fills uniformly with the slides closed.

At step **415**, the substrate **406** is further lowered toward the substrate **402** (or the substrate **402** is raised up toward the substrate **406**) and the dropped side of the substrate **406** may contact and may urge the reagent medium toward the side opposite the dropped side and creating a linear or low curvature flow front that may prevent or reduce bubble trapping between the substrates.

At step **420**, the reagent medium **401** fills the gap between the substrate **406** and the substrate **402**. The linear flow front of the liquid reagent may form by squeezing the **401** volume along the contact side of the substrate **402** and/or the substrate **406**. Additionally, capillary flow may also contribute to filling the gap area.

In some embodiments, the reagent medium (e.g., **105** in **FIG. 1A**) comprises a permeabilization agent. In some embodiments, following initial contact between the biological sample and a permeabilization agent, the permeabilization agent can be removed from contact with the biological sample (e.g., by opening sample holder). Suitable agents for this purpose include, but are not limited to, organic solvents (e.g., acetone, ethanol, and methanol), cross-linking agents (e.g., paraformaldehyde), detergents (e.g., saponin, Triton X-100™, Tween-20™, or sodium dodecyl sulfate (SDS)), and enzymes (e.g., trypsin, proteases (e.g., proteinase K). In some embodiments, the detergent is an anionic detergent (e.g., SDS or N-lauroylsarcosine sodium salt solution).

In some embodiments, the reagent medium comprises a lysis reagent. Lysis solutions can include ionic surfactants such as, for example, sarkosyl and sodium dodecyl sulfate (SDS). More generally, chemical lysis agents can include, without limitation, organic solvents, chelating agents, detergents, surfactants, and chaotropic agents. In some embodiments, the reagent medium comprises a protease. Exemplary proteases include, e.g., pepsin, trypsin, elastase, and proteinase K. In some embodiments, the reagent medium comprises a nuclease. In some embodiments, the nuclease comprises an RNase. In some embodiments, the RNase is selected from RNase A, RNase C, RNase H, and RNase I. In

some embodiments, the reagent medium comprises one or more of sodium dodecyl sulfate (SDS) or a sodium salt thereof, proteinase K, pepsin, N-lauroylsarcosine, and RNase.

In some embodiments, the reagent medium comprises polyethylene glycol (PEG). In some embodiments, the PEG is from about PEG 2K to about PEG 16K. In some embodiments, the PEG is PEG 2K, 3K, 4K, 5K, 6K, 7K, 8K, 9K, 10K, 11K, 12K, 13K, 14K, 15K, or 16K. In some embodiments, the PEG is present at a concentration from about 2% to 25%, from about 4% to about 23%, from about 6% to about 21%, or from about 8% to about 20% (v/v).

In certain embodiments a dried permeabilization reagent is applied or formed as a layer on the first substrate or the second substrate or both prior to contacting the biological sample and the array. For example, a permeabilization reagent can be deposited in solution on the first substrate or the second substrate or both and then dried.

In some instances, the aligned portions of the biological sample and the array are in contact with the reagent medium for about 1 minute, about 5 minutes, about 10 minutes, about 12 minutes, about 15 minutes, about 18 minutes, about 20 minutes, about 25 minutes, about 30 minutes, about 36 minutes, about 45 minutes, or about an hour. In some instances, the aligned portions of the biological sample and the array are in contact with the reagent medium for about 1-60 minutes.

In some instances, the device is configured to control a temperature of the first and second substrates. In some embodiments, the temperature of the first and second members is lowered to a first temperature that is below room temperature.

There are at least two methods to associate a spatial barcode with one or more 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). 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 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., Section (II)(b)(vii) of PCT Publication No. WO2020/176788 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 the template.

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 by a reverse transcriptase. In some embodiments, the capture probe is extended using one or more DNA polymerases. In some embodiments, the extended capture probes include the sequence of the capture domain, the sequence of the spatial barcode of the capture probe, and the complementary sequence of the template used for extension 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., sequencing. In some embodiments, extended capture probes (e.g., DNA molecules) can act as 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 Section (II)(a) of PCT Publication No. WO2020/176788 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 using the capture analyte as a template, 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 Section (II)(g) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. Some quality control measures are described in Section (II)(h) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

Spatial information can provide information of medical importance. For example, the methods 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. Exemplary methods for identifying spatial information of biological and/or medical importance can be found in U.S. Patent Application Publication Nos. 2021/0140982, 2021/0198741, and 2021/0199660.

Spatial information can provide information of biological importance. For example, the methods 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 or proximity based 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 types and their co-localization in healthy and diseased 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 Section (II)(c) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663. Exemplary features and geometric attributes of an array can be found in Sections (II)(d)(i), (II)(d)(iii), and (II)(d)(iv) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

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) 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 Section (II)(e) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663.

FIG. 5 is a schematic diagram showing an exemplary capture probe, as described herein. As shown, the capture probe **502** is optionally coupled to a feature **501** by a cleavage domain **503**, such as a disulfide linker. The capture probe can include a functional sequence **504** that are useful for subsequent processing. The functional sequence **504** can include all or a part of sequencer specific flow cell attachment sequence (e.g., a P5 or P7 sequence), all or a part of a sequencing primer sequence, (e.g., a R1 primer binding site, a R2 primer binding site), or combinations thereof. The capture probe can also include a spatial barcode **505**. The capture probe can also include a unique molecular identifier (UMI) sequence **506**. While **FIG. 5** shows the spatial barcode **505** as being located upstream (5') of UMI sequence **506**, it is to be understood that capture probes wherein UMI sequence **506** is located upstream (5') of the spatial barcode **505** is also suitable for use in any of the methods described herein. The capture probe can also include a capture domain **507** to facilitate capture of a target analyte. The capture domain can have a sequence complementary to a sequence of a nucleic acid analyte. The capture domain can have a sequence complementary to a connected probe described herein. The capture domain can have a sequence complementary to a capture handle sequence present in an analyte capture agent. The capture domain can have a sequence complementary to a splint oligonucleotide. Such splint oligonucleotide, in addition to having a sequence complementary to a capture domain of a capture probe, can have a sequence complementary to a sequence of a nucleic acid analyte, a portion of a connected probe described herein, a capture handle sequence described herein, and/or a methylated adaptor described herein.

FIG. 6 is a schematic illustrating a cleavable capture probe, wherein the cleaved capture probe can enter into a non-permeabilized cell and bind to analytes within the sample. The capture probe 6**01** contains a cleavage domain **602**, a cell penetrating peptide **603**, a reporter molecule **604**, and a disulfide bond (-S-S-). **605** represents all other parts of a capture probe, for example a spatial barcode and a capture domain.

FIG. 7 is a schematic diagram of an exemplary multiplexed spatially-barcoded feature. In **FIG. 7**, the feature **701** 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 **702**. One type of capture probe associated with the feature includes the spatial barcode **702** in combination with a poly(T) capture domain **703**, designed to capture mRNA target analytes. A second type of capture probe associated with the feature includes the spatial barcode **702** in combination with a random N-mer capture domain **704** for gDNA analysis. A third type of capture probe associated with the feature includes the spatial barcode **702** in combination with a capture domain complementary to the analyte capture agent of interest **705**. A fourth type of capture probe associated with the feature includes the spatial barcode **702** in combination with a capture probe that can specifically bind a nucleic acid molecule **706** that can function in a CRISPR assay (e.g., CRISPR/Cas9). While only four different capture probe-barcoded constructs are shown in **FIG. 7**, 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. 7** 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.

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, including non-commercialized sequencing systems.

In some embodiments, the spatial barcode **505** and functional sequences **504** is common to all of the probes attached to a given feature. In some embodiments, the UMI sequence **506** of a capture probe attached to a given feature is different from the UMI sequence of a different capture probe attached to the given feature.

FIG. 8 depicts an exemplary arrangement of barcoded features within an array. From left to right, **FIG. 8** shows (L) a slide including six spatially-barcoded arrays, (C) an enlarged schematic of one of the six spatially-barcoded arrays, showing a grid of barcoded features in relation to a biological sample, and (R) 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.).

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 Section (IV) of PCT Publication No. WO2020/176788 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 Section (III) of PCT Publication No. WO2020/176788 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 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 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 probe binding domain (e.g., a poly(A) sequence, a non-homopolymeric sequence). After hybridization to the analyte, a ligase (e.g., a T4 RNA ligase (Rnl2), a PBCV-1 DNA Ligase or *Chlorella* virus DNA Ligase, a single-stranded DNA ligase, or a T4 DNA ligase) ligates the two oligonucleotides together, creating a ligation product. In some instances, the two oligonucleotides hybridize to 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). In some instances, the ligation product is removed using heat. In some instances, the ligation product is removed using KOH. The released ligation 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.

In some instances, one or both of the oligonucleotides may hybridize to genomic DNA (gDNA) which can lead to false positive sequencing data from ligation events on gDNA (off target) in addition to the desired (on target) ligation events on target nucleic acids, (e.g., mRNA). Thus, in some embodiments, the disclosed methods can include contacting the biological sample with a deoxyribonuclease (DNase). The DNase can be an endonuclease or exonuclease. In some embodiments, the DNase digests single- and/or double-stranded DNA. Suitable DNases include, without limitation, a DNase I and a DNase II. Use of a DNase as described can mitigate false positive sequencing data from off target gDNA ligation events.

A non-limiting example of templated ligation methods disclosed herein is depicted in **FIG. 9A**. After a biological sample is contacted with a substrate including a plurality of capture probes and contacted with (a) a first probe **901** having a target-hybridization sequence **903** and a primer sequence **902** and (b) a second probe **904** having a target-hybridization sequence **905** and a capture domain (e.g., a poly-A sequence) **906**, the first probe **901** and a second probe **904** hybridize **910** to an analyte **907**. A ligase **921** ligates **920**

the first probe to the second probe thereby generating a ligation product **922**. The ligation product is released **930** from the analyte **931** by digesting the analyte using an endoribonuclease **932**. The sample is permeabilized **940** and the ligation product **941** is able to hybridize to a capture probe on the substrate. Methods and composition for spatial detection using templated ligation have been described in PCT Publ. No. WO 2021/133849 A1, U.S. Pat. Nos. 11,332,790 and 11,505,828, each of which is incorporated by reference in its entirety.

In some embodiments, as shown in **FIG. 9B**, the ligation product **9001** includes a capture probe capture domain **9002**, which can bind to a capture probe **9003** (e.g., a capture probe immobilized, directly or indirectly, on a substrate **9004**). In some embodiments, methods provided herein include contacting **9005** a biological sample with a substrate **9004**, wherein the capture probe **9003** is affixed to the substrate (e.g., immobilized to the substrate, directly or indirectly). In some embodiments, the capture probe capture domain **9002** of the ligated product specifically binds to the capture domain **9006**. The capture probe can also include a unique molecular identifier (UMI) **9007**, a spatial barcode **9008**, a functional sequence **9009**, and a cleavage domain **9010**.

In some embodiments, methods provided herein include permeabilization of the biological sample such that the capture probe can more easily capture the ligation products (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 **9011** to produce spatially-barcoded full-length cDNA **9012** and **9013** from the captured ligation products (e.g., ligation products).

In some embodiments, the extended ligation products can be denatured **9014** from the capture probe and transferred (e.g., to a clean tube) for amplification, and/or library construction. The spatially-barcoded ligation products can be amplified **9015** via PCR prior to library construction. P5 **9016** and P7 **9019** can be used as sequences that are complementary to sequencing probes for immobilization of the library on the sequencing flow cell and i5 **9017** and i7 **9018** can be used as sample indexes. The amplicons can then be sequenced using paired-end sequencing using TruSeq Read 1 and TruSeq Read 2 as sequencing primer sites.

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 Section (II)(b)(ix) of PCT Publication No. WO2020/176788 and/or Section (II)(b)(viii) U.S. Patent Application Publication No. 2020/0277663.

FIG. 10 is a schematic diagram of an exemplary analyte capture agent **1002** comprised of an analyte-binding moiety **1004** and an analyte-binding moiety barcode domain **1008**. The exemplary analyte-binding moiety **1004** is a molecule capable of binding to an analyte **1006** and the analyte capture agent is capable of interacting with a spatially-barcoded capture probe. The analyte-binding moiety can bind to the analyte **1006** with high affinity and/or with high specificity. The analyte capture agent can include an analyte-binding moiety barcode domain **1008** which serves to identify the analyte binding moiety, and a capture domain which can hybridize to at least a portion or an entirety of a capture domain of a capture probe. The analyte-binding moiety **1004** can include a polypeptide and/or an aptamer. The analyte-binding moiety **1004** can include an antibody or antibody fragment (e.g., an antigen-binding fragment).

FIG. 11 is a schematic diagram depicting an exemplary interaction between a feature-immobilized capture probe **1124** and an analyte capture agent **1126**. The feature-immobilized capture probe **1124** can include a spatial barcode **1108** as well as functional sequences **1106** and a UMI **1110**, as described elsewhere herein. The capture probe can be affixed **1104** to a feature such as a bead **1102**. The capture probe can also include a capture domain **1112** that is capable of binding to an analyte capture agent **1126**. The analyte-binding moiety barcode domain of the analyte capture agent **1126** can include a functional sequence **1118**, analyte binding moiety barcode **1116**, and an analyte capture sequence **1114** that is capable of binding (e.g., hybridizing) to the capture domain **1112** of the capture probe **1124**. The analyte capture agent can also include a linker **1120** that allows the analyte-binding moiety barcode domain (e.g., including the functional sequence **1118**, analyte binding barcode **1116**, and analyte capture sequence **1114**) to couple to the analyte binding moiety **1122**. In some

embodiments, the linker is a cleavable linker. In some embodiments, the cleavable linker is a photo-cleavable linker, a UV-cleavable linker, or an enzyme cleavable linker. In some instances, the cleavable linker is a disulfide linker. A disulfide linker can be cleaved by use of a reducing agent, such as dithiothreitol (DTT), Beta-mercaptoethanol (BME), or Tris (2-carboxyethyl) phosphine (TCEP).

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 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 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 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 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 the Exemplary Embodiments section of PCT Publication No. WO2020/176788 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, the sample can be immersed..." of PCT Publication No. WO2020/176788 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 F, dated January 2022); and/or the Visium Spatial Gene Expression Reagent Kits - Tissue Optimization User Guide (e.g., Rev E, dated February 2022).

In some embodiments, spatial analysis can be performed using dedicated hardware and/or software, such as any of the systems described in Sections (II)(e)(ii) and/or (V) of PCT Publication No. WO2020/176788 and/or U.S. Patent Application Publication No. 2020/0277663, or any of one or more of the devices or methods described in Sections *Control Slide for Imaging*, *Methods of Using Control Slides and Substrates for*, *Systems of Using Control Slides and Substrates for Imaging*, and/or *Sample and Array Alignment Devices and Methods*, *Informational labels* of PCT Publication No. WO2020/123320.

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. 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 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 Publication No. WO2021/102003 and/or U.S. Patent Application Publication No. 2021/0150707, each of which is incorporated herein by reference in their entireties.

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 Publication No. WO2020/053655 and spatial analysis methods are generally described in PCT Publication No. WO2021/102039 and/or U.S. Patent Application Publication No. 2021/0155982, each of which is incorporated herein by reference in their entireties.

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 the *Substrate Attributes* Section, *Control Slide for Imaging* Section of PCT Publication Nos. WO2020/123320, WO 2021/102005, and/or U.S. Patent Application Publication No. 2021/0158522, each of which is incorporated herein by reference in their entireties. 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.

