

## **BIOMOLECULAR PROBES AND METHODS OF DETECTING GENE AND PROTEIN EXPRESSION**

### **BACKGROUND OF THE INVENTION**

[0001] This application claims priority to, and the benefit of, U.S. Provisional Application No. 62/629,180, filed February 12, 2018 and U.S. Provisional Application No. 62/771,212, filed November 26, 2018. The contents of each of the aforementioned patent applications are incorporated herein by reference in their entireties.

### **SEQUENCE LISTING**

[0002] The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on February 1, 2019, is named "NATE-037\_001WO.txt" and is 48.4 KB in size.

### **BACKGROUND OF THE INVENTION**

[0003] Standard immunohistochemical and *in situ* hybridization methods allow for simultaneous detection of, at most, six to ten protein or nucleic acid targets, with three to four targets being typical. There exists a need for probes, compositions, methods, and kits for simultaneous, multiplexed detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell. Furthermore, there is a need for such systems to be adaptable for use with existing sequencing technologies that are already available to a large number of end users.

### **SUMMARY OF THE INVENTION**

[0004] The present disclosure relates to probes, compositions, methods, and kits for simultaneous, multiplexed, spatial detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell.

[0005] The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a

unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the ligation product produced in step (4); and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0006]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises: a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias, and wherein at least one of the first or second amplification primer binding sites identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; (5) amplifying the ligation product produced in step (4); and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0007]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a

unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises: a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the extension product produced in step (4) using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0008]** The nucleic acid adapter of step (4) can be a partially double-stranded nucleic acid molecule. A partially double-stranded nucleic acid adapter can comprise a double-stranded annealed region, a first single-stranded mismatched region and a second single-stranded mismatched region. The first single-stranded mismatched region and the second single stranded mismatched region can be present on opposing sides of the double-stranded annealed region.

**[0009]** The nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter.

**[0010]** The constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter.

**[0011]** A unique molecular identifier can be present in at least one of the first or second single-stranded mismatched regions of a partially double-stranded nucleic acid adapter.

**[0012]** The first amplification primer binding site can be present in the first single-stranded mismatched region of a partially double-stranded nucleic acid adapter and the second amplification primer binding site can be present in the second single-stranded mismatched region of the same partially double-stranded nucleic acid adapter.

**[0013]** The methods of the present disclosure described in the preceding can further comprise prior to step (4), performing an end repair reaction. The methods can also further comprise prior to step (4), performing a tailing reaction to attach a single nucleotide overhang to the 3' ends of the identifier oligonucleotide. The methods can further comprise, prior to step (4), performing an end repair reaction and a tailing reaction to attach a single nucleotide overhang to the 3' ends of the identifier oligonucleotide. The tailing reaction and the end repair reaction can be performed sequentially or concurrently.

**[0014]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a first amplification primer binding site, and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier, a second amplification primer binding site and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the ligation product produced in step (4); and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0015]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises: a first amplification primer binding site and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) ligating to the released identifier oligonucleotide at least one nucleic acid adapter, wherein the nucleic acid adapter comprises a nucleic acid sequence comprising a unique

molecular identifier, a second amplification primer binding site, and optionally, a constant nucleic acid sequence to minimize ligation bias; (5) amplifying the extension product produced in step (4) using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; and (6) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0016]** The nucleic acid adapter of step (4) can be a partially double-stranded nucleic acid molecule. A partially double-stranded nucleic acid adapter can comprise a double-stranded annealed region and a single-stranded mismatched region.

**[0017]** The nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter.

**[0018]** The constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of a partially double-stranded nucleic acid adapter. The constant nucleic acid sequence can also comprise a cleavable moiety. The cleavable moiety can be an enzymatically cleavable moiety. The enzymatically cleavable moiety can be a USER sequence.

**[0019]** A unique molecular identifier can be present in the single-stranded mismatched region of a partially double-stranded nucleic acid adapter.

**[0020]** The second amplification primer binding site can be present in the single-stranded mismatched region of a partially double-stranded nucleic acid adapter.

**[0021]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a first amplification primer binding site, and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a single stranded

nucleic acid template, wherein the nucleic acid template comprises a region complementary to the unique nucleic acid sequence of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier, a nucleic acid sequence complementary to a second amplification primer binding site and optionally, an affinity molecule; (5) extending the identifier oligonucleotide of step (4) to form an extension product complementary to the single stranded nucleic acid template, wherein the extension product comprises the identifier oligonucleotide, the nucleic acid sequence complementary to the unique molecular identifier, and the second amplification primer binding site; (6) amplifying the extension product produced in step (5) using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0022]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a first amplification primer binding site and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a single stranded nucleic acid template, wherein the nucleic acid template comprises a region complementary to the unique nucleic acid sequence of the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier, a nucleic acid sequence complementary to a second amplification primer binding site and optionally, an affinity molecule; (5) extending the identifier oligonucleotide of step (4) to form an extension product complementary to the single stranded nucleic acid template, wherein the extension product comprises the identifier oligonucleotide, the nucleic acid sequence complementary to the nucleic

acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, the nucleic acid sequence complementary to the unique molecular identifier and the second amplification primer binding site; (6) amplifying the extension product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0023]** The single stranded nucleic acid template can further comprise an affinity molecule. In aspects in which the single stranded nucleic acid template comprises an affinity molecule, the methods of the present disclosure described in the preceding can further comprise an affinity purification step between steps (4) and (5).

**[0024]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding site, and wherein the first the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0025]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a second amplification primer binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0026]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the

identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding site, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0027]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a first amplification primer binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide and a second amplification primer binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying

the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0028]** In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released is located on the first nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be located 5' to the first amplification primer binding site.

**[0029]** In aspects in which a unique molecular identifier is located on the second nucleic acid probe, the unique molecular identifier can be located 3' to the second amplification primer binding site.

**[0030]** In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a unique molecular identifier are present in the first nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a unique molecular identifier can be located 5' to the first amplification primer binding site.

**[0031]** In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a unique molecular identifier are present in the second nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the unique molecular identifier can be located 3' to the second amplification primer binding site.

**[0032]** In aspects in which a unique molecular identifier is present in the first nucleic acid probe and the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released is present in the second nucleic acid probe, the unique molecular identifier can be located 5' to the first amplification primer binding site and the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be located 3' to the second amplification binding site.

**[0033]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding

domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier, a nucleic acid sequence comprising a third unique molecular identifier and a second flow cell binding site, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0034]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier oligonucleotide, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to a portion of the identifier

oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier and a second flow cell binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a third unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) performing nick repair such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0035]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier, a nucleic acid sequence comprising a third unique molecular identifier and a second flow cell binding site, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6)

amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0036]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a first amplification primer binding site, a nucleic acid sequence comprising a first unique molecular identifier and a first flow cell binding site, and wherein the second nucleic acid probe comprises a nucleic acid complementary to the identifier oligonucleotide, a nucleic acid sequence comprising a second unique molecular identifier and a second flow cell binding site, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence comprising a third unique molecular identifier, and wherein at least one of the first or second nucleic acid probes comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping; (5) performing a gap extension and nick repair reaction such that the hybridized first and second nucleic acid probes are ligated together; (6) amplifying the ligation product produced in step (5); and (7) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0037]** In aspects in which the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the first unique molecular identifier are present in the first nucleic acid probe, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was

released and the first unique molecular identifier can be located 5' to the first flow cell binding site.

**[0038]** In aspects in which the second and the third unique molecular identifiers are present in the second nucleic acid probe, the second and the third unique molecular identifiers can be located 3' to the second flow cell binding site.

