

Provisional Patent Application No. 62/812,219, filed February 28, 2019, U.S. Provisional Patent Application No. 62/819,458, filed March 15, 2019, U.S. Provisional Patent Application No. 62/839,223, filed April 26, 2019, U.S. Provisional Patent Application No. 62/839,320, filed April 26, 2019, U.S. Provisional Patent Application No. 62/839,346, filed April 26, 2019, U.S. Provisional Patent Application No. 62/842,463, filed May 2, 2019, U.S. Provisional Patent Application No. 62/860,993, filed June 13, 2019, U.S. Provisional Patent Application No. 62/839,526, filed April 26, 2019 and U.S. Provisional Patent Application No. 62/858,331, filed on June 7, 2019. The contents of each of these applications are incorporated herein by reference in their entireties.

10

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, and 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 contact of an intact tissue or a portion of a tissue, or provide a lot of analyte data for single cells, but fail to provide information regarding the position of the single cell in a parent biological sample (e.g., tissue sample).

Chromatin structure can be different between cells in a biological sample or between biological samples from the same tissue. Assaying differences in accessible chromatin can be indicative of transcriptionally active sequences, e.g., genes, in a particular cell. Further understanding the transcriptionally active regions within chromatin will enable identification of which genes contribute to a cell's function and/or phenotype.

SUMMARY

The present disclosure generally describes methods for spatially analyzing genomic DNA present in a biological sample. In one example, the method comprises providing an array with a plurality of capture probes such that a capture probe of the plurality comprises a spatial barcode

and a capture domain; permeabilizing the biological sample under conditions sufficient to make the genomic DNA in the biological sample accessible to a transposon insertion; providing a transposon sequence and a transposase enzyme to the biological sample under conditions wherein the transposon sequence is inserted into the genomic DNA; allowing the transposase enzyme to excise the inserted transposon sequence from the genomic DNA thus generating fragmented genomic DNA; contacting the biological sample comprising the fragmented genomic DNA with an array under conditions such that a capture probe interacts with the fragmented genomic DNA; and correlating the location of the capture probe on the array to a location in the biological sample, thereby spatially analyzing the fragmented genomic DNA.

In one aspect, the present disclosure refers to a method for spatial analysis of genomic DNA present in a biological sample, the method comprising: providing an array comprising a plurality of capture probes, wherein a capture probe of the plurality of capture probes comprises a spatial barcode and a capture domain; permeabilizing the biological sample under conditions sufficient to make the genomic DNA in the biological sample accessible to transposon insertion; providing a transposon sequence and a transposase enzyme to the biological sample under conditions wherein the transposon sequence is inserted into the genomic DNA; allowing the transposase enzyme to excise the inserted transposon sequence from the genomic DNA, thereby generating fragmented genomic DNA; contacting the biological sample comprising the fragmented genomic DNA with the array under conditions where the capture probe interacts with the fragmented genomic DNA; and correlating the location of the capture probe on the array to a location in the biological sample, thereby spatially analyzing the fragmented genomic DNA.

In some examples, the array comprising a plurality of capture probes are provided on a substrate. In some examples, the array comprising the plurality of capture probes is provided on a feature. In some examples, the capture probe is directly or indirectly attached. In some examples, the array comprising the plurality of capture probes is provided on the feature on the substrate. In some examples, the substrate comprises a microfluidic channel. In some examples, the capture probe further comprises one or more of a cleavage domain, a functional domain, and a unique identifier, or combinations thereof.

In some examples, a further migration step comprising a step wherein the fragmented genomic DNA is migrated to the substrate. In some examples, the migration step is an active migration step comprising applying an electric field to the fragmented genomic DNA. In some examples, the migration step is a passive migration step comprising diffusion. In some examples, the migration of the fragmented genomic DNA from the biological sample comprises exposing the biological sample and the feature to heat. In some examples, the biological sample is immobilized on the substrate.