B. ***Spatial Analysis in Biological Samples***

The preservation of biological samples is useful since preservation maintains cellular morphology and detail of biological samples. Preserving biological samples also creates an archived library that can be further examined or interrogated retrospectively. Different types of preservation methods are known, including, for example formalin-fixed paraffin-

embedded (FFPE) fixation, paraformaldehyde-fixation (PFA), methanol-fixation, ethanol-fixation, and acetone-fixation.

Fixed (e.g., FFPE, PFA) processed biological samples are also advantageous because they can be stored long term (e.g., years) at ambient (e.g., room temperature), thus reducing costs associated with storing biological samples (e.g., frozen tissue) at ultra-low temperatures that can run the risk of thawing unexpectedly. Further, fixed (e.g., FFPE, PFA) biological samples can be stored long-term (e.g., years) with no significant effect on the quantity and/or quality of extracted analytes (e.g., nucleic acid, protein, etc.) (See, Kokkat, T.J., et. al., Archived Formalin-Fixed Paraffin-Embedded (FFPE) Blocks: A Valuable Underexploited Resource for Extraction of DNA, RNA, and Protein, *Biopreserv Biobank*, 11(2), 101-106, 10.1089/bio.2012.0052 (2013)).

However, fixation, such as formaldehyde or formalin fixation may lead to degradation of nucleic acids in a biological sample. Thus, there remains a need for improved methods of spatially capturing analytes from preserved biological samples, such as fixed biological samples. In general, target nucleic acids are captured from biological samples (e.g., fixed biological samples) via a templated-ligation reaction where a pair of probes hybridize to the target nucleic acid, are then ligated to one another, and captured by a capture probe on an array where one of the probes includes an analyte capture sequence capable of hybridizing to the capture domain. Generally, this method does not allow for xenografts and probe pairs are typically only available for either human or mouse tissue samples. Moreover, since the probe pairs are designed to hybridize to a target nucleic acid, novel or unknown target nucleic acids or single-nucleotide polymorphisms (SNPs) are unlikely to be detected with such a templated ligation approach.

Thus, the present disclosure features methods, compositions, and kits for the spatial capture and/or processing of target nucleic acids from a biological sample by incorporating an analyte capture sequence into the target nucleic acids in the biological sample. Typically, the analyte capture sequence is a 5' pre-adenylated oligonucleotide and includes a 3' blocked end. In some embodiments, the biological sample on a substrate is aligned with an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes a capture domain and a spatial barcode.

Provided herein are methods for determining a location of a target nucleic acid in a biological sample, where the target nucleic acid is a non-polyadenylated target nucleic acid or a truncated polyadenylated target nucleic acid, the method including: (a) providing an array including a plurality of capture probes, where a capture probe of the plurality of capture

probes includes: (i) a spatial barcode and (ii) a capture domain; (b) incorporating an analyte capture sequence to a 3' end of the target nucleic acid in the biological sample, where the biological sample is disposed on a substrate, and where the analyte capture sequence includes a pre-adenylated 5' end and a blocked 3' end, thereby generating a product; (c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe; (d) extending the capture probe using the product as an extension template, thereby generating an extended capture probe; and (e) determining (i) the sequence of the spatial barcode or a complement thereof, and (ii) all or a portion of the sequence of the target nucleic acid or a complement thereof and using the determined sequences of (i) and (ii) to determine the location of the target nucleic acid in the biological sample.

Also provided herein are methods for processing a target nucleic acid in a biological sample, the method including: (a) providing an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) incorporating an analyte capture sequence to a 3' end of the target nucleic acid in the biological sample, where the biological sample is disposed on a substrate, and where the analyte capture sequence includes a pre-adenylated 5' end and a blocked 3' end, thereby generating a product; (c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe; and (d) extending the capture probe using the product as an extension template, thereby generating an extended capture probe.

Also provided herein are methods for enriching RNA from a pool of nucleic acids in a biological sample, where the RNA is a non-polyadenylated nucleic acid or a truncated polyadenylated nucleic acid, the method including: (a) providing an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) incorporating an analyte capture sequence to a 3' end of the RNA in the biological sample, where the biological sample is disposed on a substrate, and the analyte capture sequence includes a pre-adenylated 5' end and a blocked 3' end, thereby generating a product; and (c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe, thereby enriching the RNA from a pool of nucleic acids in the biological sample.

Also provided herein are methods for determining a location of a target nucleic acid in a biological sample, where the target nucleic acid includes a non-native poly(A) sequence acid, the method including: (a) providing an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) hybridizing the non-native poly(A) sequence of the target

nucleic acid to the capture domain of the capture probe; (c) extending the capture probe using the target nucleic as an extension template, thereby generating an extended capture probe; and (d) determining (i) the sequence of the spatial barcode or a complement thereof, and (ii) all or a portion of the sequence of the target nucleic acid or a complement thereof and using the determined sequences of (i) and (ii) to determine the location of the target nucleic acid in the biological sample.

In some embodiments, the non-native poly(A) sequence was previously incorporated at a 3' end of the target nucleic acid in the biological sample (*in situ*), wherein the biological sample is disposed on a substrate, and optionally wherein the non-native poly(A) sequence includes a blocked 3' end.

Non-limiting examples of target nucleic acids include nucleic acids such as DNA or RNA. Non-limiting examples of DNA analytes include genomic DNA, methylated DNA, specific methylated DNA sequences, fragmented DNA, mitochondrial DNA, chloroplast DNA, bacterial DNA, *in situ* synthesized PCR products, and viral DNA.

Non-limiting examples of RNA analytes include various types of coding and non-coding RNA. Examples of RNA analytes include non-polyadenylated RNA and/or truncated polyadenylated RNA. Examples of the different types of RNA analytes include messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), microRNA (miRNA), lincRNA, antisense RNA, and viral RNA. The RNA can be a transcript (e.g., present in a tissue section). The RNA can be small (e.g., less than 200 nucleic acid bases in length) or large (e.g., RNA greater than 200 nucleic acid bases in length). Small RNAs mainly include 5.8S ribosomal RNA (rRNA), 5S rRNA, transfer RNA (tRNA), microRNA (miRNA), small interfering RNA (siRNA), small nucleolar RNA (snoRNAs), Piwi-interacting RNA (piRNA), tRNA-derived small RNA (tsRNA), and small rDNA-derived RNA (srRNA). The RNA can be double-stranded RNA or single-stranded RNA. The RNA can be circular RNA. The RNA can be a bacterial rRNA (e.g., 16s rRNA or 23s rRNA). The RNA can be from an RNA virus, for example RNA viruses from Group III, IV or V of the Baltimore classification system. The RNA can be from a retrovirus, such as a virus from Group VI of the Baltimore classification system.

As used herein "non-polyadenylated" refers to target nucleic acids (e.g., DNA nucleic acids described herein or RNA nucleic acids described herein) that do not have a poly(A) tail at their 3' end. For example, it is generally understood that rRNAs do not possess a poly(A) tail at their 3' end; however, rRNA is capable of being polyadenylated. In contrast, it is well-known that mRNAs include a 5' cap that includes ribonucleotides that provide binding sites

for proteins that initiate polypeptide synthesis, a middle portion that encodes the polypeptide, and a "polyadenylated" portion that includes nucleotides at the 3' end (e.g., a poly(A) tail) that regulate the stability and processing of the mRNA. As used herein, "polyadenylated" refers to nucleic acids that include a polyadenylated sequence at its 3' end. The polyadenylated sequence can include about 100, 150, 200, 250, 300, 350, 400, 450, 500, or more nucleotides.

As used herein "truncated polyadenylated target nucleic acid" refers to target nucleic acids that are typically polyadenylated in a biological example (e.g., mRNA), but have a shortened poly(A) tail. For example, a typical mammalian mRNA has a length of about 250 adenosine nucleotides (see e.g., Kuhn, U. et al, Poly(A) Tail Length Is Controlled by the Nuclear Poly(A)-binding Protein Regulating the Interaction between Poly(A) Polymerase and the Cleavage and Polyadenylation Specificity Factor, *The Journal of Biological Chemistry*, 284:34, pp. 22803-22814 (2009)). However, in biological samples such as fixed biological samples the poly(A) tail of polyadenylated target nucleic acids can be damaged or truncated as a result of fixing, decrosslinking steps, and/or the age of the sample. Thus, a truncated polyadenylated target nucleic acid has a poly(A) tail shorter than its corresponding polyadenylated version. Thus, in some embodiments, a truncated polyadenylated target nucleic acid includes less than about 100, 150, 200, 250, 300, 350, 400, 450, 500, or more nucleotides. In some embodiments, a truncated polyadenylated target nucleic acid has a poly(A) tail of no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 adenosines. In some embodiments, a truncated polyadenylated target nucleic acid has a poly(A) tail of no more than about 10-20 (e.g., 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20) adenine containing nucleotides. As used herein "truncated polyadenylated target nucleic acid" can also refer to polyadenylated target nucleic acids that no longer have a poly(A) tail. However, the 3' end of the truncated polyadenylated target nucleic acid can be repaired e.g., a poly(A) added and made available for ligation.

As used herein "non-native poly(A) sequence" refers to a sequence of poly(A) nucleotides that is incorporated at an end (e.g., a 3' end) of a non-polyadenylated target nucleic acid or a truncated polyadenylated target nucleic acid. Typically, the sequence of poly(A) nucleotides is incorporated *in situ.* In preferred embodiments, the sequence of poly(A) nucleotides includes a homopolymeric sequence as described herein. In some embodiments, the non-native poly(A) sequence is about 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, 40, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, or more nucleotides long.

The biological sample can be applied to any of the variety of the substrates described herein. In some embodiments, the substrate is or includes the array. In some embodiments, the substrate does not include the array. In some embodiments, the substrate is a glass slide. In some embodiments, the method includes aligning the substrate containing the biological sample with the array, such that at least a portion of the biological sample is aligned with at least a portion of the array, optionally wherein the array is comprised on a second substrate (e.g., a sandwich format as described herein).

In some embodiments, the array includes one or more features. In some embodiments, features are directly or indirectly attached or fixed to a substrate. In some embodiments, the features are not directly or indirectly attached or fixed to a substrate, but instead, for example, are disposed within an enclosed or partially enclosed three-dimensional space (e.g., wells or divots). For example, the plurality of capture probes can be located on features on a substrate. In some embodiments, features include, but are not limited to, a spot, an inkjet spot, a masked spot, a pit, a post, a well, a ridge, a divot, a hydrogel pad, and a bead (e.g., a hydrogel bead).

In some embodiments, the plurality of capture probes include in a 5' to a 3' direction, a spatial barcode and a capture domain. In some embodiments, the capture domain hybridizes to the analyte capture sequence. In some embodiments, the capture domain is a poly(T) capture domain. In some embodiments, the capture probe includes a cleavage domain, one or more functional domains (e.g., a primer binding site, a sequencing site), a unique molecular identifier, or a combination thereof. A functional domain typically includes a functional nucleotide sequence for a downstream analytical step in the overall analysis procedure. In some embodiments, the functional domain can include a sequencing site. In some embodiments, the functional domain can include an amplification (e.g., PCR) site. In some embodiments, a capture probe can include a unique molecular identifier as described herein. In some embodiments, the unique molecular identifier is located 5' to the capture domain in the capture probe.

In some embodiments, incorporating the analyte capture sequence includes ligating the analyte capture sequence to the 3' end of the target nucleic acid. In some embodiments, the ligating includes the use of a ligase. In some embodiments, the ligase includes an RNA ligase. In some embodiments, the RNA ligase is a T4 RNA ligase, preferably T4 RNA ligase 2. In some embodiments, the ligase is thermostable 5' App DNA/RNA ligase. In some embodiments, the ligase is a combination of T4 RNA ligase 2 and thermostable 5' App DNA/RNA ligase.

In some embodiments, the analyte capture sequence includes a homopolymeric sequence. In some embodiments, the homopolymeric sequence is DNA. In some embodiments, the homopolymeric sequence is RNA. In some embodiments, the homopolymeric sequence is about 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, 40, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, or more nucleotides long. In some embodiments, the homopolymeric sequence is from about 20 nucleotides to about 75 nucleotides long. In some embodiments, the homopolymeric nucleotide sequence is from about 25 nucleotides to about 35 nucleotides long. In some embodiments, the homopolymeric nucleotide sequence is 30 nucleotides long. In some embodiments, the homopolymeric nucleotide sequence includes a poly(A) sequence.

In some embodiments, the method includes migrating the product from the biological sample to the array. In some embodiments, the migrating includes passive or active migration. In some embodiments, passive migration includes diffusion. In some embodiments, the migrating includes electrophoresis.

In some embodiments, the analyte capture sequence includes a 3' blocked end. The 3' blocked end prevents unwanted ligation between one or more analyte capture sequences, thus providing specificity in the ligation reaction. In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more carbon atoms (e.g., 2, 3, 4, 5, or more carbon atoms). In some embodiments, the 3' blocked end of the analyte capture sequence includes a biotin moiety. In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more inverted nucleotides. The combination of either T4 RNA ligase 2 and/or thermostable 5' DNARNA ligase with the 5' pre-adenylated poly(A) oligonucleotide also prevents unwanted ligation events since the ligase will only ligate the 5' pre-adenylated poly(A) oligonucleotides and not non-specifically ligate target nucleic acids in the biological sample. In some embodiments, the analyte capture sequence includes one or more phosphorothioate bonds (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10 or more phosphorothioate bonds).

In some embodiments, the method includes contacting the biological sample with one or more ribosomal RNA depletion probes. In some embodiments, the one or more ribosomal RNA depletion probes includes nucleic acid probes complementary to ribosomal RNA present in the biological sample and, optionally, a binding moiety. In some embodiments, the one or more ribosomal RNA depletion probes hybridize to the ribosomal RNA, thereby generating a ribosomal depletion probe/ribosomal RNA complex. In some embodiments, the binding moiety is biotin.

In some embodiments, the method includes contacting the biological sample with one or more mitochondrial RNA depletion probes. In some embodiments, the one or more mitochondrial RNA depletion probes includes nucleic acid probes complementary to mitochondrial RNA and, optionally, a binding moiety. In some embodiments, the one or more mitochondrial RNA depletion probes hybridize to the mitochondrial RNA, thereby generating a mitochondrial depletion probe/mitochondrial RNA complex. In some embodiments, the binding moiety is biotin.

In some embodiments, the ribosomal depletion probe/ribosomal RNA complexes and/or the mitochondrial depletion probe/mitochondrial RNA complexes are removed. In some embodiments, the removal includes the use of an RNase. In some embodiments, the RNase is RNase H1, RNase H2, or a thermostable RNase H. In some embodiments, the removal includes the use of streptavidin. For example, in embodiments where the depletion probe (e.g., rRNA or mtRNA depletion probes) includes a binding moiety such as biotin, the ribosomal depletion probe/ribosomal RNA complexes and/or the mitochondrial depletion probe/mitochondrial RNA complexes can be removed by contacting the biological sample or a pool of target nucleic acids derived from the biological sample with streptavidin, thereby preventing said complexes from interacting with capture probes on the array.