**[0039]** In some aspects, the first unique molecular identifier can be present in the first nucleic acid probe and can be located 5' to the first flow cell binding site. In other aspects, the second unique molecular identifier can be present in the second nucleic acid probe and can be located 3' to the second flow cell binding site.

**[0040]** In some aspects, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the third unique molecular identifier can be present in the first nucleic acid probe and can be located 5' to the first flow cell binding site.

**[0041]** In some aspects, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and the third unique molecular identifier can be present in the second nucleic acid probe and can be located 3' to the second flow cell binding site.

**[0042]** In some aspect, the third unique molecular identifier can be present in the first nucleic acid probe and can be located 5' to the first flow cell binding site. In this same aspect, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can be present in the second nucleic acid probe and can be located 3' to the second flow cell binding site.

**[0043]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, and a second amplification primer binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) amplifying the released identifier oligonucleotide using a

first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released; (5) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0044]** The present disclosure also provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a first amplification primer binding site and a second amplification primer binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) amplifying the released identifier oligonucleotide using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site, wherein at least one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released, and wherein at least one of the amplification primers comprises a nucleic acid sequence comprising a unique molecular identifier; and (5) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0045]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain and a capture probe binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier

oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a capture probe, wherein the capture probe comprises an affinity molecule and a region complementary to the capture probe binding site; and (5) Identifying the released identifier oligonucleotide by sequencing the amplified hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0046]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a capture probe binding site and a multiplexing probe binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a capture probe and a multiplexing probe, wherein the capture probe comprises an affinity molecule and a region complementary to the capture probe binding site, and wherein the multiplexing probe comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released and a region complementary to the multiplexing probe binding site; and (5) Identifying the released identifier oligonucleotide by sequencing the hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0047]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe, wherein the first nucleic acid probe comprises: a nucleic acid complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, and wherein the second nucleic acid probe comprises: a

nucleic acid complementary to a portion of the identifier oligonucleotide, and a second amplification primer binding site, and wherein the first and the second nucleic acid probes hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent but not overlapping; (5) ligating the hybridized first and second nucleic acid probes together; (6) amplifying the ligation product produced in step (5); and (7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0048]** The present disclosure provides a method for spatially detecting at least one target analyte in at least one cell from a tissue sample comprising: (1) contacting at least one target analyte in at least one cell in a tissue sample with at least one probe comprising a target binding domain and an identifier oligonucleotide, wherein the identifier oligonucleotide comprises: a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a nucleic acid sequence comprising a unique molecular identifier, a first amplification primer binding site, and a second amplification primer binding site; (2) providing a force to a location of the tissue sample sufficient to release the identifier oligonucleotide; (3) collecting the released identifier oligonucleotide; (4) amplifying the collected identifier oligonucleotide; (5) Identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[0049]** A In all methods of the present disclosure, the ligation process can be a nick ligation process. The nick ligation process can be a nick repair process.

**[0050]** In all methods of the present disclosure, the sequencing can be an enzyme free sequencing method.

**[0051]** In all methods of the present disclosure, the identifier oligonucleotide can be double-stranded. In aspects in which the identifier oligonucleotide is double-stranded, at least one of the two strands of the identifier oligonucleotide can comprise at least two separate nucleic acid molecules. Alternatively, at least one 3' end of an identifier oligonucleotide can comprise a single nucleotide overhang.

**[0052]** In all methods of the present disclosure, the identifier oligonucleotide can be single-stranded.

**[0053]** In all methods of the present disclosure, the unique nucleic acid sequence which identifies the target analyte bound to a target binding domain can comprise between about 5 nucleotides and about 40 nucleotides preferably about 35 nucleotides, preferably still about 10 nucleotides.

**[0054]** In all methods of the present disclosure, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can comprise between about 6 nucleotides and about 15 nucleotides, preferably about 12 nucleotides, preferably still about 10 nucleotides.

**[0055]** In all methods of the present disclosure, at least one of a first nucleic acid probe or a second nucleic acid probe can comprise an affinity molecule. For example, at least one of a first nucleic acid probe or a second nucleic acid probe can comprise a biotin.

**[0056]** In all methods of the present disclosure, an amplification primer binding site can comprise between about 18 nucleotides and about 40 nucleotides, preferably about 32 nucleotides, preferably still about 25 nucleotides. An amplification primer binding site can comprise an i7 sequence, wherein the i7 sequence comprises the sequence set forth in SEQ ID NO: 1. An amplification primer binding site can comprise an i5 sequence, wherein the i5 sequence comprises the sequence set forth in SEQ ID NO: 2.

**[0057]** In all methods of the present disclosure, an amplification primer can comprise a flow cell adapter sequence, wherein the flow cell adapter sequence is suitable for sequencing. An amplification primer can comprise a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3. An amplification primer can comprise a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

**[0058]** In all methods of the present disclosure, a flow cell binding site can comprise a flow cell adapter sequence, wherein the flow cell adapter sequence is suitable for sequencing. A flow cell binding site can comprise a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3. A flow cell binding site can comprise a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

**[0059]** In all methods of the present invention, at least one of the amplification primers can comprise an affinity molecule. For example, at least one of the amplification primers can comprise a biotin.

**[0060]** In all methods of the present disclosure, amplification can comprise performing PCR. Performing PCR can comprise an amplification primer.

**[0061]** An amplification primer can comprise a flow cell binding site. An amplification primer can comprise a nucleic sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released. An amplification primer can comprise a nucleic acid sequence complementary to an amplification primer binding site.

**[0062]** Any of the above aspects can be combined with any other aspect.

**[0063]** Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. In the Specification, the singular forms also include the plural unless the context clearly dictates otherwise; as examples, the terms “a,” “an,” and “the” are understood to be singular or plural and the term “or” is understood to be inclusive. By way of example, “an element” means one or more element. Throughout the specification the word “comprising,” or variations such as “comprises” or “comprising,” will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term “about.”

**[0064]** Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. The references cited herein are not admitted to be prior art to the claimed invention. In the case of conflict, the present Specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting. Other features and advantages of the disclosure will be apparent from the following detailed description and claim.

## BRIEF DESCRIPTION OF THE DRAWINGS

- [0065] The above and further features will be more clearly appreciated from the following detailed description when taken in conjunction with the accompanying drawings.
- [0066] Figure 1 is a schematic of a two-ended adapter ligation method of the present disclosure.
- [0067] Figure 2 is a schematic of a one-ended adapter ligation method of the present disclosure.
- [0068] Figure 3 is a schematic of a templated-primer extension method of the present disclosure.
- [0069] Figure 4 is a schematic of a template-extended identifier oligonucleotide of the present disclosure.
- [0070] Figure 5 is a schematic of a short probe hybridization method of the present disclosure.
- [0071] Figure 6 is a schematic of a short probe hybridization method of the present disclosure.
- [0072] Figure 7 is a schematic of a short probe hybridization method of the present disclosure.
- [0073] Figure 8 is a schematic of a long probe hybridization method of the present disclosure.
- [0074] Figure 9 is a schematic of a long probe hybridization method of the present disclosure.
- [0075] Figure 10 is a schematic of a long probe hybridization method of the present disclosure.
- [0076] Figure 11 is a schematic of a direct-PCR method of the present disclosure.
- [0077] Figure 12 is a schematic of an enzyme free method of the present disclosure.
- [0078] Figure 13 is a schematic of a multiplexed enzyme free method of the present disclosure.
- [0079] Figure 14 is a schematic of a probe of the present disclosure indirectly binding to a target nucleic acid.
- [0080] Figure 15 is a schematic of an identifier oligonucleotide-short nucleic acid probe complex of the present disclosure.
- [0081] Figure 16 is a schematic of a short probe hybridization method of the present disclosure.
- [0082] Figure 17 is a schematic of an identifier oligonucleotide-short nucleic acid probe complex of the present disclosure.
- [0083] Figure 18 is a schematic of a short probe hybridization method of the present disclosure.
- [0084] Figure 19 is a schematic of a direct-PCR method of the present disclosure.
- [0085] Figure 20 is a schematic overview of the methods of the present disclosure.
- [0086] Figure 21 shows the spatial detection of protein target analytes using the methods of the present disclosure.
- [0087] Figure 22 shows the spatial detection of RNA target analytes using the methods of the present disclosure.