In some examples, the transposase enzyme is a dimer comprised of a first monomer complexed with a first adapter comprising a transposon end sequence and a sequence complementary to the capture domain and wherein a second monomer is complexed with a second adapter comprising a transposon end sequence and a second adapter sequence, wherein the transposase enzyme ligates the first adapter and the second adapter to the fragmented genomic DNA. In some examples, the first adapter and the second adapter have a 5' end and a 3' end, wherein the 5' end is phosphorylated in situ. In some examples, prior to fragmenting the DNA, the 5' end of the first adapter complexed with the first monomer and the

second adapter complexed with the second monomer are phosphorylated. In some examples, the step of phosphorylating the 5' end of the first adapter complexed with the first monomer and the second adapter complexed with the second monomer comprises contacting a first monomer: first adapter complex and a second monomer: second adapter complex with a polynucleotide kinase in the presence of ATP.

In some examples, the capture domain of the capture probe comprises a sequence that hybridizes to the sequence complementary to the capture domain of the first adapter. In some examples, the capture probe is a partially double stranded molecule comprising a first strand comprising the capture domain hybridized to a second strand, and wherein the first strand templates the ligation of the first adapter to the second strand. In some examples, the first adapter sequence complementary to the capture domain, or portion thereof, hybridized to the capture probe templates the ligation and ligating the 5' end of the first adapter to the 3' end of the capture probe. In some examples, the capture probe comprises a surface probe and a splint oligonucleotide and the splint oligonucleotide comprises a sequence complementary to a hybridization domain of the surface probe. In some examples, the splint oligonucleotide comprises the capture domain with a sequence complementary to the first adapter, or portion thereof. In some examples, the splint oligonucleotide hybridizes to the first adapter, or portion thereof, and to the hybridization domain of the surface probe, or portion thereof. In some examples, ligation is performed in the presence of the splint oligonucleotide, thereby ligating the surface probe of the capture probe and the first adapter.

In some examples, the fragmented genomic DNA hybridized to the capture probe by the first adapter is an extension template used to produce an extended capture probe that comprises the sequences of the spatial barcode and a sequence complementary to the fragmented genomic DNA. In some examples, the capture probe hybridized to the fragmented genomic DNA is extended with a DNA polymerase. In some examples, the DNA polymerase has strand displacement activity. In some examples, a further step of gap repair of single stranded breaks in the fragmented genomic DNA.

In some examples, the sequence complementary to the capture domain is a unique sequence. In some examples, the capture probe is ligated to the fragmented genomic DNA by a DNA ligase enzyme. In some examples, the transposase enzyme is a Tn5 transposase, or a functional derivative thereof. In some examples, the Tn5 transposase enzyme comprises a

sequence having at least 80% identity to SEQ ID NO: 1. In some examples, the transposase enzyme is a Mu transposase, or the functional derivative thereof. In some examples, the Mu transposase enzyme comprises a sequence having at least 80% identity to SEQ ID NO: 2. In some examples, the transposon end sequence comprises a sequence having at least 80% identity to SEQ ID NO: 8. In some examples, the transposon end sequence comprises a sequence having at least 80% identity to any one of SEQ ID NO: 9 to 14.

In some examples, permeabilizing the biological sample is performed under a chemical permeabilization condition, an enzymatic permeabilization condition, or both. In some examples, the chemical permeabilization condition comprises contacting the biological sample with an alkaline solution. In some examples, the enzymatic permeabilization condition comprises contacting the biological sample with an acidic solution comprising a protease enzyme. In some examples, the protease enzyme is an aspartyl protease, preferably a pepsin enzyme, a pepsin-like enzyme, or the functional equivalent thereof. In some examples, the pepsin enzyme, the pepsin-like enzyme, or the functional equivalent thereof, comprises a sequence having at least 80% identity to SEQ ID NO: 3 or 4.

In some examples, the enzymatic permeabilization condition comprises contacting the biological sample with a zinc endopeptidase, a collagenase enzyme, a collagenase-like enzyme, or a functional equivalent thereof; a serine protease, a proteinase K enzyme, a proteinase K-like enzyme, or a functional equivalent thereof; or both. In some examples, the collagenase enzyme, the collagenase-like enzyme, or the functional equivalent thereof comprises a sequence having at least 80% identity to SEQ ID NO: 5 or 6. In some examples, the proteinase K enzyme, the proteinase K-like enzyme, or the functional equivalent thereof comprises a sequence having at least 80% identity to SEQ ID NO: 7.