In some embodiments, the biological sample can be stained. In some embodiments, the biological sample is stained after fixation. In some embodiments, the biological sample is stained before fixation. In some embodiments, the staining includes optical labels as described herein, including, but not limited to, fluorescent (e.g., fluorophore), radioactive (e.g., radioisotope), chemiluminescent (e.g., a chemiluminescent compound), a bioluminescent compound, calorimetric, or colorimetric detectable labels. In some embodiments, the staining includes a fluorescent antibody directed to a target analyte (e.g., cell surface or intracellular proteins) in the biological sample. In some embodiments, the staining includes an immunohistochemistry stain directed to a target analyte (e.g., cell surface or intracellular proteins) in the fresh-frozen biological sample. In some embodiments, the staining includes a chemical stain, such as hematoxylin and eosin (H&E) or periodic acid-schiff (PAS). In some embodiments, staining the fresh-frozen biological sample includes the use of a biological stain including, but not limited to, acridine orange, Bismarck brown, carmine, coomassie blue, cresyl violet, DAPI, eosin, ethidium bromide, acid fuchsine, hematoxylin, Hoechst stains, iodine, methyl green, methylene blue, neutral red, Nile blue, Nile red, osmium tetroxide, propidium iodide, rhodamine, safranin, or any combination

thereof. In some embodiments, significant time (e.g., days, months, or years) can elapse between staining and/or imaging the biological sample.

In some embodiments, the biological sample is imaged. In some embodiments, the biological sample is imaged after fixation. In some embodiments, the biological sample is imaged before fixation. In some embodiments, imaging includes one or more of expansion microscopy, bright field microscopy, dark field microscopy, phase contrast microscopy, electron microscopy, fluorescence microscopy, reflection microscopy, interference microscopy and confocal microscopy.

In some embodiments, the biological sample is permeabilized. Permeabilization of a biological sample (e.g., a fixed biological sample) can occur on a substrate where the substrate is aligned with the array such that at least a portion of the biological sample is aligned with at least a portion of the array or directly on an array including a plurality of capture probes. In some embodiments, the biological sample is permeabilized with a protease. In some embodiments, the protease is one or more of pepsin, Proteinase K, and collagenase. In some embodiments, the biological sample is contacted with a DNase. In some embodiments, the DNase digests single-stranded nucleic acid molecules. In some embodiments, the DNase digests double-stranded DNA molecules. In some embodiments, the biological sample is contacted with both a single-stranded DNase and a double-stranded DNase.

The biological sample can be any of the biological samples described herein. For example, in some embodiments, the biological sample is a tissue sample. In some embodiments, the tissue sample is a fresh-frozen tissue sample. In some embodiments, the tissue sample is a fixed tissue sample. In some embodiments, the fixed tissue sample is a formalin-fixed paraffin-embedded (FFPE) tissue sample, an acetone-fixed tissue sample, a methanol-fixed tissue sample, or a paraformaldehyde-fixed tissue sample. In some embodiments, the biological sample is a tissue section. In some embodiments, tissue section is a fresh-frozen tissue section. In some embodiments, the tissue section is a fixed tissue section. In some embodiments, the fixed tissue section is a FFPE tissue section, an acetone-fixed tissue section, a methanol-fixed tissue section, or a paraformaldehyde-fixed tissue section.

In some embodiments, the fixed biological sample is an organoid, embryonic stem cells, pluripotent stem cells, or any combination thereof. Non-limiting examples of an organoid include a cerebral organoid, an intestinal organoid, a stomach organoid, a lingual organoid, a thyroid organoid, a thymic organoid, a testicular organoid, a hepatic organoid, a

pancreatic organoid, an epithelial organoid, a lung organoid, a kidney organoid, a gastruloid, a cardiac organoid, a retinal organoid, or any combination thereof. In other example embodiments, the fixed biological sample can include diseased cells, fetal cells, immune cells, cellular macromolecules, organelles, extracellular polynucleotides, or any combination thereof.

In some embodiments, the biological sample (e.g., tissue sample) is a tissue microarray (TMA). A tissue microarray contains multiple representative tissue samples – which can be from different tissues or organisms – assembled on a single histologic slide. The TMA can therefore allow for high throughput analysis of multiple specimens at the same time. Tissue microarrays are paraffin blocks produced by extracting cylindrical tissue cores from different paraffin donor blocks and re-embedding these into a single recipient (microarray) block at defined array coordinates.

In some embodiments, the 3' end of the capture probe can be extended to add a sequence that corresponds (e.g., complementary) to the target nucleic acid, the spatial barcode, the unique molecular identifier, one or more functional domains, and the complement of the analyte capture sequence. In some embodiments, a second strand can be generated (e.g., a second strand sequence complementary to the extended capture probe). In some embodiments, the extending includes the use of a reverse transcriptase. In some embodiments, the method includes the use of a template switch oligonucleotide. In some embodiments, generating the second strand includes adding at least 3 non-templated nucleotides during the extension step. For example, a homopolynucleotide sequence can be included (e.g., CCC) or a heteropolynucleotide sequence (e.g., CGG, CGC) can be included at the 3'end of the extended capture probe. In some embodiments, the method includes hybridizing the template switch oligonucleotide to the at least 3 non-templated nucleotides. The extended capture probe can be further extended using the template switch oligonucleotide as a template. In some embodiments, the method includes generating a second strand complementary to the extended capture probe (e.g., the extended capture probe containing a complement of the template switch oligonucleotide). For example, the extended capture probe containing a complement of the template switch oligonucleotide can be used as a template to generate a second strand using a primer identical or similar in sequence to the template switch oligonucleotide. In some embodiments, the second strand can include a sequence complementary to the spatial barcode and the target nucleic acid. In some embodiments, the second strand including the sequence of the target nucleic acids and the complementary sequences of the spatial barcode, UMI etc. are removed (e.g., melted away

with increased temperature or KOH) and the second strand can be, optionally amplified, and used to generate a sequencing library.

In some embodiments, the sequence of the spatial barcode or a complement thereof and the sequence of all or a portion of the target nucleic acid, or a complement thereof, can be determined. In some embodiments, the determined sequences can be used to identify the location of the target nucleic acid in the biological sample. The resulting products from the target nucleic acids (e.g., second strands) can be denatured from the extended capture probe and transferred (e.g., to a clean tube) for amplification, and/or library construction as described herein. The spatially-barcoded, full-length products or complements thereof can be amplified via PCR prior to library construction. The amplified products can then be enzymatically fragmented and size-selected in order to optimize the amplicon size. P5, P7, i7, and i5 can be incorporated into the library as for downstream sequencing (as described herein), and additional library sequencing regions, such as TruSeq Read 2, can be added via End Repair, A-tailing, Adaptor Ligation, and PCR. The fragments can then be sequenced using paired-end sequencing using TruSeq Read 1 and TruSeq Read 2 as sequencing primer sites. In some instances, the library is sequenced using any method described herein, such that different sequencing domains specific to other sequencing methods and techniques can be incorporated into a capture probe or introduced during library preparation. In some instances, the sequence of the target nucleic acid or a complement thereof is determined via sequencing. In some instances, the sequencing is high-throughput sequencing. In some instances, the spatial barcode is sequenced, providing the location of the target nucleic acid.

C. ***Kits***

The present disclosure also features kits used in any of the methods described herein. Thus, provided herein are kits including: (a) an array including a plurality of capture probes, where a capture probe of the plurality of capture probes includes: (i) a spatial barcode and (ii) a capture domain; (b) a plurality of analyte capture sequences, where an analyte capture sequence of the plurality of analyte capture sequences is capable of hybridizing to the capture domain of the capture probe and includes a pre-adenylated 5' end and a blocked 3' end; and (c) a ligase.

In some embodiments, the capture probe includes one or more functional domains (e.g., a primer binding site, a sequencing site), a cleavage domain, a unique molecular identifier, and combinations thereof.

In some embodiments, the analyte capture sequence includes a homopolymeric sequence. In some embodiments, the homopolymeric sequence is a DNA sequence. In some embodiments, the homopolymeric sequence includes a RNA sequence. In some embodiments, the homopolymeric sequence includes a poly(A) sequence.

In some embodiments, the ligase includes T4 RNA ligase 2. In some embodiments, the ligase is thermostable 5' App DNA/RNA ligase.

In some embodiments, the kit includes one or more permeabilization reagents. In some embodiments, the one or more permeabilization reagents includes a protease. In some embodiments, the protease includes proteinase K, pepsin, and/or collagenase. In some embodiments, the kit includes a DNase. In some embodiments, the DNase is a single-stranded DNase. In some embodiments, the DNase is a double-stranded DNase. In some embodiments, the kit includes both a single-stranded DNase and a double-stranded DNase.

In some embodiments, the kit includes one or more ribosomal RNA depletion probes (e.g., any of the ribosomal RNA depletion probes described herein). In some embodiments, the kit includes one or more mitochondrial RNA depletion probes (e.g., any of the mitochondrial RNA depletion probes described herein.

In some embodiments, the 3' blocked end of the analyte capture sequence is blocked. For example, the 3' end can be blocked by one or more carbon atoms (e.g., 2, 3, 4, 5, or more carbon atoms), a biotin moiety, and/or one or more inverted nucleotides. In some embodiments, the kit includes instructions for performing any of the methods described herein.

D. ***Compositions***

In addition to the methods and kits described herein, the present disclosure also features compositions including non-polyadenylated nucleic acids or truncated polyadenylated nucleic acids. For example, provided herein are composition including (a) a target nucleic acid, where the target nucleic acid is a non-polyadenylated nucleic acid or a truncated polyadenylated nucleic acid; (b) an analyte capture sequence comprising a 5' pre-adenylated end and a 3' blocked end, where the analyte capture sequence is capable of hybridizing to a capture domain of a capture probe; and (c) a ligase.

In some embodiments, the ligase is T4 RNA ligase 2. In some embodiments, the ligase is thermostable 5' App DNA/RNA ligase.

In some embodiments, the analyte capture sequence is covalently attached to a 3' end of the target nucleic acid.

In some embodiments, the analyte capture sequence includes a 3' blocked end. The 3' blocked end prevents unwanted ligation between one or more analyte capture sequences, thus providing specificity in the ligation reaction. In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more carbon atoms (e.g., 2, 3, 4, 5, or more carbon atoms). In some embodiments, the 3' blocked end of the analyte capture sequence includes a biotin moiety. In some embodiments, the 3' blocked end of the analyte capture sequence includes one or more inverted nucleotides.

In some embodiments, the analyte capture sequence comprises a homopolymeric sequence. In some embodiments, the homopolymeric sequence is a DNA sequence. In some embodiments, the homopolymeric sequence is a RNA sequence. In some embodiments, the homopolymeric sequence is from about 20 nucleotides to about 50 nucleotides. In some embodiments, the homopolymeric nucleotide sequence is from about 20 nucleotides to about 50 nucleotides. In some embodiments, the homopolymeric nucleotide sequence comprises about 30 nucleotides. In embodiments, where the homopolymeric sequence is a RNA sequence, the homopolymeric sequence includes ribonucleotides instead of nucleotides. In some embodiments, the homopolymeric sequence is a poly(A) sequence.

In some embodiments, the composition includes a DNase. In some embodiments, the DNase is a single-stranded DNase. In some embodiments, the DNase is a double-stranded DNase. In some embodiments, the composition includes both a single-stranded DNase and a double-stranded DNase.

In some embodiments, the composition includes one or more ribosomal RNA depletion probes (e.g., any of the ribosomal depletion probes described herein). In some embodiments, the composition includes one or more mitochondrial RNA depletion probes (e.g., any of the mitochondrial depletion probes described herein).

In some embodiments, the composition includes an array including a plurality of capture probes, wherein a capture probe of the plurality of capture probes includes a spatial barcode. In some embodiments, the capture probe includes a cleavage domain, one or more functional domains (e.g., a primer binding site, a sequencing site), a unique molecular identifier, and combinations thereof.

In some embodiments, the analyte capture sequence is hybridized to the capture domain of the capture probe.

In some embodiments, the capture probe has been extended using the analyte capture sequence covalently attached to the 3' end of the target nucleic acid as an extension template.

In some embodiments, the target nucleic acid is an RNA. In some embodiments, the RNA is mRNA, rRNA, tRNA, miRNA, viral RNA, siRNA, snoRNA, or piRNA.

EXAMPLES

Example 1. 5' Pre-Adenylated Poly(A) Oligonucleotide Ligation in Fixed Biological Samples

RNA molecules in fixed biological samples (e.g., FFPE samples) generally have degraded or missing poly(A) tails. Typical capture on a spatial array relies on the presence of a poly(A) sequence present in the RNA molecule to hybridize with a poly(T) capture domain.

In brief, experiments were performed according to the following protocol. Breast cancer FFPE samples were hematoxylin & eosin stained, decrosslinked in citrate buffer at 95°C, and treated with DNase. Next, 5' pre-adenylated 30mer oligonucleotides were ligated to RNA molecules present in the FFPE samples with T4 RNA ligase 2 (T4 Rnl2). T4 Rnl2 specifically ligates the pre-adenylated 5' end of either DNA or RNA oligonucleotides to the 3' end of RNA molecules. The T4 Rnl2 enzyme does not require ATP for ligation, however, it does require a pre-adenylated substrate for successful ligation which also confers specificity during the ligation reaction. Additionally, the pre-adenylated oligonucleotide included a block at its 3' end to prevent ligation of the pre-adenylated oligonucleotide to another pre-adenylated oligonucleotide. In this experiment, a three carbon atom spacer was used as a 3' blocker, however, other 3' blocking mechanisms are known in the art and described herein. Ligation was performed for 2 hours at 25°C in saline-sodium citrate (SSC) buffer and 25% polyethylene glycol (PEG) 8000 to improve ligation efficiency. Ligation was also performed in a sandwich format (as described herein), where the breast cancer tissue samples were disposed on separate substrates from the spatial arrays.

After ligation, the ligated products were captured on the spatial array, reverse transcribed to generate an extended capture probe which was further extended with a template switching oligonucleotide. A second strand complementary to the extended capture probe was generated by extension of the template switching oligonucleotide. The second strand was denatured from the extended capture probe with KOH, amplified, and a sequencing library was prepared.

FIG. 12A shows quantitative PCR results from a control reaction, where no 5' pre-adenylated oligonucleotide ligation was performed on a breast cancer FFPE tissue sample. The peak demonstrates an average length of about 550 base pairs and quantitation cycle (Cq) of 14.05. In contrast, **FIG. 12B** shows an increase in the number of shorter base pair species and a Cq of 10.51. The lower the Cq value means there were a higher initial copy numbers of the target nucleic acids, suggesting that ligating a pre-adenylated poly(A) oligonucleotide to RNA molecules in the breast cancer sample increased the overall population of RNA molecules which were then captured on the spatial array.

FIG. 13 further corroborates the data from **FIG. 12B**. **FIG. 13** shows ribosomal RNA and mitochondrial RNA sequencing data. The first column is a control reaction where typical templated ligation is performed with a pair of probes that hybridize to target nucleic acids. Generally, templated ligation is performed in fixed tissue samples for the reasons described herein. The pair of probes are ligated to one another and one of the probes includes an analyte capture sequence that hybridizes to the capture domain of the capture probe on the array. The middle column is another control reaction where 3' capture of RNA molecules is performed without the 5' pre-adenylated poly(A) oligonucleotide or T4 Rnl2 ligase. The data demonstrate that almost no ribosomal RNA was captured and some mitochondrial RNA was captured. The final column shows data from the 5' pre-adenylated poly(A) ligation described herein. The fraction of ribosomal RNA reads is significantly higher under the ligation reaction conditions than either the templated-ligation method or the control reaction without the ligase or pre-adenylated oligonucleotide. The second and third columns also show sequencing data from either GEX (gene expression) or VDJ sequencing library preparation methods. That is, the experiments were performed the same way for both methods and the difference between GEX and VDJ data points is only a result in the preparation of the sequencing libraries.

Overall, the data demonstrate successful capture of the ligation products generated by ligating a 5' pre-adenylated poly(A) oligonucleotide to rRNA molecules present in a FFPE biological sample.