[0088] Figure 23 shows the spatial detection of protein target analytes using the methods of the present disclosure.

[0089] Figure 24 shows the spatial detection of RNA target analytes using the methods of the present disclosure.

[0090] Figure 25 shows the spatial detection of protein target analytes using the methods of the present disclosure.

[0091] Figure 26 is a schematic of a probe of the present disclosure.

[0092] Figure 27 shows the use of probe tiling in the methods of the present disclosure.

[0093] Figure 28 shows the regions of interest selected on a tissue microarray.

[0094] Figure 29 is a series of graphs showing the read depth achieved using the methods of the present disclosure.

[0095] Figure 30 is a series of graphs showing the spatial detection of RNA target analytes in negative control samples using the methods of the present disclosure.

[0096] Figure 31 is a series of graphs showing the spatial detection of RNA target analytes in a HEK293 sample (top panel) and a Jurkat cell sample (bottom panel) using the methods of the present disclosure.

[0097] Figure 32 is a series of graphs showing the spatial detection of RNA target analytes in sixteen FFPE samples using the methods of the present disclosure.

[0098] Figure 33 is a graph showing the spatial detection of RNA target analytes in a HEK293 sample using the methods of the present disclosure.

[0099] Figure 34 is a graph showing the spatial detection of RNA target analytes in a Jurkat cell sample using the methods of the present disclosure.

## DETAILED DESCRIPTION OF THE INVENTION

[00100] The present disclosure is based in part on probes, compositions, methods, and kits for simultaneous, multiplexed spatial detection and quantification of protein and/or nucleic acid expression in a user-defined region of a tissue, user-defined cell, and/or user-defined subcellular structure within a cell using existing sequencing methods.

[00101] The present disclosure provides a comparison of the identity and abundance of target proteins and/or target nucleic acids present in a first region of interest (*e.g.*, tissue type, a cell (including normal and abnormal cells), and a subcellular structure within a cell) and the identity

and abundance of target proteins and/or target nucleic acids present in a second region of interest. There is no pre-defined upper limit to the number of regions of interest and comparisons that can be made; the upper limit relates to the size of the region of interest relative the size of the sample. As examples, when a single cell represents a region of interest, then a section may have hundreds to thousands of regions of interest; however, if a tissue section includes only two cell types, then the section may have only two regions of interest (each including only one cell type).

**[00102]** The present disclosure provides a higher degree of multiplexing than is possible with standard immunohistochemical or *in situ* hybridization methods. Standard immunohistochemical methods allow for maximal simultaneous detection of six to ten protein targets, with three to four protein targets being more typical. Similarly, *in situ* hybridization methods are limited to simultaneous detection of fewer than ten nucleic acid targets. The present disclosure provides detection of large combinations of nucleic acid targets and/or protein targets from a defined region of a sample. The present disclosure provides an increase in objective measurements by digital quantification and increased reliability and consistency, thereby enabling comparison of results among multiple centers.

**[00103]** Various compositions and methods of the present disclosure are described in full detail herein.

**[00104]** In one aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a “two-ended adapter ligation method”.

**[00105]** A two ended-adapter ligation method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The probes and samples of the present disclosure are described in further detail herein. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide

can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain.

**[00106]** Following contacting the at least one target analyte with the at least one probe, a two-ended adapter ligation method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00107]** Following release of the identifier oligonucleotide, a two-ended adapter ligation method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location, thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00108]** Following collection of the released identifier oligonucleotide, a two-ended adapter ligation method can further comprise: (4) ligating to the released identifier oligonucleotide collected in step (3) at least one nucleic acid adapter.

**[00109]** The nucleic acid adapter can comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", the nucleic acid adapter would comprise a nucleic acid sequence that corresponds to "ROI #1".

**[00110]** The nucleic acid adapter can also comprise a unique molecular identifier.

**[00111]** The nucleic acid adapter can also comprise a first amplification primer binding site. In other aspects, the nucleic acid adapter can also comprise a second amplification primer binding site.

**[00112]** In some aspects, the nucleic acid adapter can also comprise a constant nucleic acid sequence to minimize ligation bias caused by differences in sequences of particular identifier oligonucleotides.

**[00113]** The nucleic acid adapter can be a partially double-stranded nucleic acid molecule. In aspects in which the nucleic acid adapter is partially double-stranded, the nucleic acid adapter comprises a double-stranded annealed region, a first single-stranded mismatched region, and a second single-stranded mismatched region. The first single-stranded mismatched region and the second single stranded mismatched region can be present on opposing sides of the double-stranded annealed region.

**[00114]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of the nucleic acid adapter.

**[00115]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a constant nucleic acid sequence to minimize ligation bias, the constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of the nucleic acid adapter.

**[00116]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a unique molecular identifier, the unique molecular identifier can be present in at least one of the first or second single-stranded mismatched regions of the nucleic acid adapter.

**[00117]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a first and a second amplification primer binding site, the first amplification primer binding site can be present in the first single-stranded mismatched region of the nucleic acid adapter and the second amplification primer binding site can be present in the second single-stranded mismatched region of the nucleic acid adapter.

**[00118]** After ligation of the at least one nucleic acid adapter, a two-ended adapter ligation method can further comprise: (5) amplifying the ligation product produced in step (4) using amplification primers that bind to the first and second amplification primer binding sites; and (6) identifying the released oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the sample.

**[00119]** A two-ended adapter ligation method of the present disclosure can further comprise, prior to step (4), performing an end repair reaction using methods known in the art. The method can also further comprise, prior to step (4), performing a tailing reaction to attach a single nucleotide

overhang to the 3' ends of the identifier oligonucleotide using methods known in the art. In aspects, the end repair reaction and the tailing reaction can be performed sequentially or concurrently.

**[00120]** In preferred aspects of a two-ended adapter ligation method, a nucleic acid adapter is ligated to both ends of the released and collected identifier oligonucleotide.

**[00121]** In other aspects of a two-ended adapter ligation method, at least one of the amplification primers used in step (5) to amplify the ligation product produced in step (4) comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", at least one of the amplification primers would comprise a nucleic acid sequence that corresponds to "ROI #1".

**[00122]** Figure 1 shows a schematic of a preferred aspect of a two-ended adapter ligation method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising an antibody that binds to a target protein. In the upper left panel, the probe binds to the target protein. In the upper right panel, a UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target protein bound to the target binding domain. In the bottom panel, a nucleic acid adapter is ligated to both ends of the identifier oligonucleotide. In this non-limiting example, the nucleic acid adapter is partially double-stranded and comprises a double-stranded annealed region, a first single-stranded mismatched region, and a second single-stranded mismatched region. Present in the double-stranded annealed region is a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. Present in the first single-stranded mismatched region is a first amplification primer binding site. Present in the second single-stranded mismatched region is a unique molecular identifier and a second amplification primer binding site. Following ligation of the nucleic acid adapters to the identifier oligonucleotide, the product is amplified using amplification primers that bind the first and the second amplification primer binding sites and sequenced to identify the target protein bound by the probe.