In some examples, the fragmented genomic DNA hybridized to the capture probe as the extension template generates a DNA molecule. In some examples, the fragmented genomic DNA hybridized to the capture probe acts as a ligation template to generate a DNA molecule. In some examples, the step comprising a step of analyzing the generated DNA molecule. In some examples, the step of analyzing the DNA molecule includes sequencing. In some examples, the step of correlating the spatial barcode of the capture probe with the fragmented genomic DNA associated with the capture probe spatially analyzes the fragmented

genomic DNA. In some examples, the biological sample is imaged before or after contacting the biological sample with the substrate.

In another example, the present disclosure generally describes a kit for use in a method of spatially detecting nucleic acids of a biological sample, wherein the kit comprises any two or more of an array on which plurality of capture probes are present; one or more biological sample permeabilization reagents; one or more transposase enzymes; one or more reverse transcriptases; and one or more cleavage enzymes.

In another aspect, the present disclosure refers to a kit for use in a method of spatially detecting nucleic acids of a biological sample as disclosed herein, wherein the kit comprises any two or more of: (i) an array on which a plurality of capture probes are present; (ii) one or more biological sample permeabilization reagents; (iii) one or more transposase enzymes; (iv) one or more reverse transcriptases; and (v) one or more cleavage enzymes.

In a different example, the present disclosure generally describes a method for spatial analysis of genomic DNA and RNA present in a biological sample wherein an array is provided and the array comprises a plurality of capture probes, wherein a first capture probe of the plurality of capture probes comprises a spatial barcode and a first capture domain, and wherein a second capture probe of the plurality of capture probes comprises the spatial barcode and a second capture domain; permeabilizing the biological sample under conditions sufficient to make the genomic DNA in the biological sample accessible to transposon insertion; providing a transposon sequence and a transposase enzyme to the biological sample under conditions wherein the transposon sequence is inserted into the genomic DNA; allowing the transposase enzyme to excise the inserted transposon sequence from the genomic DNA, thereby generating fragmented genomic DNA; contacting the biological sample comprising the fragmented genomic DNA and RNA with the array under conditions where the first capture domain interacts with the fragmented genomic DNA and the second capture domain interacts with the RNA; and correlating the location of the first capture probe on the array to a location in the biological sample and correlating the location of the second capture probe on the array to a location in the biological sample, thereby spatially analyzing the fragmented genomic DNA and RNA at the location in the biological sample.

In another aspect, the present disclosure refers to a method for spatial analysis of genomic DNA and RNA present in a biological sample, the method comprising: providing an array, wherein the array comprises a plurality of capture probes, wherein a first capture probe of the plurality of capture probes comprises a spatial barcode and a first capture domain, and wherein a second capture probe of the plurality of capture probes comprises the spatial barcode and a second capture domain; permeabilizing the biological sample under conditions sufficient to make the genomic DNA in the biological sample accessible to transposon insertion; providing a transposon sequence and a transposase enzyme to the biological sample under conditions wherein the transposon sequence is inserted into the genomic DNA; allowing the transposase enzyme to excise the inserted transposon sequence from the genomic DNA, thereby generating fragmented genomic DNA; contacting the biological sample comprising the fragmented genomic DNA and RNA with the array under conditions where the first capture domain interacts with the fragmented genomic DNA and the second capture domain interacts with the RNA; and correlating the location of the first capture probe on the array to a location in the biological sample and correlating the location of the second capture probe on the array to a location in the biological sample, thereby spatially analyzing the fragmented genomic DNA and RNA at the location in the biological sample.