What is Claimed is:

1. A method for determining a location of a target nucleic acid in a biological sample, wherein the target nucleic acid is a non-polyadenylated target nucleic acid or a truncated polyadenylated target nucleic acid, the method comprising:

(a) providing an array comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises: (i) a spatial barcode and (ii) a capture domain;

(b) incorporating an analyte capture sequence to a 3' end of the target nucleic acid in the biological sample, wherein the biological sample is disposed on a substrate, and wherein the analyte capture sequence comprises a pre-adenylated 5' end and a blocked 3' end, thereby generating a product;

(c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe;

(d) extending the capture probe using the product as an extension template, thereby generating an extended capture probe; and

(e) determining (i) the sequence of the spatial barcode or a complement thereof, and (ii) all or a portion of the sequence of the target nucleic acid or a complement thereof and using the determined sequences of (i) and (ii) to determine the location of the target nucleic acid in the biological sample.

2. The method of claim 1, wherein the target nucleic acid is an RNA.

3. The method of claim 2, wherein the RNA is an mRNA, rRNA, tRNA, miRNA, lincRNA, antisense RNA, viral RNA, siRNA, snoRNA, or piRNA.

4. The method of claim 1, wherein the target nucleic acid is DNA.

5. The method of claim 4, wherein the DNA is genomic DNA.

6. The method of any one of claims 1-5, wherein the substrate comprises the array.

7. The method of any one of claims 1-5, wherein the substrate is a glass slide.

8. The method of claim 7, further comprising aligning the substrate with the array, such that at least a portion of the biological sample is aligned with at least a portion of the array, optionally wherein the array is comprised on a second substrate.

9. The method of any one of claims 1-8, wherein incorporating the analyte capture sequence comprises ligating the analyte capture sequence to the 3' end of the target nucleic acid.

10. The method of claim 9, wherein the ligating comprises the use of a ligase.

11. The method of claim 10, wherein the ligase comprises an RNA ligase, preferably a T4 RNA ligase, more preferably T4 RNA ligase 2.

12. The method of any one of claims 1-11, wherein the extending comprises the use of a reverse transcriptase.

13. The method of any one of claims 1-12, further comprising the use of a template switch oligonucleotide.

14. The method of any one of claims 1-13, further comprising generating a second strand complementary to the extended capture probe.

15. The method of claim 14, wherein generating the second strand comprises incorporating at least 3 non-templated nucleotides during the extension in step (d).

16. The method of claim 15, further comprising hybridizing the template switch oligonucleotide to the at least 3 non-templated nucleotides and further extending the extended capture probe using the template switch oligonucleotide as a template.

17. The method of any one of claims 1-16, further comprising imaging the biological sample.

18. The method of any one of claims 1-17, further comprising staining the biological sample.

19. The method of claim 18, wherein the staining comprises hematoxylin and/or eosin staining.

20. The method of claim 18, wherein the staining comprises the use of a detectable label selected from the group consisting of a radioisotope, a fluorophore, a chemiluminescent compound, a bioluminescent compound, or a combination thereof.

21. The method of any one of claims 1-20, wherein the capture probe further comprises a cleavage domain, one or more functional domains, a unique molecular identifier, or a combination thereof.

22. The method of any one of claims 1-21, further comprising permeabilizing the biological sample.

23. The method of claim 22, wherein the permeabilizing comprises use of a protease.

24. The method of claim 23, wherein the protease comprises pepsin or proteinase K.

25. The method of any one of claims 1-24, wherein the array comprises one or more features selected from the group consisting of: a spot, an inkjet spot, a masked spot, a pit, a post, a well, a ridge, a divot, a hydrogel pad, and a bead.

26. The method of any one of claims 1-25, further comprising migrating the product from the biological sample to the array.

27. The method of claim 26, wherein the migrating comprises electrophoresis.

28. The method of any one of claims 1-27, wherein the biological sample is a tissue sample.

29. The method of claim 28, wherein the tissue sample is a fresh-frozen tissue sample.

30. The method of claim 28, wherein the tissue sample is a fixed tissue sample, and optionally, wherein the fixed tissue sample is a formalin-fixed paraffin-embedded tissue sample, an acetone-fixed tissue sample, a methanol-fixed tissue sample, or a paraformaldehyde-fixed tissue sample.

31. The method of any one of claims 1-27, wherein the biological sample is a tissue section.

32. The method of claim 31, wherein the tissue section is a fresh-frozen tissue section.

33. The method of claim 31, wherein the tissue section is a fixed tissue section.

34. The method of claim 33, wherein the fixed tissue section is a formalin-fixed paraffin-embedded tissue section, an acetone-fixed tissue section, a methanol-fixed tissue section, or a paraformaldehyde-fixed tissue section.

35. The method of any one of claims 1-34, further comprising contacting the biological sample with a DNase.

36. The method of any one of claims 1-35, wherein the 3' blocked end of the analyte capture sequence comprises one or more carbon atoms.

37. The method of any one of claims 1-35, wherein the 3' blocked end of the analyte capture sequence comprises a biotin moiety.

38. The method of any one of claims 1-35, wherein the 3' blocked end of the analyte capture sequence comprises one or more inverted nucleotides.

39. The method of any one of claims 1-38, wherein the analyte capture sequence comprises a homopolymeric nucleotide sequence comprising DNA or RNA.

40. The method of claim 39, wherein the homopolymeric nucleotide sequence comprises from about 20 nucleotides to about 50 nucleotides.

41. The method of claim 40, wherein the homopolymeric nucleotide sequence comprises from about 25 nucleotides to about 35 nucleotides.

42. The method of any one of claims 39-41, wherein the homopolymeric nucleotide sequence comprises a poly(A) sequence.

43. The method of any one of claims 39-42, wherein the capture domain of the capture probe comprises a poly(T) sequence.

44. The method of any one of claims 1-43, further comprising contacting the biological sample with one or more ribosomal RNA depletion probes.

45. The method of claim 44, wherein the one or more ribosomal RNA depletion probes comprises nucleic acid probes complementary to ribosomal RNA and a binding moiety.

46. The method of claim 45, wherein the one or more ribosomal RNA depletion probes hybridize to the ribosomal RNA, thereby generating a ribosomal depletion probe/ribosomal RNA complex.

47. The method of claim 45 or 46, wherein the binding moiety is biotin.

48. The method of any one of claims 1-47, further comprising contacting the biological sample with one or more mitochondrial RNA depletion probes.

49. The method of claim 48, wherein the one or more mitochondrial RNA depletion probes comprises nucleic acid probes complementary to mitochondrial RNA and a binding moiety.

50. The method of claim 49, wherein the one or more mitochondrial RNA depletion probes hybridize to the mitochondrial RNA, thereby generating a mitochondrial depletion probe/mitochondrial RNA complex.

51. The method of claim 49, wherein the binding moiety is biotin.

52. The method of claim 46 or 50, wherein the ribosomal depletion probe/ribosomal RNA complex and/or the mitochondrial depletion probe/mitochondrial RNA complex are removed.

53. The method of claim 52, wherein the removal comprises the use of an RNase.

54. The method of claim 53, wherein the RNase is RNase H1, RNase H2, or a thermostable RNase H.

55. The method of claim 52, wherein the removal comprises the use of streptavidin.

56. The method of any one of claims 1-55, further comprising generating a polynucleotide comprising: (i) the spatial barcode or a complement thereof; (ii) the analyte capture sequence or a complement thereof; and (iii) all or a portion of the sequence of the target nucleic acid or a complement thereof.

57. The method of claim 56, wherein the polynucleotide is generated by amplification of the extended capture probe or a complement thereof.

58. The method of claim 57, wherein the complement of the extended capture probe is removed from the array.

59. The method of claim 58, wherein the removing comprises use of KOH.

60. The method of claim 58, wherein the removing comprises use of heat.

61. A kit comprising:
(a) an array comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises: (i) a spatial barcode and (ii) a capture domain;
(b) a plurality of analyte capture sequences, wherein an analyte capture sequence of the plurality of analyte capture sequences is capable of hybridizing to the capture domain and comprises a pre-adenylated 5' end and a blocked 3' end; and
(c) a ligase.

62. The kit of claim 61, wherein the capture probe further comprises one or more functional domains, a cleavage domain, a unique molecular identifier, or a combination thereof.

63. The kit of claim 61 or 62, wherein the analyte capture sequence comprises a homopolymeric nucleotide sequence.

64. The kit of claim 63, wherein the homopolymeric nucleotide sequence comprises a poly(A) sequence.

65. The kit of any one of claims 61-64, wherein the ligase comprises an RNA ligase, preferably a T4 RNA ligase, more preferably T4 RNA ligase 2.

66. The kit of any one of claims 61-65, further comprising one or more permeabilization reagents.

67. The kit of claim 66, wherein the one or more permeabilization reagents comprises a protease, and optionally, wherein the protease comprises proteinase K, pepsin, or collagenase.

68. The kit of any one of claims 61-67, further comprising a DNase.

69. The kit of any one of claims 61-68, further comprising one or more ribosomal RNA depletion probes.

70. The kit of any one of claims 61-69, further comprising one or more mitochondrial RNA depletion probes.

71. The kit of any one of claims 61-70, further comprising instructions for performing the method of any one of claims 1-60.

72. The kit of any one of claims 61-71, wherein the 3' blocked end of the analyte capture sequence comprises one or more carbon atoms, a biotin moiety, and/or one or more inverted nucleotides.

73. A method for enriching RNA from a pool of nucleic acids in a biological sample, wherein the RNA is a non-polyadenylated nucleic acid or a truncated polyadenylated nucleic acid, the method comprising:

(a) providing an array comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises: (i) a spatial barcode and (ii) a capture domain;

(b) incorporating an analyte capture sequence to a 3' end of the RNA in the biological sample, wherein the biological sample is disposed on a substrate, and the analyte capture sequence comprises a pre-adenylated 5' end and a blocked 3' end, thereby generating a product; and

(c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe, thereby enriching the RNA from a pool of nucleic acids in the biological sample.

74. The method of claim 73, wherein the RNA is an mRNA, rRNA, tRNA, miRNA, lincRNA, antisense RNA, viral RNA, siRNA, snoRNA, or piRNA.

75. The method of claim 73 or 74, wherein the substrate comprises the array.

76. The method of any one of claims 73-75, wherein the substrate is a glass slide.

77. The method of claim 76, further comprising aligning the substrate with the array, such that at least a portion of the biological sample is aligned with at least a portion of the array, optionally wherein the array is comprised in a second substrate.

78. The method of any one of claims 73-77, wherein incorporating the analyte capture sequence comprises ligating the analyte capture sequence to the 3' end of the RNA.

79. The method of claim 78, wherein the ligating comprises the use of a ligase.

80. The method of claim 79, wherein the ligase comprises an RNA ligase, preferably a T4 RNA ligase, more preferably T4 RNA ligase 2.

81. The method of any one of claims 73-80, further comprising extending the capture probe using the product as an extension template, thereby generating an extended capture probe.

82. The method of claim 81, wherein the extending comprises the use of a reverse transcriptase.

83. The method of any one of claims 73-82, further comprising the use of a template switch oligonucleotide.

84. The method of any one of claims 81-83, further comprising generating a second strand complementary to the extended capture probe.

85. The method of claim 84, wherein generating the second strand comprises adding at least 3 non-templated nucleotides during the extension.

86. The method of claim 85, further comprising hybridizing the template switch oligonucleotide to the at least 3 non-templated nucleotides and further extending the extended capture probe using the capture probe as a template.

87. The method of any one of claims 73-86, further comprising imaging the biological sample.

88. The method of any one of claims 73-87, further comprising staining the biological sample.

89. The method of claim 88, wherein the staining comprises hematoxylin and/or eosin staining.

90. The method of claim 88, wherein the staining comprises the use of a detectable label selected from the group consisting of a radioisotope, a fluorophore, a chemiluminescent compound, a bioluminescent compound, or a combination thereof.

91. The method of any one of claims 73-90, wherein the capture probe further comprises a cleavage domain, one or more functional domains, a unique molecular identifier, or a combination thereof.

92. The method of any one of claims 73-91, further comprising permeabilizing the biological sample.

93. The method of claim 92, wherein the permeabilizing comprises the use of a protease.

94. The method of claim 93, wherein the protease comprises pepsin, collagenase, or proteinase K.

95. The method of any one of claims 73-94, wherein the array comprises one or more features selected from the group consisting of: a spot, an inkjet spot, a masked spot, a pit, a post, a well, a ridge, a divot, a hydrogel pad, and a bead.

96. The method of any one of claims 73-95, further comprising migrating the product from the biological sample to the array.

97. The method of claim 96, wherein the migrating comprises electrophoresis.

98. The method of any one of claims 73-97, wherein the biological sample is a tissue sample.

99. The method of claim 98, wherein the tissue sample is a fresh-frozen tissue sample.

100. The method of claim 98, wherein the tissue sample is a fixed tissue sample, and optionally, wherein the fixed tissue sample is a formalin-fixed paraffin-embedded tissue

sample, an acetone-fixed tissue sample, a methanol-fixed tissue sample, or a paraformaldehyde-fixed tissue sample.

101. The method of any one of claims 73-97, wherein the biological sample is a tissue section.

102. The method of claim 101, wherein the tissue section is a fresh-frozen tissue section.

103. The method of claim 101, wherein the tissue section is a fixed tissue section.

104. The method of claim 103, wherein the fixed tissue section is a formalin-fixed paraffin-embedded tissue section, an acetone-fixed tissue section, a methanol-fixed tissue section, or a paraformaldehyde-fixed tissue section.

105. The method of any one of claims 73-104, further comprising contacting the biological sample with a DNase.

106. The method of any one of claims 73-105, wherein the 3' blocked end of the analyte capture sequence comprises one or more carbon atoms.

107. The method of any one of claims 73-105, wherein the 3' blocked end of the analyte capture sequence comprises a biotin moiety.

108. The method of any one of claims 73-105, wherein the 3' blocked end of the analyte capture sequence comprises one or more inverted nucleotides.

109. The method of any one of claims 73-108, wherein the analyte capture sequence comprises a homopolymeric nucleotide sequence comprising DNA or RNA.

110. The method of claim 109, wherein the homopolymeric nucleotide sequence comprises from about 20 nucleotides to about 50 nucleotides.

111. The method of claim 110, wherein the homopolymeric nucleotide sequence comprises from about 25 nucleotides to about 35 nucleotides.

112. The method of any one of claims 109-111, wherein the homopolymeric nucleotide sequence comprises a poly(A) sequence.

113. The method of any one of claims 73-112, further comprising contacting the biological sample with one or more ribosomal RNA depletion probes.

114. The method of claim 113, wherein the one or more ribosomal RNA depletion probes comprises nucleic acid probes complementary to ribosomal RNA and a binding moiety.

115. The method of claim 114, wherein the one or more ribosomal RNA depletion probes hybridize to the ribosomal RNA, thereby generating a ribosomal depletion probe/ribosomal RNA complex.

116. The method of claim 114 or 115, wherein the binding moiety is biotin.

117. The method of any one of claims 73-116, further comprising contacting the biological sample with one or more mitochondrial RNA depletion probes.

118. The method of claim 117, wherein the one or more mitochondrial RNA depletion probes comprises nucleic acid probes complementary to mitochondrial RNA and a binding moiety.

119. The method of claim 118, wherein the one or more mitochondrial RNA depletion probes hybridize to the mitochondrial RNA, thereby generating a mitochondrial depletion probe/mitochondrial RNA complex.

120. The method of claim 118, wherein the binding moiety is biotin.

121. The method of claims 115 or 119, wherein the ribosomal depletion probe/ribosomal RNA complex and/or the mitochondrial depletion probe/mitochondrial RNA complex are removed.