**[00123]** In one aspect, the present disclosure provides a composition of an identifier oligonucleotide dually ligated to two nucleic acid adapters for spatially detecting at least one target analyte in a

sample. An identifier oligonucleotide dually ligated to two nucleic acid adapters comprises an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which is capable of identifying a target analyte in a sample. Each end of the identifier oligonucleotide is attached to a nucleic acid adapter molecule, wherein the nucleic acid adapter molecule is partially double-stranded and comprises a double-stranded annealed region, a first single-stranded mismatched region and a second single-stranded mismatched region. The first single-stranded mismatched region and the second single stranded mismatched region are present on opposing sides of the double-stranded annealed region. The double-stranded mismatch region comprises a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence which is capable of identifying a specific location of a sample. The first single-stranded mismatched region comprises a first amplification primer binding site. The second single-stranded mismatched region comprises a second amplification primer binding site and a nucleic acid sequence comprising a unique molecular identifier. A schematic of an identifier oligonucleotide dually ligated to two nucleic acid adapters is shown in the bottom panel of Figure 1.

**[00124]** In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a “one-ended adapter ligation method”.

**[00125]** A one-ended adapter ligation method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a first amplification primer binding site. In some aspects, the identifier oligonucleotide also comprises at least one 3' end with a single nucleotide overhang.

**[00126]** Following contacting the at least one target analyte with the at least one probe, a one-ended adapter ligation method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00127]** Following release of the identifier oligonucleotide, a one-ended adapter ligation method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00128]** Following collection of the released identifier oligonucleotide, a one-ended adapter ligation method can further comprise: (4) ligating to the released oligonucleotide collected in step (3) at least one nucleic acid adapter;

**[00129]** The nucleic acid adapter can comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", the nucleic acid adapter would comprise a nucleic acid sequence that corresponds to "ROI #1". The nucleic acid adapter can also comprise a unique molecular identifier. The nucleic acid adapter can also comprise a second amplification primer binding site.

**[00130]** In some aspects, the nucleic acid adapter can also comprise a constant nucleic acid sequence to minimize ligation bias caused by differences in sequences of particular identifier oligonucleotides. The constant nucleic acid sequence can comprise a cleavable moiety. The cleavable moiety can be enzymatically cleavable. In a non-limiting example, the enzymatically cleavable moiety can be a USER sequence, wherein the USER sequence comprises the sequence GUGUATUG.

**[00131]** The nucleic acid adapter can comprise any combination of the features described above.

**[00132]** The nucleic acid adapter can be a partially double-stranded nucleic acid molecule. In aspects in which the nucleic acid adapter is partially double-stranded, the nucleic acid adapter comprises a double-stranded annealed region and a single-stranded mismatched region.

**[00133]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released can be present in the double-stranded annealed region of the nucleic acid adapter.

**[00134]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a constant nucleic acid sequence to minimize ligation bias, the constant nucleic acid sequence to minimize ligation bias can be present in the double-stranded annealed region of the nucleic acid adapter.

**[00135]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a unique molecular identifier, the unique molecular identifier can be present in the single-stranded mismatched region.

**[00136]** In aspects in which the nucleic acid adapter is partially double-stranded and comprises a second amplification primer binding site, the second amplification primer binding site can be present in the single-stranded mismatched region of the nucleic acid adapter.

**[00137]** After ligation of the at least one nucleic acid adapter, a two-ended adapter ligation method can further comprise: (5) amplifying the ligation product produced in step (4); and (6) identifying the released oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the sample.

**[00138]** In other aspects of a one-ended adapter ligation method of the present disclosure, at least one of the amplification primers used in step (5) to amplify the ligation product produced in step (4) comprises a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated "ROI #1", at least one of the amplification primers would comprise a nucleic acid sequence that corresponds to "ROI #1".

**[00139]** Figure 2 shows a schematic of a preferred aspect of a one-ended adapter ligation method of the present disclosure. In this aspect, the probe comprises a target binding domain that is an antibody that binds to a target protein. In the upper left panel, the probe binds to the target

protein. In upper right panel, the UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target protein and a first amplification primer binding site. In this non-limiting example, the identifier oligonucleotide is double-stranded with one strand that comprises three separate nucleic acid molecules. The identifier oligonucleotide also comprises one 3' end with a single nucleotide overhang.

**[00140]** In the bottom panel of Figure 2, a nucleic acid adapter is ligated to the end of the identifier oligonucleotide that comprises the 3' single nucleotide overhang. In this non-limiting example, the nucleic acid adapter is partially double-stranded and comprises a double-stranded annealed region and a single-stranded mismatched region. Present in the double-stranded annealed region is a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. Present in the single-stranded mismatched region is a unique molecular identifier and a second amplification primer binding site. Following ligation of the nucleic acid adapter to the identifier oligonucleotide, the product is amplified using amplification primers that bind to the first and the second amplification primer binding sites and sequenced to identify the target protein bound by the probe.

**[00141]** In one aspect, the present disclosure provides a composition of an identifier oligonucleotide ligated to one nucleic acid adapter for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide ligated to one nucleic acid adapter comprises an identifier oligonucleotide, wherein the identifier oligonucleotide comprises a unique nucleic acid sequence which is capable of identifying a target analyte in a sample and a first amplification primer binding site. One end of the identifier oligonucleotide is attached to a nucleic acid adapter molecule, wherein the nucleic acid adapter molecule is partially double-stranded and comprises a double-stranded annealed region and a single-stranded mismatched region and a second single-stranded mismatched region. The double-stranded mismatch region comprises a constant nucleic acid sequence to minimize ligation bias and a nucleic acid sequence nucleic acid sequence which is capable of identifying a specific location of a sample. The single-stranded mismatched region comprises a second amplification primer binding site. A schematic of an identifier oligonucleotide ligated to one nucleic acid adapter is shown in the bottom panel of Figure 2.

**[00142]** In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a “templated-primer extension method”.

**[00143]** A templated-primer extension method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a first amplification primer binding site.

**[00144]** Following contacting the at least one target analyte with the at least one probe, a templated-primer extension method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00145]** Following release of the identifier oligonucleotide, a templated-primer extension method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00146]** Following collection of the released identifier oligonucleotide, a templated-primer extension method can further comprise: (4) hybridizing to the released identifier oligonucleotide collected in step (3) a single stranded nucleic acid template.

**[00147]** The single stranded nucleic acid template can comprise a region complementary to the unique nucleic acid sequence of the identifier oligonucleotide, thereby allowing for the hybridization of the single stranded nucleic acid template and the collected identifier oligonucleotide.

**[00148]** The single stranded nucleic acid template can also comprise a nucleic acid sequence comprising a unique molecular identifier.

**[00149]** The single stranded nucleic acid template can also comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released.

**[00150]** The single stranded nucleic acid template can also comprise a nucleic acid sequence that is complementary to a second amplification primer binding site.

**[00151]** The single stranded nucleic acid template can comprise any combination of the features described above.

**[00152]** Following hybridization of the identifier oligonucleotide to the single stranded nucleic acid template, a templated-primer extension method can further comprise: (5) extending the identifier oligonucleotide of step (4) to form an extension product complementary to the single stranded nucleic acid template, wherein the extension product comprises the identifier oligonucleotide and the sequence complementary to the single stranded nucleic acid template; (6) amplifying the extension product of step (6) using amplification primers that hybridize to the first and second amplification primer binding sites; and (7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the sample.

**[00153]** In some aspects, the single stranded nucleic acid template can comprise an affinity molecule. In aspects in which the single stranded nucleic acid template comprises an affinity molecule, a templated-primer extension method can further comprise an affinity purification step between steps (4) and (5).