In some examples, the RNA is a mRNA. In some examples, the first capture domain and the second capture domain are identical. In some examples, the first capture domain and the second capture domain comprise a homopolymeric poly (T) sequence. In some examples, the first capture domain and the second capture domain are different. In some examples, the first capture domain comprises a random sequence and the second capture domain comprises a poly (T) sequence. In some examples, the array comprising the plurality of capture probes is provided on a substrate. In some examples, the array comprising the

plurality of capture probes is provided on a feature. In some examples, the feature comprises the first capture probe, the second capture probe, or both. In some examples, the first capture probe, the second capture probe, or both, are directly or indirectly attached. In some examples, the array comprising the plurality of capture probes is provided on the feature on the substrate. In some

examples, the substrate comprises a microfluidic channel. In some examples, the first capture probe, the second capture probe, or both, comprise one or more of a cleavage domain, a functional domain, and a unique identifier, or combinations thereof.

In some examples, there is a migration step wherein the fragmented genomic DNA and the RNA are migrated to the substrate. In some examples, the migration step is an active migration step. In some examples, the migration step is a passive migration step. In some examples, the migration of the fragmented genomic DNA and the RNA from the biological sample comprises exposing the biological sample to heat. In some examples, the biological sample is immobilized on the substrate.

In some examples, the fragmented genomic DNA is repaired by ligating breaks with a ligase enzyme. In some examples, single stranded breaks in the fragmented genomic DNA undergo gap repair. In some examples, a sequence complementary to the first capture domain of the first capture probe is introduced to the fragmented genomic DNA. In some examples, the first capture domain of the first capture probe hybridizes to the sequence complementary to the capture domain introduced to the fragmented genomic DNA. In some examples, the random sequence of the first capture domain hybridizes the fragmented genomic DNA. In some examples, the second capture domain of the second capture probe hybridizes to a complementary sequence in the mRNA. In some examples, the sequence complementary to the first capture domain and the complementary sequence in the mRNA is a homopolymeric sequence. In some examples, the homopolymeric sequence is a poly(A) sequence.

In some examples, extension of the first capture probe using the fragmented genomic DNA as an extension template, and extension of the second capture probe using the RNA as an extension template is performed. In some examples, extending the first capture probe is performed with a DNA polymerase. In some examples, extending the second capture probe is performed with reverse transcriptase.

In some examples, transposase is a Tn5 transposase, or a functional derivative thereof. In some examples, the Tn5 transposase enzyme comprises a sequence having at least

80% identity to SEQ ID NO: 1. In some examples, the transposase enzyme is a Mu transposase enzyme, or a functional derivative thereof. In some examples, the Mu transposase enzyme comprises a sequence having at least 80% identity to SEQ ID NO: 2. In some examples, the transposase enzyme is complexed with an adapter comprising a transposon end sequence. In
5 some examples, the transposon end sequence comprises a sequence having at least 80% identity to SEQ ID NO: 8. In some examples, the transposon end sequence comprises a sequence having at least 80% identity to any one of SEQ ID NO: 9 to 14.

In some examples, a step of permeabilizing the biological sample is performed. In some examples, 7. The method of any one of claims 51 to 86, wherein permeabilizing the biological
10 sample is performed under a chemical permeabilization condition, an enzymatic permeabilization condition, or both. In some examples, the chemical permeabilization condition comprises contacting the biological sample with an alkaline solution. In some examples, the enzymatic permeabilization condition comprises contacting the biological sample with an acidic solution comprising a protease enzyme. In some examples, the protease enzyme is an aspartyl
15 protease, preferably a pepsin enzyme, a pepsin-like enzyme, or a functional equivalent thereof. In some examples, the pepsin enzyme, the pepsin-like enzyme, or functional equivalent thereof, comprises a sequence having at least 80% identity to SEQ ID NO: 3 or 4. In some examples, the enzymatic permeabilization condition comprises contacting the biological sample with a zinc endopeptidase, a collagenase enzyme, a collagenase-like enzyme, or a functional equivalent
20 thereof; a serine protease, a proteinase K enzyme, a proteinase K-like enzyme, or a functional equivalent thereof; or both. In some examples, the collagenase enzyme, the collagenase-like enzyme, or the functional equivalent thereof comprises a sequence having at least 80% identity to SEQ ID NO: 5 or 6. In some examples, the proteinase K enzyme, the proteinase K-like enzyme, or the functional equivalent thereof comprises a sequence having at least 80% identity to SEQ ID
25 NO: 7.