122. The method of claim 121, wherein the removal comprises the use of an RNase.

123. The method of claim 122, wherein the RNase is RNase H1, RNase H2, or a thermostable RNase H.

124. The method of claim 122, wherein the removal comprises the use of streptavidin.

125. A composition comprising:

(a) a target nucleic acid, wherein the target nucleic acid is a non-polyadenylated nucleic acid or a truncated polyadenylated nucleic acid;

(b) an analyte capture sequence comprising a 5' pre-adenylated end and a 3' blocked end, where the analyte capture sequence is capable of hybridizing to a capture domain of a capture probe; and

(c) a ligase.

126. The composition of claim 125, wherein the ligase comprises RNA ligase 2.

127. The composition of claim 125 or 126, wherein the analyte capture sequence is covalently attached to a 3' end of the target nucleic acid.

128. The composition of any one of claims 125-127, wherein the 3' blocked end of the analyte capture sequence comprises one or more carbon atoms.

129. The composition of any one of claims 125-127, wherein the 3' blocked end of the analyte capture sequence comprises a biotin moiety.

130. The composition of any one of claims 125-127, wherein the 3' blocked end of the analyte capture sequence comprises one or more inverted nucleotides.

131. The composition of any one of claims 125-130, wherein the analyte capture sequence comprises a homopolymeric nucleotide sequence comprising DNA or RNA.

132. The composition of claim 131, wherein the homopolymeric nucleotide sequence comprises from about 20 nucleotides to about 50 nucleotides.

133. The composition of claim 132, wherein the homopolymeric nucleotide sequence comprises about 30 nucleotides.

134. The composition of any one of claims 131-133, wherein the homopolymeric nucleotide sequence comprises a poly(A) sequence.

135. The composition of any one of claims 125-134, further comprising a DNase.

136. The composition of any one of claims 125-135, further comprising one or more ribosomal RNA depletion probes.

137. The composition of any one of claims 125-136, further comprising one or more mitochondrial RNA depletion probes.

138. The composition of any one of claims 125-137, further comprising an array comprising the capture probe, wherein the capture probe further comprises a spatial barcode.

139. The composition of claim 138, wherein the capture probe further comprises a cleavage domain, one or more functional domains, a unique molecular identifier, and combinations thereof.

140. The composition of claim 125-139, wherein the analyte capture sequence is hybridized to the capture domain of the capture probe.

141. The composition of claim 140, wherein the capture probe has been extended using the analyte capture sequence covalently attached to the 3' end of the target nucleic acid as an extension template.

142. The composition of any one of claims 125-141, wherein the target nucleic acid is an RNA.

143. The composition of claim 142, wherein the RNA is mRNA, rRNA, tRNA, miRNA, viral RNA, siRNA, snoRNA, or piRNA.

144. A method for processing a target nucleic acid in a biological sample, the method comprising:

(a) providing an array comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises: (i) a spatial barcode and (ii) a capture domain;

(b) incorporating an analyte capture sequence to a 3' end of the target nucleic acid in the biological sample, wherein the biological sample is disposed on a substrate, and wherein the analyte capture sequence comprises a pre-adenylated 5' end and a blocked 3' end, thereby generating a product;

(c) hybridizing the analyte capture sequence of the product to the capture domain of the capture probe; and

(d) extending the capture probe using the product as an extension template, thereby generating an extended capture probe.

145. The method of claim 144, wherein the target nucleic acid is a non-polyadenylated target nucleic acid or a truncated polyadenylated target nucleic acid.

146. The method of claim 144 or 145, wherein the substrate comprises the array.

147. The method of any one of claims 144-146, wherein the substrate is a glass slide.

148. The method of claim 147, further comprising aligning the substrate with the array, such that at least a portion of the biological sample is aligned with at least a portion of the array, optionally wherein the array is comprised in a second substrate.

149. The method of any one of claims 144-148, wherein incorporating the analyte capture sequence comprises ligating the analyte capture sequence to the 3' end of the target nucleic acid.

150. The method of claim 149, wherein the ligating comprises the use of a ligase.

151. The method of claim 150, wherein the ligase comprises an RNA ligase, preferably a T4 RNA ligase, more preferably T4 RNA ligase 2.

152. The method of any one of claims 144-151, wherein the extending comprises the use of a reverse transcriptase.

153. The method of any one of claims 144-152, further comprising the use of a template switch oligonucleotide.

154. The method of any one of claims 144-153, further comprising generating a second strand complementary to the extended capture probe.

155. The method of claim 154, wherein generating the second strand comprises adding at least 3 non-templated nucleotides during the extension in step (d).

156. The method of claim 155, further comprising hybridizing the template switch oligonucleotide to the at least 3 non-templated nucleotides and further extending the extended

capture probe using the extended capture probe as a template.

157. The method of any one of claims 144-156, wherein the analyte capture sequence comprises a homopolymeric nucleotide sequence.

158. The method of claim 157, wherein the homopolymeric nucleotide sequence comprises a poly(A) sequence, optionally from about 20 nucleotides to about 50 nucleotides.

159. A method for determining a location of a target nucleic acid in a biological sample, wherein the target nucleic acid comprises a non-native poly(A) sequence acid, the method comprising:

(a) providing an array comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises: (i) a spatial barcode and (ii) a capture domain;

(b) hybridizing the non-native poly(A) sequence of the target nucleic acid to the capture domain of the capture probe;

(c) extending the capture probe using the target nucleic as an extension template, thereby generating an extended capture probe; and

(d) determining (i) the sequence of the spatial barcode or a complement thereof, and (ii) all or a portion of the sequence of the target nucleic acid or a complement thereof and using the determined sequences of (i) and (ii) to determine the location of the target nucleic acid in the biological sample.

160. The method of claim 159, wherein the non-native poly(A) sequence was previously incorporated at a 3' end of the target nucleic acid in the biological sample (*in situ*), wherein the biological sample is disposed on a substrate, and optionally wherein the non-native poly(A) sequence comprises a blocked 3' end.

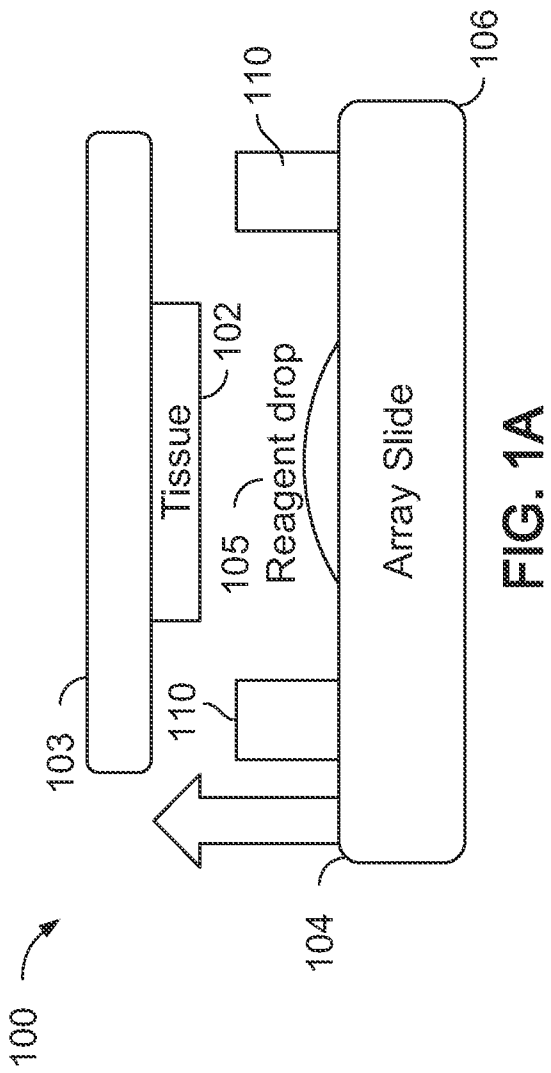

100
103
110
Tissue
102
105
Reagent drop
110
Array Slide
106
104

FIG. 1A

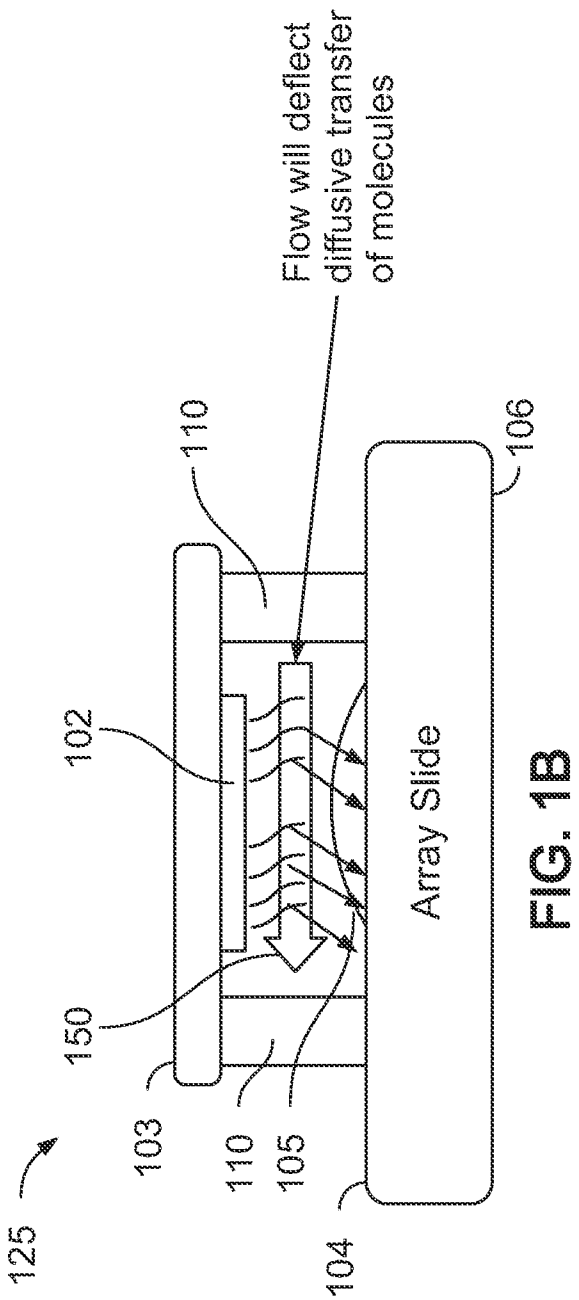

Flow will deflect diffusive transfer of molecules
125
150
102
103
110
105
110
Array Slide
106
104