**[00154]** Figure 3 shows a schematic of a preferred aspect of a templated-primer extension method of the present disclosure. In this aspect, the probe comprises a target binding domain that is an antibody that binds to a target protein. In the upper left panel, the probe binds to the target protein. In upper right panel, the UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The

identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target protein and a first amplification primer binding site. In the lower right panel, the identifier oligonucleotide is hybridized to a single stranded nucleic acid template. In this non-limiting example, the single-stranded nucleic acid template comprises an affinity molecule, a nucleic acid sequence complementary to the unique nucleic acid sequence of the identifier oligonucleotide, a first unique molecular identifier, and a sequence complementary to a second amplification primer binding site. The identifier oligonucleotide is extended to form an extension product complementary to the single stranded nucleic acid template. As shown in the lower left panel, the extension product comprises the identifier oligonucleotide, the nucleic acid sequence complementary to the first unique molecular identifier, and the second amplification primer binding site. Following the extension reaction, the primer extension product is amplified using amplification primers that bind to the first and the second amplification primer binding sites. In this non-limiting example, one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. The amplified product is then sequenced to identify the target protein bound by the probe.

**[00155]** In one aspect, the present disclosure provides a composition of a template-extended identifier oligonucleotide for spatially detecting at least one target analyte in a sample. The template-extended identifier oligonucleotide comprises a first flow cell adapter sequence suitable for sequencing, followed by a first unique molecular identifier, followed by an identifier oligonucleotide, followed by a second unique molecular identifier, followed by a second amplification primer binding site, followed by a third unique molecular identifier, followed by a second flow cell adapter sequence suitable for sequencing. The identifier oligonucleotide comprises a first amplification primer binding site and a unique nucleic acid sequence which is capable of identifying a target analyte in a sample. A schematic of a template-extended identifier oligonucleotide is shown in the bottom panel of Figure 4.

**[00156]** In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a “short probe hybridization method”.

**[00157]** A short probe hybridization method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at

least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain.

**[00158]** Following contacting the at least one target analyte with the at least one probe, a short probe hybridization method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00159]** Following release of the identifier oligonucleotide, a short probe hybridization method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00160]** Following collection of the released identifier oligonucleotide, a short probe hybridization method can further comprise: (4) hybridizing to the release identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe.

**[00161]** The first or the second nucleic acid probe can comprise a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The first or the second nucleic acid probe can also comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. The first or the second nucleic acid probe can also comprise a nucleic acid sequence comprising unique molecular identifier. The first nucleic acid probe can comprise a first amplification primer binding site. The second nucleic acid probe can comprise a second amplification primer binding site.

**[00162]** The first or the second nucleic acid probe can comprise any combination of the features described above. In a preferred aspect depicted in Figure 5, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released is located 5' to the first amplification primer binding site and the unique molecular identifier is located 3' to the second amplification primer binding site.

**[00163]** In another preferred aspect depicted in Figure 6, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and the unique molecular identifier are located 5' to the first amplification primer binding site.

**[00164]** In another preferred aspect depicted in Figure 7, the first nucleic acid probe comprises a first amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and the unique molecular identifier are located 3' to the second amplification primer binding site.

**[00165]** In another preferred aspect depicted in Figure 15, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence comprising a unique molecular identifier is located 3' to the first amplification binding site.

**[00166]** In another preferred aspect depicted in Figure 17, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence comprising a unique molecular identifier is located 5' to the first amplification binding site.

**[00167]** The first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Alternatively, the first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping.

**[00168]** Following hybridization of the first and the second nucleic acid probe to the identifier oligonucleotide, a short probe hybridization method can further comprise: (5) in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping, ligating the first and the second nucleic acid probes together, for example, by performing a nick repair reaction. Alternatively, in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping, the method comprises ligating the first and the second nucleic acid probes together, for example, by performing a gap extension reaction and a nick repair reaction, such that the first and the second nucleic acid probes are ligated together.

**[00169]** Following ligation of the first and the second nucleic acid probe, a short probe hybridization method can further comprise: (6) amplifying the ligation product produced in step

(5) using amplification primers that hybridize to the first and second amplification primer binding sites; and (7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the sample.

**[00170]** In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-short nucleic acid probe complex for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide-short nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target analyte in a sample. The first nucleic acid probe comprises a first amplification primer binding site, followed by a unique nucleic acid sequence capable of identifying a specific location in a sample, followed by a region complementary to the identifier oligonucleotide. The second nucleic acid probe comprises a second amplification primer binding site, followed by a nucleic acid sequence comprising a unique molecular identifier, followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 5.

**[00171]** In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-short nucleic acid probe complex for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide-short nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target analyte in a sample. The first nucleic acid probe comprises a first amplification primer binding site, followed by a nucleic acid sequence comprising a unique molecular identifier, followed by a region complementary to the identifier oligonucleotide, wherein the nucleic acid sequence comprising a unique molecular identifier is located 3' to the first amplification primer binding site. The second nucleic acid probe comprises a second amplification primer binding site followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 15.

**[00172]** In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-short nucleic acid probe complex for spatially detecting at least one target

analyte in a sample. An identifier oligonucleotide-short nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target analyte in a sample. The first nucleic acid probe comprises a first amplification primer binding site, followed by a nucleic acid sequence comprising a unique molecular identifier, followed by a region complementary to the identifier oligonucleotide, wherein the nucleic acid sequence comprising a unique molecular identifier is located 5' to the first amplification primer binding site. The second nucleic acid probe comprises a second amplification primer binding site followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 17.

**[00173]** Figure 16 shows a schematic overview of an exemplary short probe hybridization method of the present disclosure. First at least one target analyte in a sample is contacted with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target domain.

**[00174]** Following contacting the at least one target analyte with the at least one probe, a force is then provided to a location of the sample sufficient to release the identifier oligonucleotide. The identifier oligonucleotide is collected following release, as shown in the top panel of Figure 16.

**[00175]** As shown in the second panel from the top of Figure 16, the released identifier oligonucleotide is then hybridized to a first nucleic acid probe and a second nucleic acid probe. In this non-limiting example, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The nucleic acid sequence comprising the unique molecular identifier is located 3' to the first amplification primer binding site. The second nucleic acid probe comprises a second amplification primer binding site and a

nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this non-limiting example, the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Following hybridization to the identifier oligonucleotide, the first and second probe are ligated together, for example, by performing a nick repair reaction.

**[00176]** Following ligation of the first and second nucleic acid probes, the ligation product is amplified via PCR using amplification primers that hybridize to the first and second amplification primer binding sites. As shown in the second panel from the bottom of Figure 16, the amplification primer that hybridizes to the second amplification primer binding site comprises a first flow cell binding site, a first nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the second amplification primer binding site. The amplification primer that hybridizes to the first amplification primer binding site comprises a second flow cell binding site, a second nucleic acid sequence which identifies the specific location from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the first amplification primer binding site. The PCR product shown in the bottom panel of Figure 16 is then sequenced to identify the released oligonucleotide, thereby spatially detecting the at least one target analyte in the sample.

**[00177]** Figure 18 shows a schematic overview of an exemplary short probe hybridization method of the present disclosure. First, at least one target analyte in a sample is contacted with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target domain.

**[00178]** Following contacting the at least one target analyte with the at least one probe, a force is then provided to a location of the sample sufficient to release the identifier oligonucleotide. The identifier oligonucleotide is collected following release, as shown in the top panel of Figure 18.