In some examples, step of analyzing the DNA molecule includes sequencing. In some examples, correlating the spatial barcode of the first capture probe with the fragmented genomic DNA associated with the first capture probe spatially analyzes the fragmented genomic DNA. In some examples, correlating the spatial barcode of the second capture probe with the mRNA
30 associated with the second capture probe spatially analyzes the mRNA. In some

examples, the biological sample is imaged before or after contacting the biological sample with the substrate.

In another aspect, the present disclosure refers to a method for determining a location of an analyte in a biological sample, comprising: (a) contacting a biological sample with a substrate
5 comprising a plurality of capture probes, wherein a capture probe of the plurality comprises a spatial barcode and a capture domain that specifically binds to an analyte from the biological sample; and (b) determining (i) all or a portion of the analyte sequence specifically bound to the capture domain, or a complement thereof, and (ii) all or a portion of the spatial barcode, or a complement thereof, and using the determined sequences of (i) or (ii) to determine the location
10 of the analyte in the biological sample, wherein the biological sample is a cell culture, an organism, or an organoid.

In another aspect, the present disclosure refers to a method for spatial analysis of a nucleic acid in a biological sample, the method comprising: (a) contacting the biological sample to an array comprising a plurality of capture probes immobilized on the array, wherein a capture
15 probe of the plurality of capture probes comprises a spatial barcode and a capture domain; (b) contacting a polymer material to the biological sample, and activating the polymer material to form a hydrogel-embedded biological sample; (c) permeabilizing the hydrogel-embedded biological sample under conditions sufficient to allow the nucleic acid in the biological sample to specifically bind to the capture domain; (d) amplifying the RNA specifically bound to the
20 capture domain to generate an amplicon having a sequence, wherein the amplicon comprises all or a portion of the spatial barcode or a complement thereof, and all or a portion of the nucleic acid sequence or a complement thereof; and (e) determining the sequence of the amplicon by sequential hybridization and detection with a plurality of labelled probes, and using the determined amplicon sequence to identify the location of the nucleic acid in the biological
25 sample.

In another aspect, the present disclosure refers to a nucleic acid comprising: (a) a first single-stranded oligonucleotide comprising: a spatial barcode; a primer binding site; a unique molecular identifier; and a first hybridization sequence; and (b) a splint oligonucleotide comprising: a second hybridization sequence that specifically binds to the first hybridization
30 sequence; and a capture domain.

In another aspect, the present disclosure refers to a composition comprising: (a) a nucleic acid disclosed herein; and (b) a single-stranded tagment comprising a first adaptor that comprises a sequence that is capable of specifically binding to the capture domain.

5 In another aspect, the present disclosure refers to a composition comprising: (a) a nucleic acid disclosed herein; (b) DNA from a biological sample; and (c) a transposase dimer comprising: (i) a first transposase monomer complexed with a first adaptor comprising a first transposon end sequence and a sequence that is capable of specifically binding to the capture domain; (ii) a second transposase monomer complexed with a second adaptor comprising a
10 second transposon end sequence and a primer binding site or a complement thereof.

In another aspect, the present disclosure refers to a composition comprising: (a) a nucleic acid disclosed herein; (b) DNA from a biological sample; and (c) a first transposase monomer complexed with a first adaptor comprising a first transposon end sequence and a sequence that is capable of specifically binding to the capture domain; and (d) a second transposase monomer
15 complexed with a second adaptor comprising a second transposon end sequence and a primer binding site or a complement thereof.

All publications, patents, patent applications, and information available on the internet and 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
20 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
25 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,
30 unless expressly stated otherwise, or unless the context of the usage clearly indicates otherwise.

Various examples of the features of this disclosure are described herein. However, it should be understood that such examples 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 examples 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. 1 shows an exemplary spatial analysis workflow.

FIG. 2 shows an exemplary spatial analysis workflow.

FIG. 3 shows an exemplary spatial analysis workflow.

FIG. 4 shows an exemplary spatial analysis workflow.

FIG. 5 shows an exemplary spatial analysis workflow.