FIG. 1B

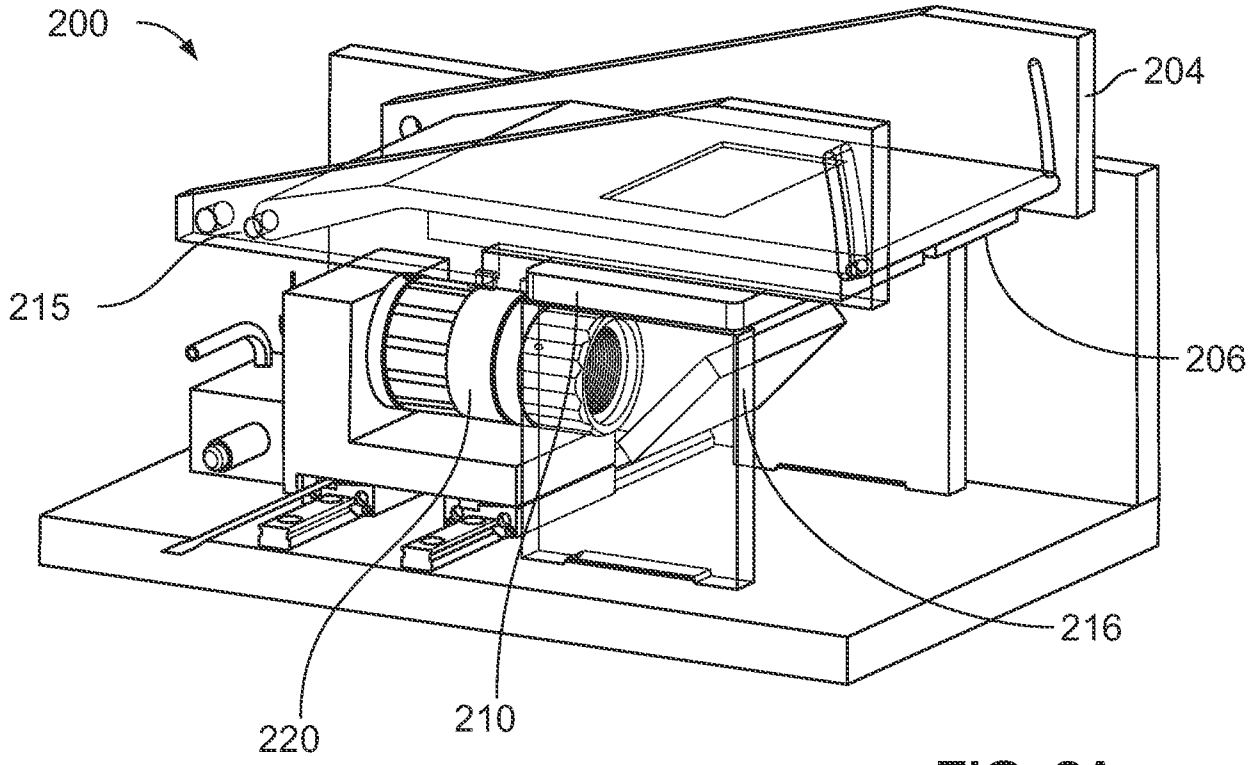

200
204
215
206
216
220
210

FIG. 2A

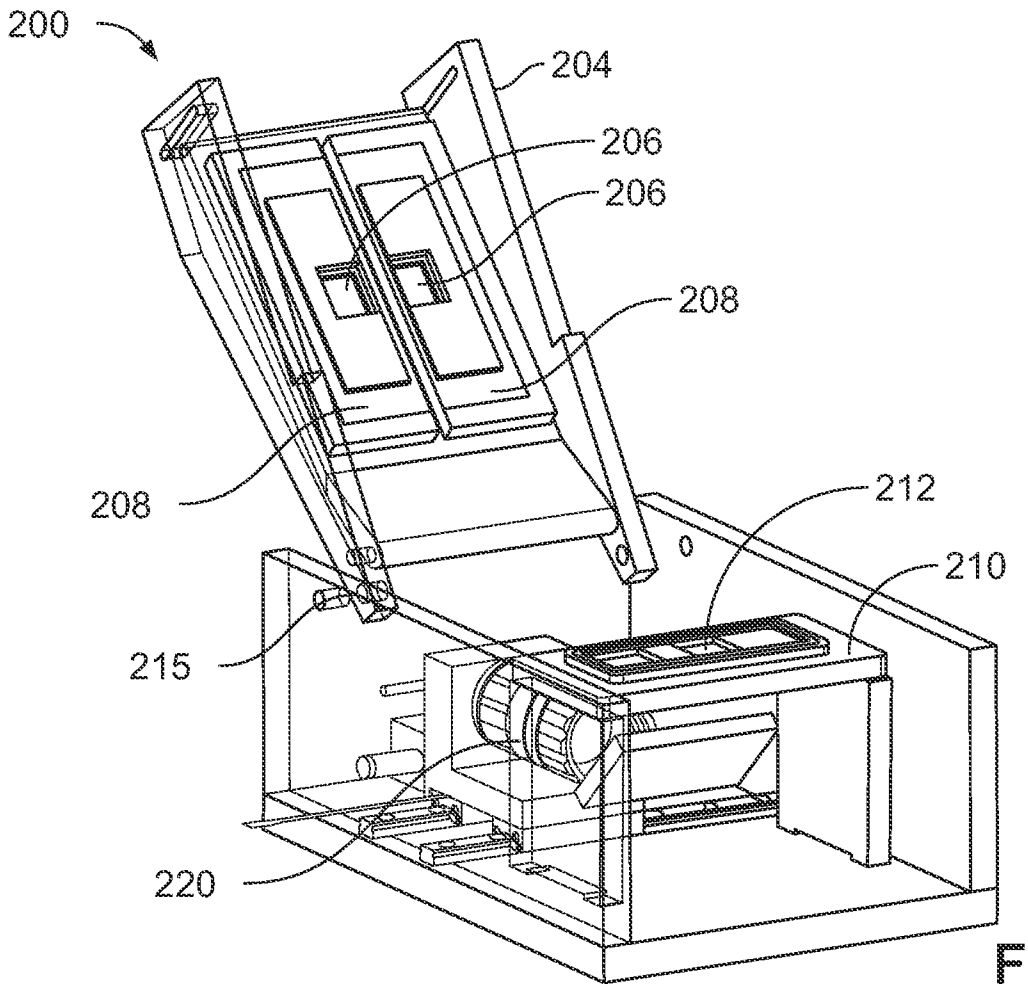

200
204
206
206
208
208
212
210
215
220

FIG. 2B

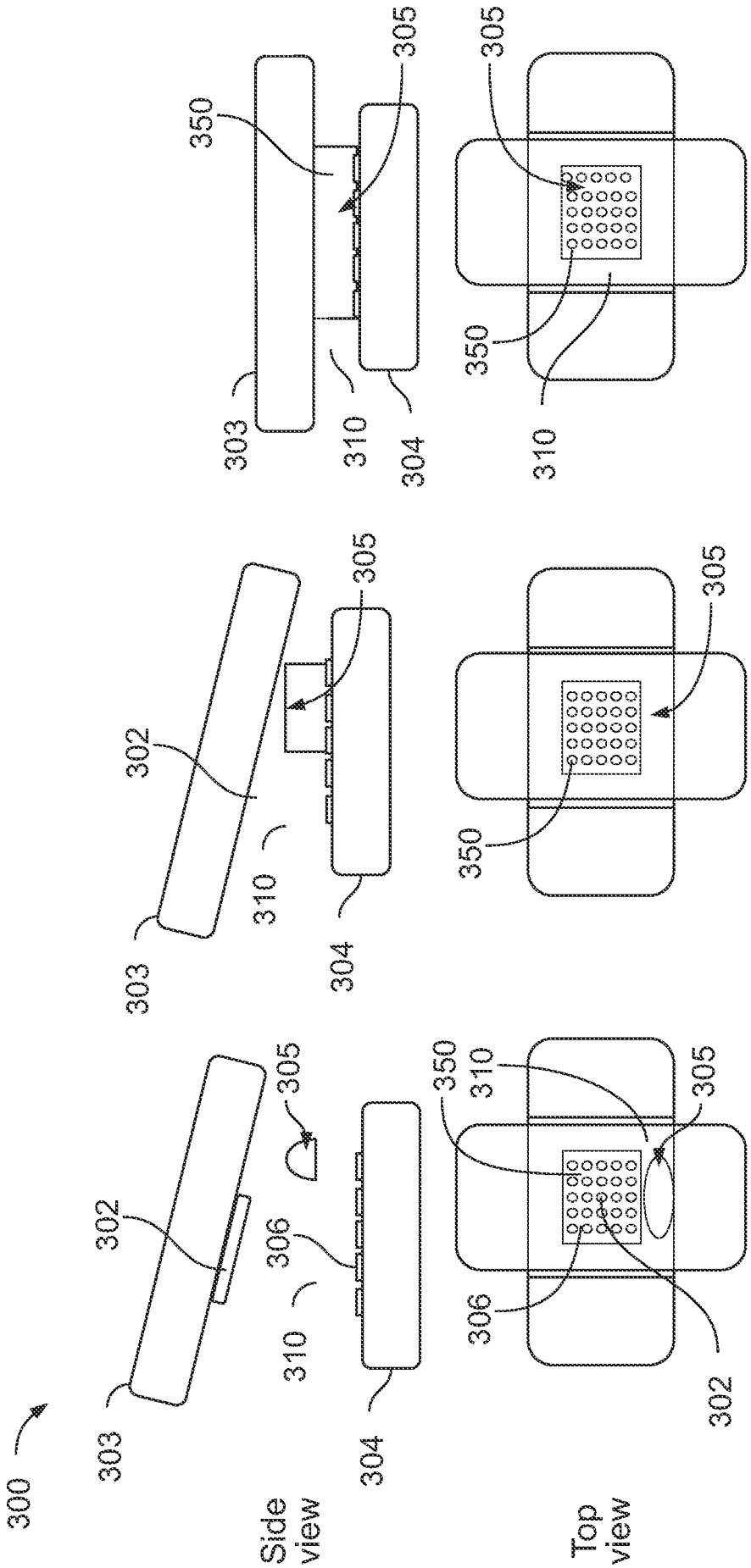


FIG. 3A

FIG. 3B

FIG. 3C

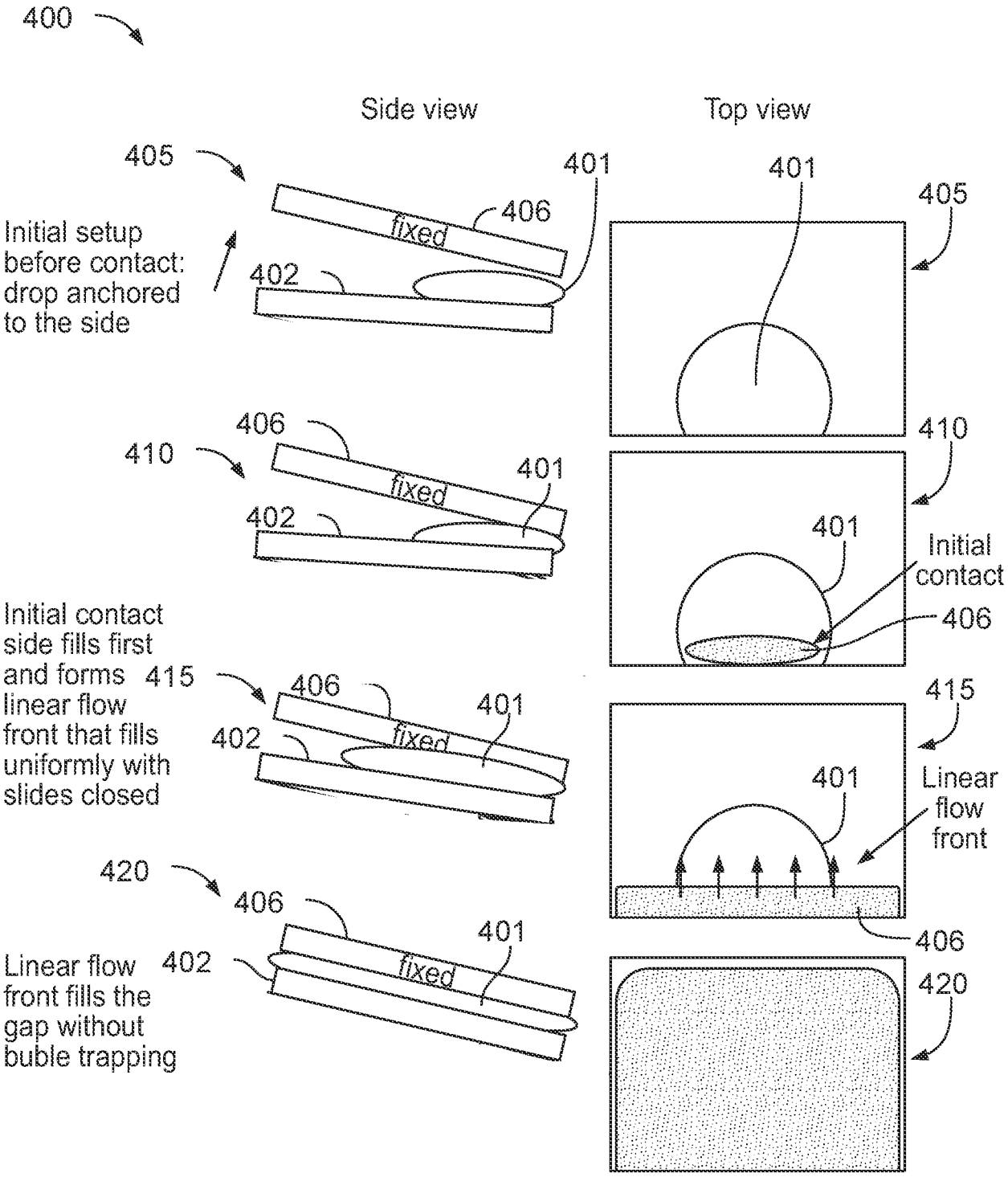

400
Side view
Top view
405
401
406
fixed
402
Initial setup before contact: drop anchored to the side
410
Initial contact side fills first and forms linear flow front that fills uniformly with slides closed
415
420
Linear flow front fills the gap without buble trapping
Initial contact
Linear flow front