**[00179]** As shown in the second panel from the top of Figure 18, the released identifier oligonucleotide is then hybridized to a first nucleic acid probe and a second nucleic acid probe. In this non-limiting example, the first nucleic acid probe comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The nucleic acid sequence comprising the unique molecular identifier is located 5' to the first amplification primer binding site. The second nucleic acid probe comprises a second amplification primer binding site and a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. In this non-limiting example, the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Following hybridization to the identifier oligonucleotide, the first and second probe are ligated together, for example, by performing a nick repair reaction.

**[00180]** Following ligation of the first and second nucleic acid probes, the ligation product is amplified via PCR using amplification primers that hybridize to the first and second amplification primer binding sites. As shown in the second panel from the bottom of Figure 18, the amplification primer that hybridizes to the second amplification primer binding site comprises a first flow cell binding site, a first nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the second amplification primer binding site. The amplification primer that hybridizes to the first amplification primer binding site comprises a second flow cell binding site, a second nucleic acid sequence which identifies the specific location from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the first amplification primer binding site. The PCR product shown in the bottom panel Figure 18 is then sequenced to identify the released oligonucleotide, thereby spatially detecting the at least one target analyte in the sample.

**[00181]** In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a "long probe hybridization method".

**[00182]** A long probe hybridization method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at

least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain.

**[00183]** Following contacting the at least one target analyte with the at least one probe, a long probe hybridization method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00184]** Following release of the identifier oligonucleotide, a long probe hybridization method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00185]** Following collection of the released identifier oligonucleotide(s), a long probe hybridization method can further comprise: (4) hybridizing to the released identifier oligonucleotide a first nucleic acid probe and a second nucleic acid probe.

**[00186]** The first or the second nucleic acid probe can comprise a nucleic acid sequence complementary to a portion of the identifier oligonucleotide. The first or the second nucleic acid probe can also comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released.

**[00187]** The first or the second nucleic acid probe can also comprise a first unique molecular identifier. The first or the second nucleic acid probe can also comprise a second unique molecular identifier. The first or the second nucleic acid probe can also comprise a third unique molecular identifier.

**[00188]** The first nucleic acid probe can comprise a first amplification primer binding site.

**[00189]** The first nucleic acid probe can also comprise a first flow cell binding site. The second nucleic acid probe can comprise a second flow cell binding site.

**[00190]** The first and the second nucleic acid probes can comprise any combination of the features described above. In a preferred aspect depicted in Figure 8, the first nucleic acid probe comprises a first flow cell binding site, a first unique molecular identifier, a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second flow cell binding site, a second unique molecular identifier, a third unique molecular identifier and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and the first unique molecular identifier are located 5' to the first flow cell binding site and the second and the third unique molecular identifiers are located 3' to the second flow cell binding site.

**[00191]** In another preferred aspect depicted in Figure 9, the first nucleic acid probe comprises a first flow cell binding site, a first unique molecular identifier, a second unique molecular identifier, a first amplification primer binding site, a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In the same preferred aspect, the second nucleic acid probe comprises a second flow cell binding site, a third unique molecular identifier and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the first unique molecular identifier, the second unique molecular identifier and the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released are located 5' to the first flow cell binding site and the third unique molecular identifier is located 3' to the second flow cell binding site.

**[00192]** In another preferred aspect depicted in Figure 10, the first nucleic acid probe comprises a first flow cell binding site, a first unique molecular identifier, a first amplification primer binding site and a nucleic acid sequence complementary to the identifier oligonucleotide. In this same preferred aspect, the second nucleic acid probe comprises a second flow cell binding site, a second unique molecular identifier, a third unique molecular identifier, a nucleic acid sequence

which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the identifier oligonucleotide. In this preferred aspect, the first unique molecular identifier is located 5' to the first flow cell binding site and the second unique molecular identifier, the third unique molecular identifier, and the nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released are located 3' to the second amplification primer binding site.

**[00193]** The first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping. Alternatively, the first nucleic acid probe and the second nucleic acid probe can hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping.

**[00194]** Following hybridization of the first and the second nucleic acid probe to the identifier oligonucleotide, a long probe hybridization method can further comprise: (5) in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are adjacent and are not overlapping, performing a nick repair reaction such that the first and the second nucleic acid probes are ligated together. Alternatively, in the aspect in which the first and the second nucleic acid probe hybridize to the identifier oligonucleotide such that the first and the second nucleic acid probes are not adjacent and are not overlapping, the method comprises performing a gap extension and a nick repair reaction such that the first and the second nucleic acid probes are ligated together.

**[00195]** The name method can further comprise: (6) amplifying the ligation product produced in step (5) using amplification primers that hybridize to the first and second amplification primer binding sites; and (7) identifying the released identifier oligonucleotide by sequencing the amplified products produced in step (6), thereby spatially detecting the at least one target analyte in the sample.

**[00196]** In one aspect, the present disclosure provides a composition of an identifier oligonucleotide-long nucleic acid probe complex for spatially detecting at least one target analyte in a sample. An identifier oligonucleotide-long nucleic acid probe complex comprises an identifier oligonucleotide hybridized to a first nucleic acid probe and a second nucleic acid probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a target

analyte in a sample. The first nucleic acid probe comprises a first flow cell binding site suitable for sequencing, followed by a first unique molecular identifier, followed by a first amplification primer binding site, followed by a unique nucleic acid sequence capable of identifying a specific location in a sample, followed by a region complementary to the identifier oligonucleotide. The second nucleic acid probe comprises a second flow cell binding site, followed by a second unique molecular identifier, followed by a third unique molecular identifier, followed by a region complementary to the identifier oligonucleotide. A schematic of an identifier oligonucleotide-short nucleic acid probe complex is depicted in Figure 8.

**[00197]** In another aspect, the present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a “direct PCR method”.

**[00198]** A direct PCR method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a first amplification primer binding site, a second amplification primer binding site, or a unique molecular identifier. The identifier oligonucleotide can comprise any combination of these features. Any of these features can also be flanked by regions comprising constant nucleic acid sequences of about 1 nucleotide to about 10 nucleotides.

**[00199]** Following contacting the at least one analyte with the at least one probe, a direct PCR method further comprises: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00200]** A direct PCR method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00201]** Following release of the identifier oligonucleotide, a direct PCR method can further comprise: (4) amplifying the released identifier oligonucleotide using a first amplification primer capable of binding to the first amplification primer binding site and a second amplification primer capable of binding to the second amplification primer binding site. In some aspects, at least one of the amplification primers comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. For example, if the identifier oligonucleotide was released from location of the sample designated “ROI #1”, at least one of the amplification primers would comprise a nucleic acid sequence that corresponds to “ROI #1”. In still other aspects, at least one of the amplification primers comprises a unique molecular identifier.

**[00202]** Following amplification, a direct PCR method of the present disclosure can further comprise: (5) identifying the released oligonucleotide by sequencing the amplified products produced in step (5), thereby spatially detecting the at least one target analyte in the sample.

**[00203]** Figure 11 shows a schematic of a preferred aspect of a direct PCR method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising a nucleic acid sequence that is complementary to a target nucleic acid. In the upper panel, the probe hybridizes to the target nucleic acid. In the lower panel, a UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a first amplification primer binding site, a second amplification primer binding site, a unique molecular identifier, and a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. Located between these four features are constant spacer regions that are 3 nucleotides in length. The identifier oligonucleotide is double-stranded and comprises a strand that comprises 3 separate nucleic acid molecules. After release, the identifier oligonucleotide is amplified using a first amplification primer that hybridizes to the first amplification primer binding site and

comprises a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a second amplification primer that hybridizes to the second amplification primer binding site. The amplified product is then sequenced to identify the target nucleic acid bound by the probe.

**[00204]** Figure 19 shows a schematic of a preferred aspect of a direct PCR method of the present disclosure. In this aspect, the identifier oligonucleotide comprises a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain and a second amplification primer binding site, as shown in the top panel of Figure 19. The identifier oligonucleotide is amplified using amplification primers that hybridize to the first and second amplification primer binding sites. As shown in the middle panel of Figure 19, the amplification primer that hybridizes to the second amplification primer binding site comprises a first flow cell binding site, a first nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the second amplification primer binding site. The amplification primer that hybridizes to the first amplification primer binding site comprises a second flow cell binding site, a second nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a nucleic acid sequence complementary to the first amplification primer binding site. The PCR product shown in the bottom panel of Figure 19 is sequenced to identify the released oligonucleotide, thereby spatially detecting the at least one target analyte in the sample.

**[00205]** In one aspect, the present disclosure provides a composition of a direct-PCR compatible identifier oligonucleotide for spatially detecting at least one target analyte in a sample. A direct-PCR compatible identifier oligonucleotide comprises a first amplification primer binding site, followed a unique nucleic acid sequence which is capable of identifying a target analyte in a sample, followed by a unique molecular identifier, followed by a second amplification primer binding site. A schematic of a direct-PCR compatible identifier oligonucleotide is depicted in the lower panel of Figure 11.

**[00206]** In one aspect, the present disclosure provides a composition of a direct-PCR compatible identifier oligonucleotide for spatially detecting at least one target analyte in a sample. A direct-PCR compatible identifier oligonucleotide comprises a first amplification primer binding site,

followed a nucleic acid sequence comprising a unique molecular identifier, followed by a unique nucleic acid sequence which is capable of identifying a target analyte in a sample, followed by a second amplification primer binding site. A schematic of a direct-PCR compatible identifier oligonucleotide is shown in the top panel of Figure 19.

**[00207]** In some aspects of the methods of the present disclosure at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten, or at least eleven, or at least twelve, or at least thirteen, or at least fourteen, or at least fifteen, or at least sixteen, or at least seventeen, or at least eighteen, or at least nineteen, or at least twenty, or at least thirty, or at least forty, or at least fifty, or at least sixty, or at least seventy, or at least eighty, or at least ninety, or at least one hundred probes can be used to a single target analyte. As used herein, the term “tiling” is used to describe when more than one probe of the present disclosure is bound to a target analyte. The top panel of Figure 27 shows the tiling of probes onto a target RNA. Tiling multiple probes onto a target analyte means that each target analyte will be individually detected multiple times, increasing the overall accuracy of the measurement. In a non-limiting example, as shown in the bottom panel of Figure 27, in the case where 10 probes are tiled onto a single target RNA, one of the probes may be incorrectly detected too many times (outlier high count probe), while another probe may be incorrectly detected too few times (outlier low count probe). However, the other 8 probes may be detected at a similar level, indicating that the two outliers should be discarded during analysis and the signals from the 8 probes used to generate a more accurate measurement of the abundance of the target RNA.

**[00208]** The present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as an “enzyme free method”.

**[00209]** An enzyme free method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest.

The probe further comprises an identifier oligonucleotide. An identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. An identifier oligonucleotide can also comprise a capture probe binding site.

**[00210]** Following contacting at least one target analyte with at least one probe, an enzyme free method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00211]** Following release of an identifier oligonucleotide, an enzyme free method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00212]** Following collection of a released identifier oligonucleotide, an enzyme free method can further comprise: (4) hybridizing to a released identifier oligonucleotide a capture probe.

**[00213]** A capture probe can comprise a region complementary to the capture probe binding site. A capture probe can also comprise an affinity molecule.

**[00214]** Following hybridization of a capture probe, an enzyme free method can further comprise: (5) Identifying a released identifier oligonucleotide by sequencing the hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[00215]** A hybridized product produced in step (4) can be sequenced using an enzyme free method of sequencing. Enzyme-free methods of sequencing have been described in, e.g., US2014946386 and US15/819,151, each of which is incorporated herein by reference in its entirety.

**[00216]** Figure 12 shows a schematic of a preferred aspect of an enzyme free method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising a nucleic acid sequence that is complementary to a target nucleic acid. In the top panel, the probe

hybridizes to the target nucleic acid. In the middle panel, a UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target nucleic acid bound to the target binding domain and a capture probe binding site. After release, the identifier oligonucleotide is hybridized to a capture probe, as depicted in the bottom panel. The capture probe comprises a nucleic acid sequence complementary to the capture probe binding site and an affinity molecule. The hybridized product is then sequenced using enzyme free sequencing methods to identify the target nucleic acid bound by the probe.

**[00217]** In one aspect, the present disclosure provides a composition of a hybridized identifier oligonucleotide-capture probe complex for spatially detecting at least one target analyte in a sample. A hybridized identifier oligonucleotide-capture probe complex comprises an identifier oligonucleotide hybridized to a capture probe. An identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a specific target analyte in a sample and a capture probe binding site. A capture probe comprises an affinity molecule and a region complementary to the capture probe binding site. A schematic of a hybridized identifier oligonucleotide-capture probe complex is depicted in the bottom of panel Figure 12.

**[00218]** The present disclosure provides compositions and methods for spatially detecting at least one target analyte in a sample using the probes of the present disclosure in a method herein referred to as a “multiplexed enzyme free method”.

**[00219]** A multiplexed enzyme free method of the present disclosure can comprise: (1) contacting at least one target analyte in a sample with at least one probe of the present disclosure. The at least one target analyte can be a target protein or a target nucleic acid. In the aspect that the at least one target analyte is a target protein, the probe can comprise a target binding domain that is a target protein-binding region that can specifically bind to a target protein of interest. In the aspect that the at least one target analyte is a target nucleic acid, the probe can comprise a target nucleic acid-binding region that can directly or indirectly hybridize to a target nucleic acid of interest. The probe further comprises an identifier oligonucleotide. The identifier oligonucleotide can comprise a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain. The identifier oligonucleotide can also comprise a capture probe binding site. The identifier oligonucleotide can also comprise a multiplexing probe binding site.

**[00220]** Following contacting the at least one target analyte with the at least one probe, a multiplexed enzyme free method can further comprise: (2) providing a force to a location of the sample sufficient to release the identifier oligonucleotide. In a non-limiting example, in aspects in which the probe comprises a photo-cleavable linker between the identifier oligonucleotide and the target binding domain, a region-of-interest (ROI) is excited with light of a sufficient wavelength capable of cleaving the photo-cleavable linker.

**[00221]** Following release of the identifier oligonucleotide, a multiplexed enzyme free method can further comprise: (3) collecting the released identifier oligonucleotide. By directing the force only to a specific location in step (2), identifier oligonucleotides are only released from probes within that location and not from probes located outside that location. Thus, identifier oligonucleotides are collected only for probes that are bound to targets within that location in step (3), thereby permitting detection of the identities and quantities of the targets (proteins and/or nucleic acids) located only within that location.

**[00222]** Following collection of the released identifier oligonucleotide, a multiplexed enzyme free method can further comprise: (4) hybridizing to the released identifier oligonucleotide a capture probe and a multiplexing probe.

**[00223]** A capture probe can comprise a region complementary to the capture probe binding site. A capture probe can also comprise an affinity molecule.

**[00224]** A multiplexing probe can comprise a region complementary to the multiplexing probe binding site. A multiplexing probe can also comprise a nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a region complementary to the multiplexing probe binding site.