FIG. 4A FIG. 4B

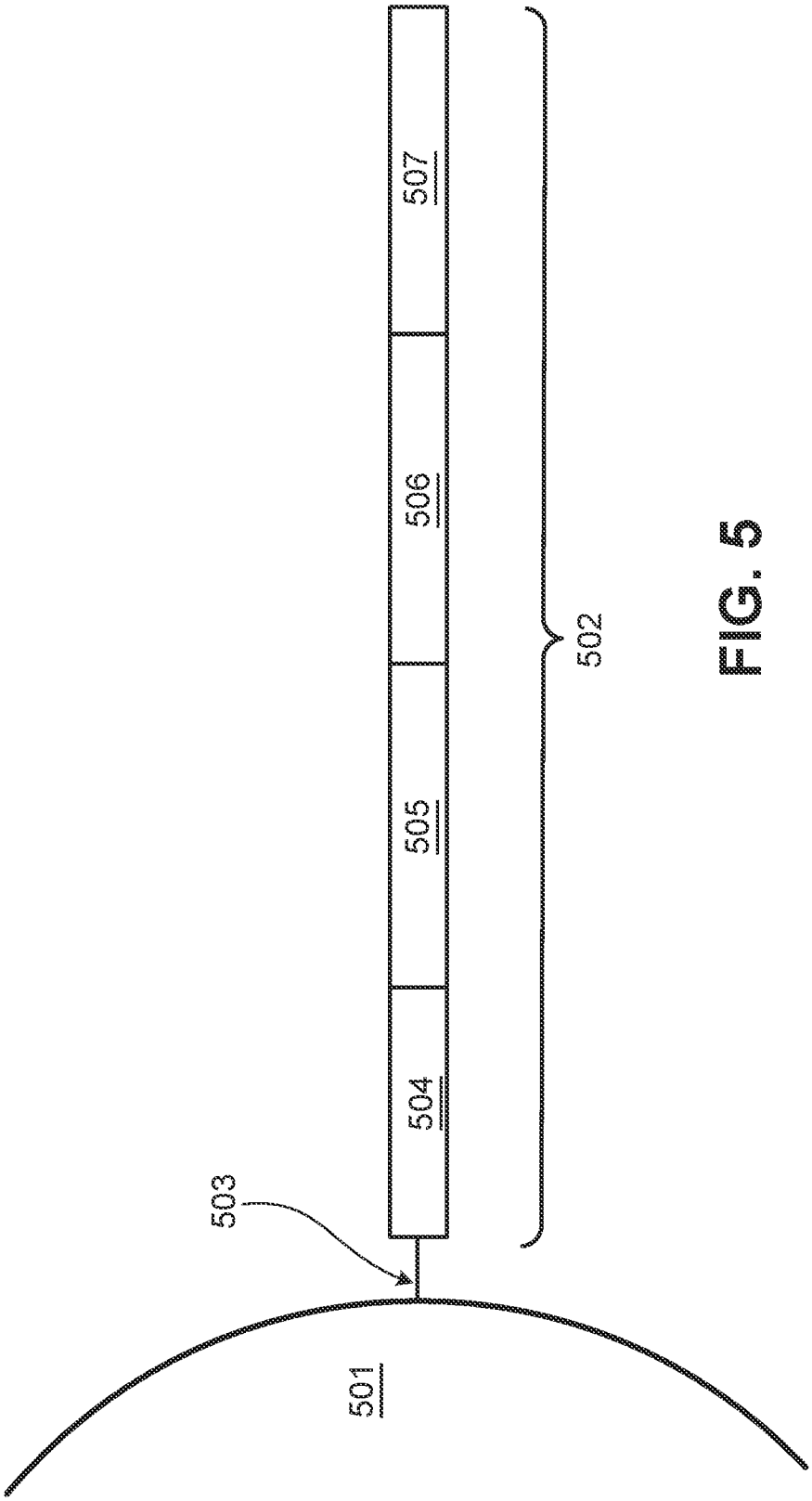


FIG. 5

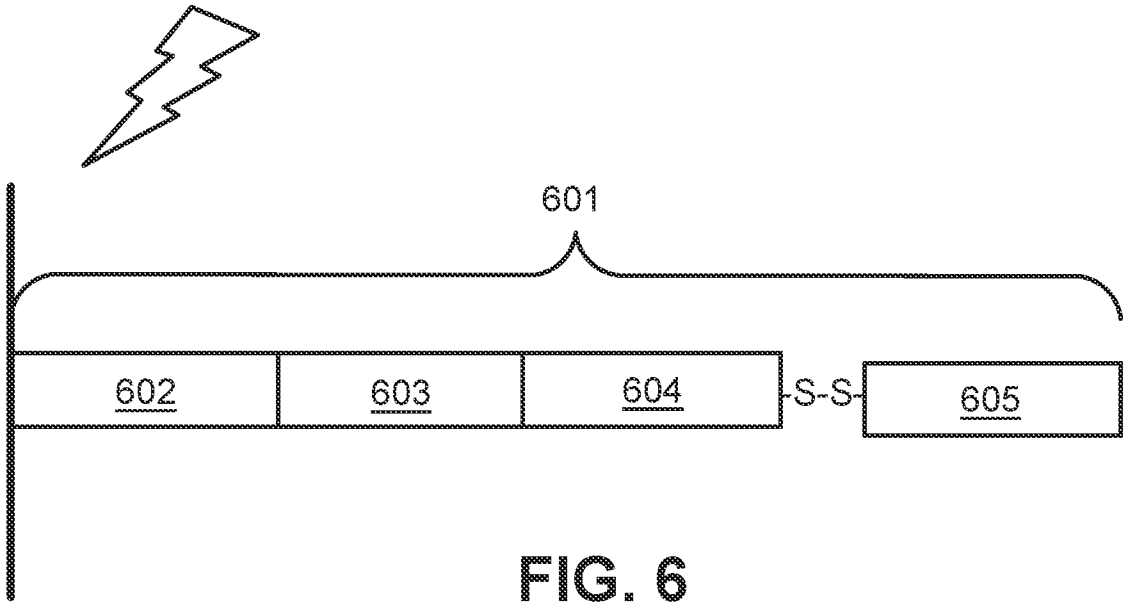


FIG. 6

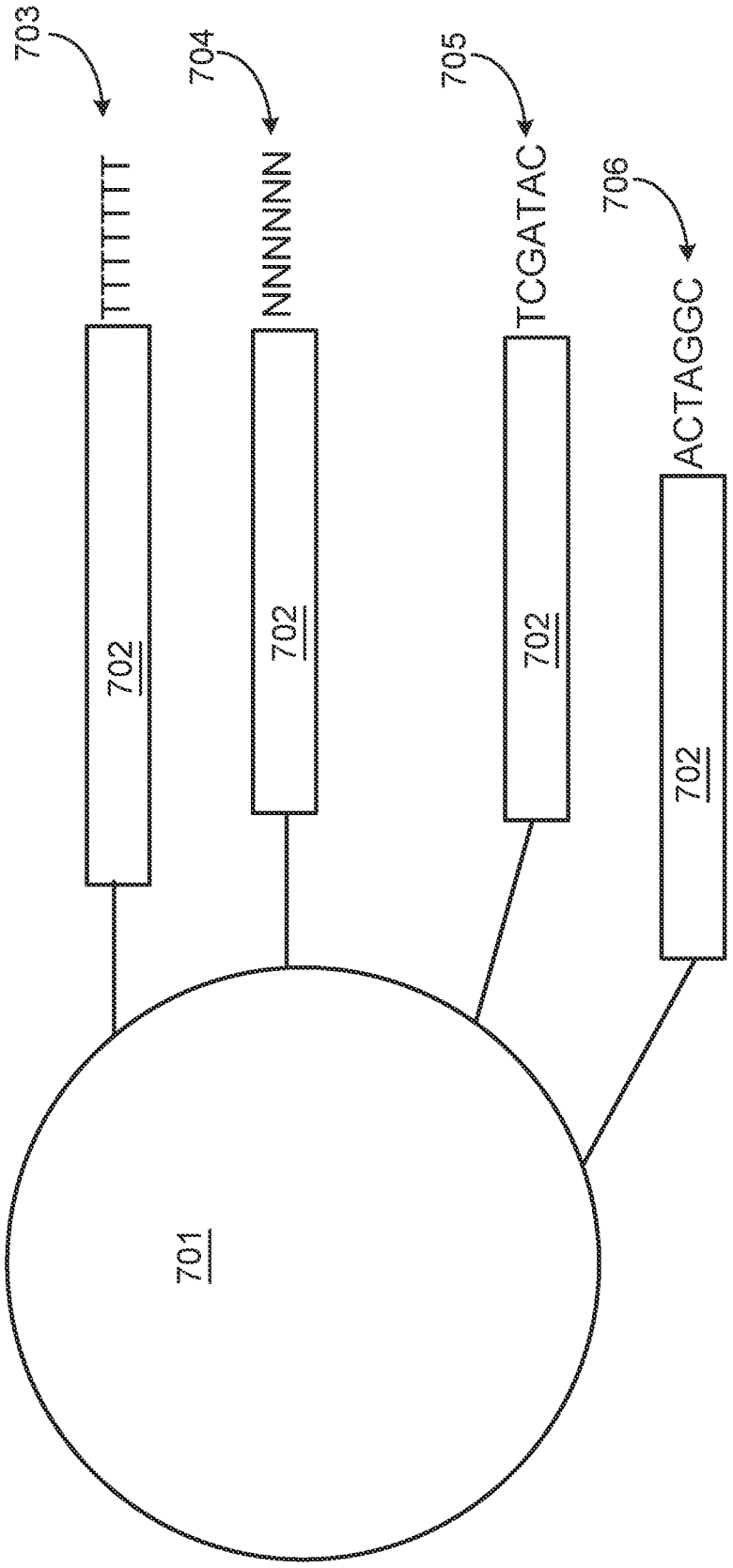

703
704
705
706
TTTTTTT
NNNNNN
TCGATAC
ACTAGGC
702
702
702
702
701

FIG. 7

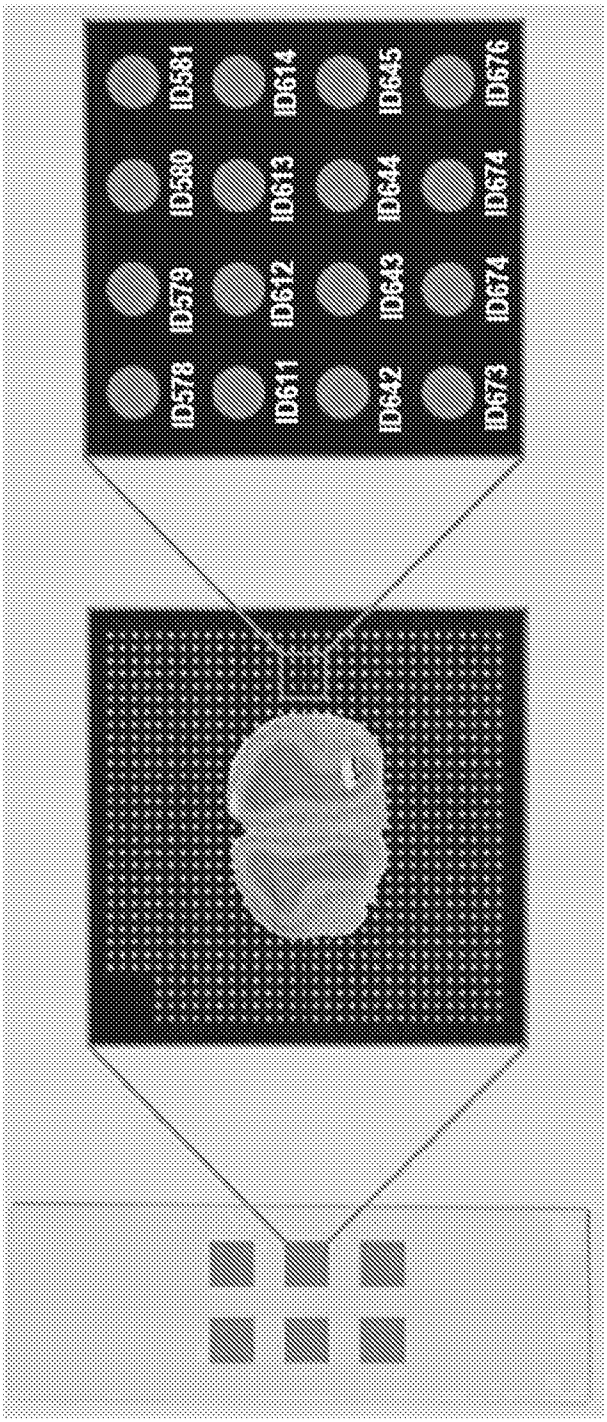

ID631
ID614
ID645
ID676
ID630
ID613
ID644
ID674
ID629
ID612
ID643
ID674
ID628
ID611
ID642
ID673

FIG. 8

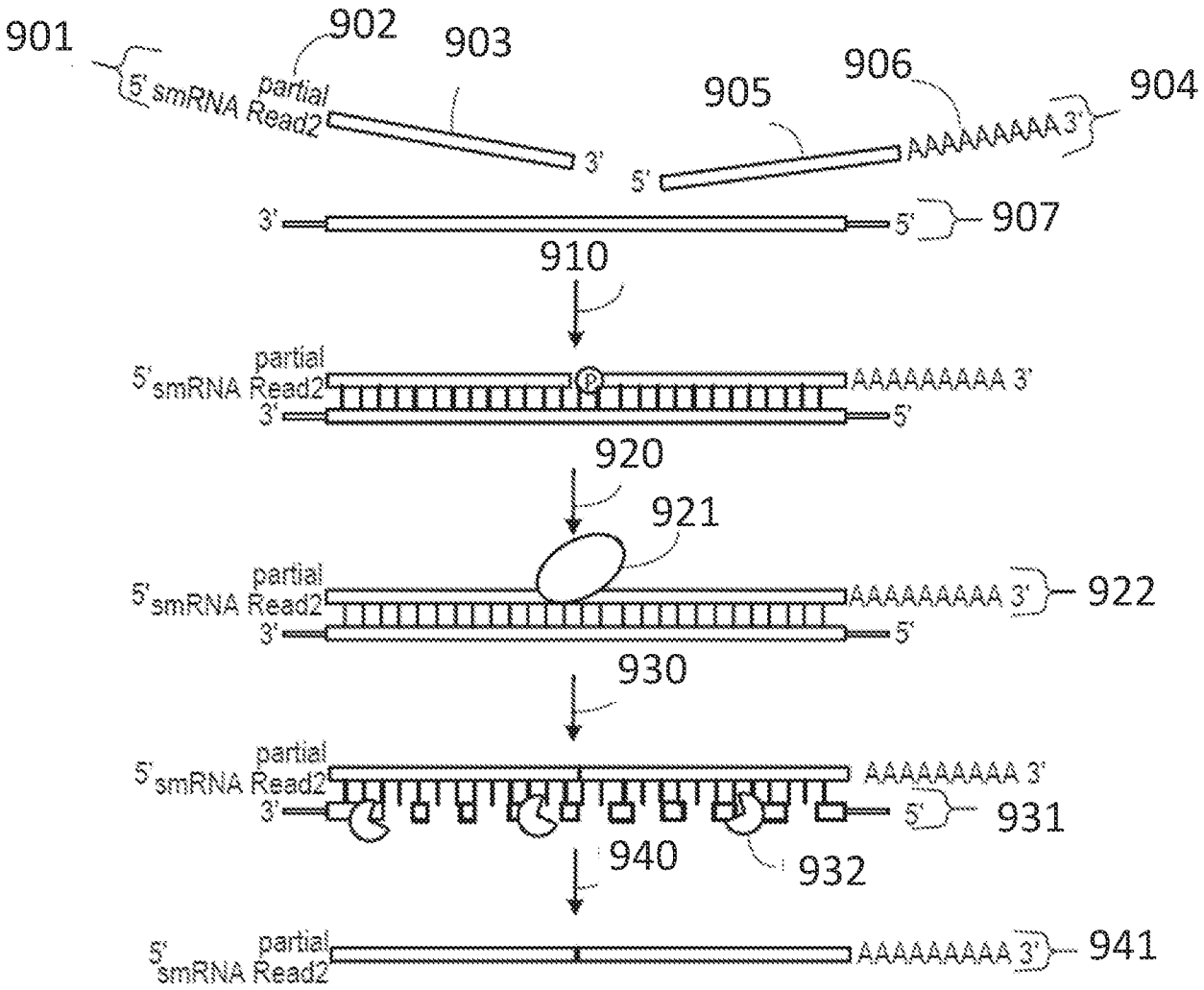

901
902
903
905
906
904
5' smRNA
partial Read2
3'
AAAAAAAAA 3'
5'
907
910
920
921
922
930
931
932
940
941

FIG. 9A

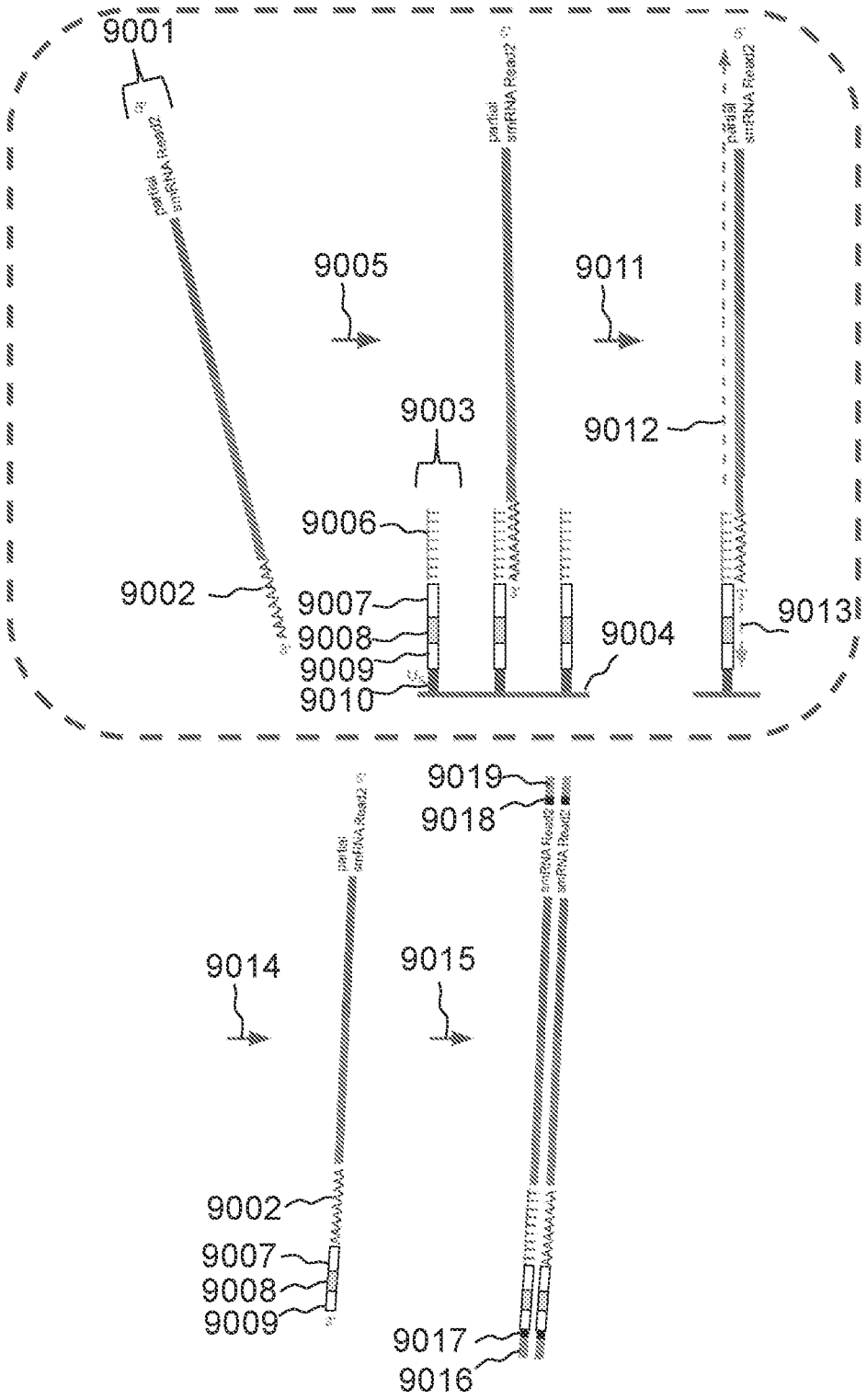

9001
partial smRNA Read2
9005
9011
partial smRNA Read2
partial smRNA Read2
9003
9012
9006
9002
9007
9008
9009
9010
9004
9013
9019
9018
partial smRNA Read2
smRNA Read2
smRNA Read2
9014
9015
9002
9007
9008
9009
9017
9016

FIG. 9B

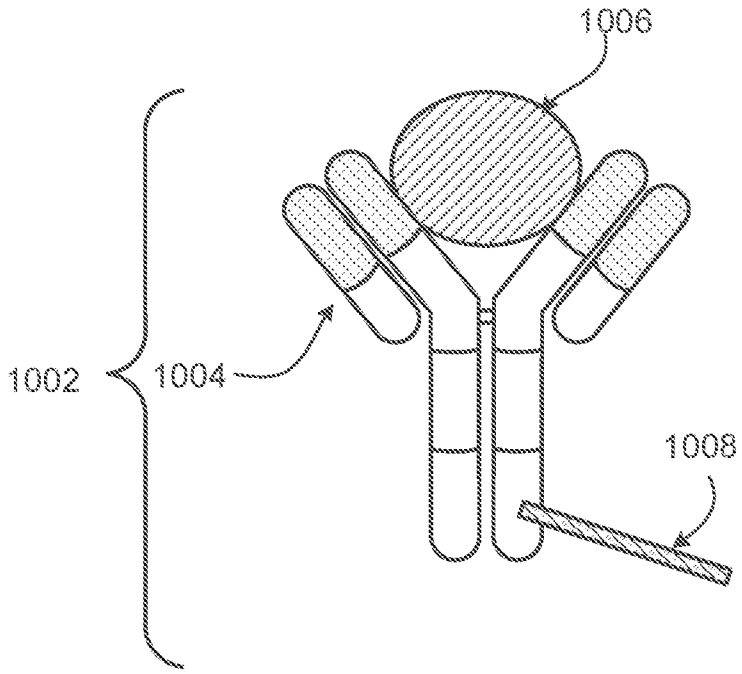


FIG. 10

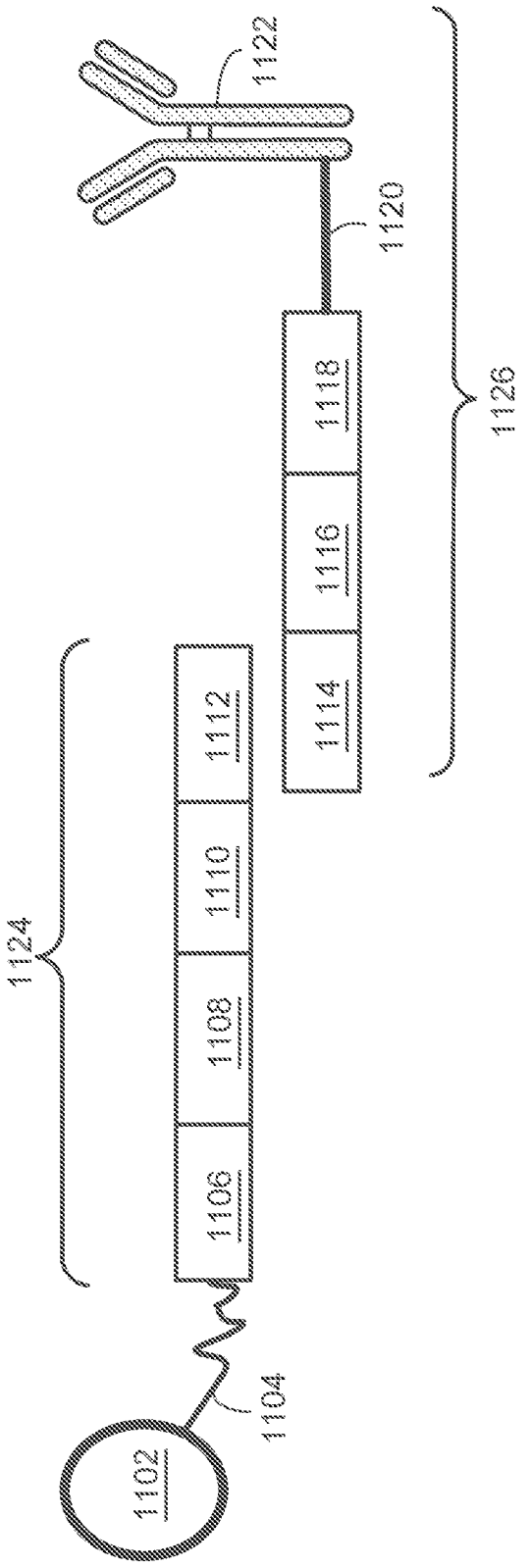
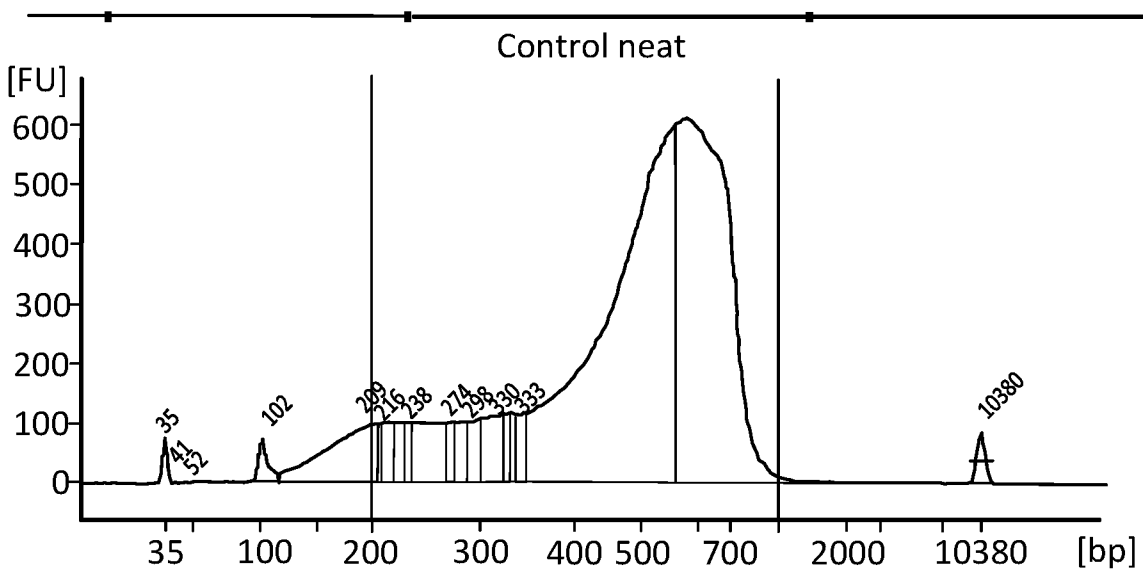

1102
1104
1106
1108
1110
1112
1114
1116
1118
1120
1122
1124
1126

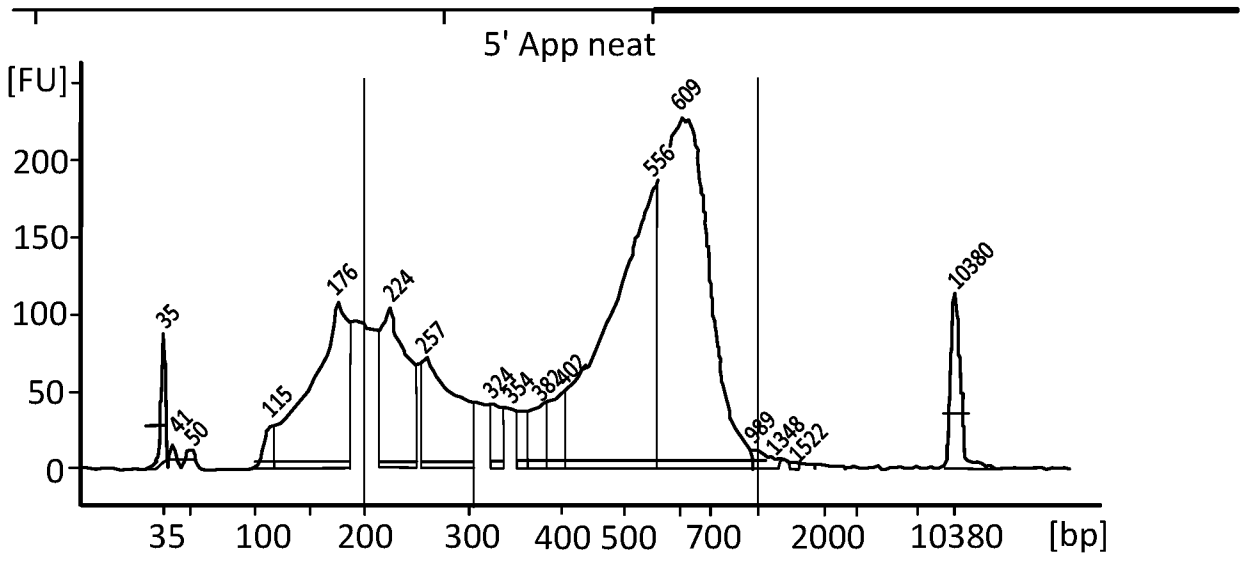
FIG. 11



Control

Cq = 14.05

FIG. 12A



5' App

Cq = 10.51

FIG. 12B

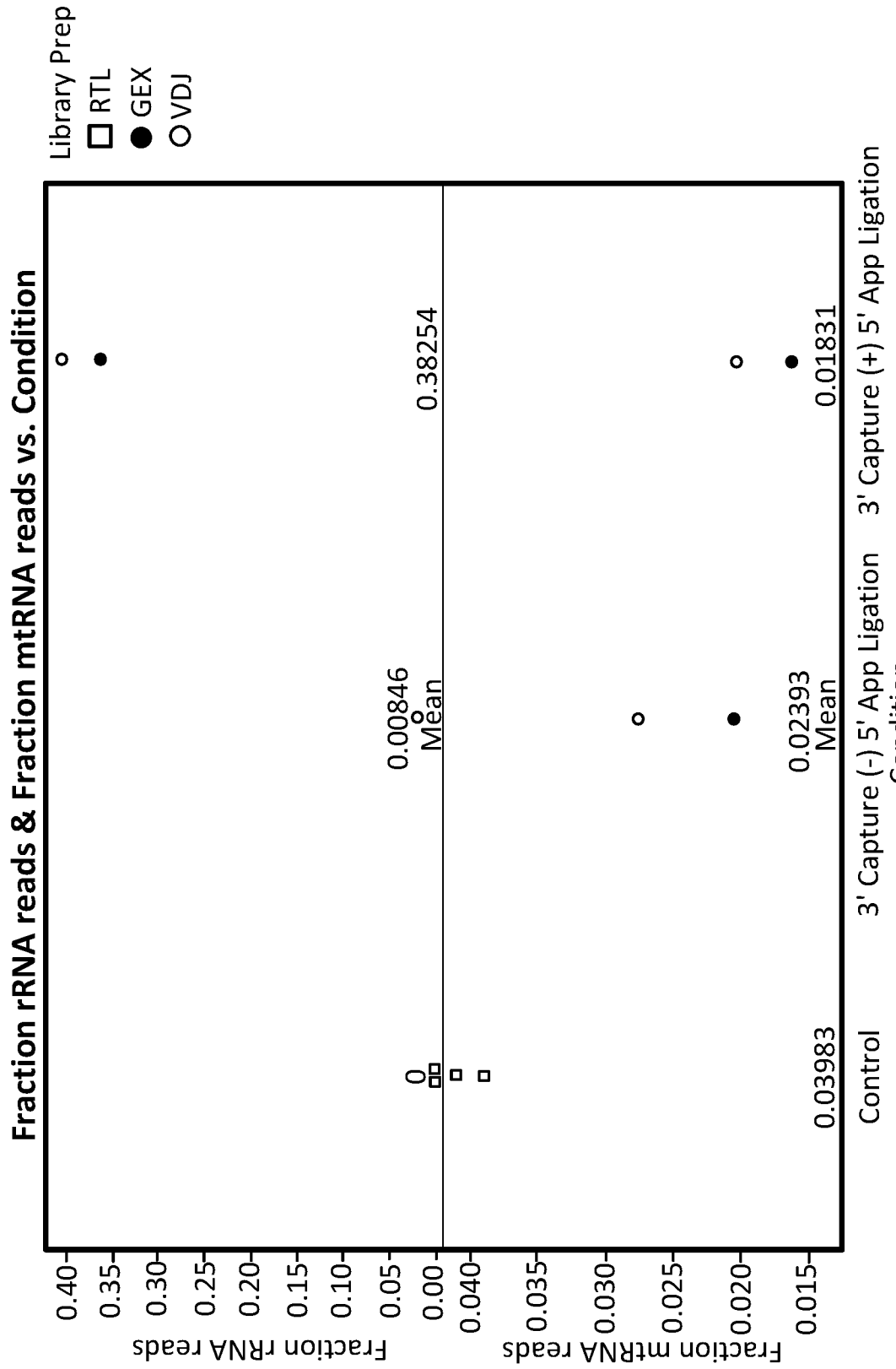

Fraction rRNA reads & Fraction mtRNA reads vs. Condition
Library Prep
RTL
GEX
VDJ
Fraction rRNA reads
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.00846
Mean
0.38254
Fraction mtRNA reads
0.040
0.035
0.030
0.025
0.020
0.015
0.02393
Mean
0.01831
0.03983
Control
3' Capture (-) 5' App Ligation
3' Capture (+) 5' App Ligation
Condition

FIG. 13

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/086181

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/6806 C12Q1/6844
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2020/123309 A1 (10X GENOMICS INC [US]; RAMACHANDRAN IYER ESWAR PRASAD [US] ET AL.) 18 June 2020 (2020-06-18)	159,160
A	page 64, paragraph 3; figures 1, 17 page 32	1-158
X	US 2022/333195 A1 (SCHNALL-LEVIN MICHAEL [US] ET AL) 20 October 2022 (2022-10-20) paragraph [1049]; example 3	159,160
X	WO 2021/142233 A1 (10X GENOMICS INC [US]; SPATIAL TRANSCRIPTOMICS AB [SE] ET AL.) 15 July 2021 (2021-07-15) page 38; claim 24; figures 6, 15	1-160
A	WO 2021/133849 A1 (10X GENOMICS INC [US]) 1 July 2021 (2021-07-01) page 65; claim 34	1-160

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

☒ See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 March 2024	19/03/2024

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Aslund, Fredrik

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2023/086181

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2020123309 A1	18-06-2020	CN 113439124 A	24-09-2021
		CN 113767176 A	07-12-2021
		CN 113767177 A	07-12-2021
		EP 3894585 A2	20-10-2021
		EP 3894586 A2	20-10-2021
		EP 3894587 A1	20-10-2021
		EP 3894588 A1	20-10-2021
		EP 3894589 A2	20-10-2021
		EP 3894590 A2	20-10-2021
		EP 3894591 A2	20-10-2021
		EP 3894592 A2	20-10-2021
		SG 11202105824R A	29-06-2021
		US 2020277663 A1	03-09-2020
		US 2020277664 A1	03-09-2020
		US 2022025446 A1	27-01-2022
		US 2022025447 A1	27-01-2022
		US 2022033888 A1	03-02-2022
		US 2022049294 A1	17-02-2022
		US 2022064630 A1	03-03-2022
		US 2022112486 A1	14-04-2022
		US 2022241780 A1	04-08-2022
		US 2022290217 A1	15-09-2022
		WO 2020123301 A2	18-06-2020
		WO 2020123305 A2	18-06-2020
		WO 2020123309 A1	18-06-2020
		WO 2020123311 A2	18-06-2020
		WO 2020123316 A2	18-06-2020
		WO 2020123317 A2	18-06-2020
		WO 2020123318 A1	18-06-2020
		WO 2020123319 A2	18-06-2020
		WO 2020123320 A2	18-06-2020
US 2022333195 A1	20-10-2022	US 2022333195 A1	20-10-2022
		US 2023340596 A1	26-10-2023
		US 2024018589 A1	18-01-2024
WO 2021142233 A1	15-07-2021	CN 115715329 A	24-02-2023
		EP 4087945 A1	16-11-2022
		EP 4339299 A2	20-03-2024
		WO 2021142233 A1	15-07-2021
WO 2021133849 A1	01-07-2021	AU 2020412766 A1	09-06-2022
		CN 114885610 A	09-08-2022
		DK 3891300 T3	22-05-2023
		EP 3891300 A1	13-10-2021
		EP 4219754 A1	02-08-2023
		ES 2946357 T3	17-07-2023
		FI 3891300 T3	10-05-2023
		SG 11202106899S A	29-09-2021
		US 2021285046 A1	16-09-2021
		US 2021348221 A1	11-11-2021
		US 2022282329 A1	08-09-2022
		US 2023064372 A1	02-03-2023
		US 2023160008 A1	25-05-2023
		US 2023392204 A1	07-12-2023
		WO 2021133849 A1	01-07-2021