**[00225]** Following hybridization of a capture probe and a multiplexing probe, a multiplexed enzyme free method can further comprise: (5) Identifying the released identifier oligonucleotide by sequencing the hybridized product produced in step (4), thereby spatially detecting the at least one target analyte in the at least one cell in a tissue sample.

**[00226]** A hybridized product produced in step (4) can be sequenced using an enzyme free method of sequencing. Enzyme-free methods of sequencing have been described in, e.g., US2014946386 and US15/819,151, each of which is incorporated herein by reference in its entirety.

**[00227]** Figure 13 shows a schematic of a preferred aspect of a multiplexed enzyme free method of the present disclosure. In this aspect, the probe comprises a target binding domain comprising a nucleic acid sequence that is complementary to a target nucleic acid. In the top panel, the probe hybridizes to the target nucleic acid. In the middle panel, a UV photo-cleavable linker located between the target binding domain and the identifier oligonucleotide is cleaved, releasing the identifier oligonucleotide. The identifier oligonucleotide comprises a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain, a capture probe binding site, and a multiplexing probe binding site, as shown in the middle panel. After release, the identifier oligonucleotide is hybridized to a capture probe and a multiplexing probe as shown in the lower panel. The capture probe comprises a nucleic acid sequence complementary to the capture probe binding site and an affinity molecule. The multiplexing probe comprises a nucleic acid sequence complementary to the multiplexing probe binding site and a nucleic acid sequence which identifies the specific location of a sample from which the identifier oligonucleotide was released. The hybridized product is then sequenced using enzyme free sequencing methods to identify the target nucleic acid bound by the probe.

**[00228]** In one aspect, the present disclosure provides a composition of a hybridized identifier oligonucleotide-capture probe-multiplex probe complex for spatially detecting at least one target analyte in a sample. A hybridized identifier oligonucleotide-capture probe-multiplex probe complex comprises an identifier oligonucleotide hybridized to a capture probe and a multiplexing probe. The identifier oligonucleotide comprises a unique nucleic acid sequence capable of identifying a specific target analyte in a sample, a capture probe binding site and a multiplexing probe binding site. The capture probe comprises an affinity molecule and a region complementary to the capture probe binding site. The multiplexing probe comprises a nucleic acid sequence which identifies the specific location of a sample from which the identifier oligonucleotide was released and a region complementary to the multiplexing probe binding site. A schematic of a hybridized identifier oligonucleotide-capture probe-multiplex probe complex is depicted in the bottom panel of Figure 13.

**[00229]** Figure 20 is an exemplary schematic of overview of the methods of the present disclosure. First, a sample on a microscope slide is contacted with a plurality of probes of the present disclosure (step 1 in Figure 20). The slide is then imaged and particular regions of interest (ROIs) are selected (step 2 in Figure 20). A specific ROI is then illuminated by UV light

to release identifier oligonucleotides from probes bound within the ROI. The released identifier oligonucleotides are then collected via aspiration with a microcapillary. Following aspiration, the identifier oligonucleotides are transferred to a particular well within a 96 well plate. Steps 4 and 5 are then repeated for each ROI identified in step 2. After all ROIs have been illuminated and all released identifier oligonucleotides collected, the identifier oligonucleotides are sequenced using next generation sequencing methods to spatially detect at least one target analyte in the sample.

**[00230]** As described in the preceding, the present disclosure provides probes for the compositions and methods of spatially detecting at least one target analyte in a sample. The present disclosure provides probes comprising a target binding domain and an identifier oligonucleotide. The target binding domain is a region of the probe that specifically binds to at least one target analyte in a sample.

**[00231]** Probes of the present disclosure can be used for spatially detecting a target nucleic acid. In this aspect, the target binding domain can be a target nucleic acid-binding region. The target nucleic acid-binding region is preferably at least 15 nucleotides in length, and more preferably is at least 20 nucleotides in length. In specific aspects, the target nucleic acid-binding region is approximately 10 to 500, 20 to 400, 25, 30 to 300, 35, 40 to 200, or 50 to 100 nucleotides in length. Probes and methods for binding and identifying a target nucleic acid have been described in, *e.g.*, US2003/0013091, US2007/0166708, US2010/0015607, US2010/0261026, US2010/0262374, US2010/0112710, US2010/0047924, and US2014/0371088, each of which is incorporated herein by reference in its entirety.

**[00232]** The target nucleic acid-binding region can directly hybridize to a target nucleic acid present in a sample. Alternatively, the probes of the present disclosure can indirectly hybridize to a target nucleic acid present in a sample (via an intermediary oligonucleotide). Figure 14 illustrates a probe (or composition) of this aspect. The probe includes a target nucleic-acid binding domain which binds to a synthetic oligonucleotide (the intermediary oligonucleotide) that in turn binds to a target nucleic acid in a biological sample. It could be said that the intermediary oligonucleotide is a probe, as defined herein, since it comprises a nucleic acid backbone and is capable of binding a target nucleic acid. In these aspects, a probe's target nucleic acid-binding region hybridizes to a region of an intermediary oligonucleotide (*i.e.*, a synthetic oligonucleotide) which is different from the target nucleic acid present in a sample.

Thus, the probe's target binding region is independent of the ultimate target nucleic acid in the sample. This allows economical and rapid flexibility in an assay design, as the target (present in a sample)-specific components of the assay are included in inexpensive and widely-available synthetic DNA oligonucleotides rather than the more expensive probes. Such synthetic oligonucleotides are simply designed by including a region that hybridizes to the target nucleic acid present in a sample and a region that hybridizes to a probe. Therefore, a single set of indirectly-binding probes can be used to detect an infinite variety of target nucleic acids (present in a sample) in different experiments simply by replacing the target-specific (synthetic) oligonucleotide portion of the assay.

**[00233]** A target nucleic acid may be DNA or RNA and preferably messenger RNA (mRNA) or miRNA.

**[00234]** Probes of the present disclosure can be used for detecting a target protein. In this aspect, the target binding domain can be a target protein-binding region. A target protein-binding region includes molecules or assemblies that are designed to bind to at least one target protein, at least one target protein surrogate, or both and can, under appropriate conditions, form a molecular complex comprising the probe and the target protein. The target-protein binding region can include an antibody, a peptide, an aptamer, or a peptoid. The antibody can be obtained from a variety of sources, including but not limited to polyclonal antibody, monoclonal antibody, monospecific antibody, recombinantly expressed antibody, humanized antibody, plantibodies, and the like. The terms protein, polypeptide, peptide, and amino acid sequence are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids or synthetic amino acids. The terms also encompass an amino acid polymer that has been modified, for example, by disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. As used herein the term amino acid refers to either natural and/or unnatural or synthetic amino acids, including but not limited to glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. Probes and methods for binding and identifying a target protein have been described, *e.g.*, in US2011/0086774, the contents of which is incorporated herein by reference in its entirety.

**[00235]** An identifier oligonucleotide is a nucleic acid molecule that identifies the target analyte bound to the target binding domain. The identifier oligonucleotide comprises a unique nucleic acid sequence that identifies the target analyte bound to the target binding domain of the probe. In a non-limiting example, a probe with a target binding domain that binds to the protein P53 comprises an identifier oligonucleotide with a unique nucleic acid sequence that corresponds to P53, while a probe with a target binding domain that binds to the protein P97 comprises an identifier oligonucleotide with a unique nucleic acid sequence that corresponds to P97.

**[00236]** An identifier oligonucleotide can be DNA, RNA, or a combination of DNA and RNA.

**[00237]** In some aspects, an identifier oligonucleotide can comprise at least one amplification primer binding site. An amplification primer binding site is a nucleic acid sequence capable of binding to an amplification primer. An amplification primer can be used to amplify the nucleic molecule to which it is bound using methods known in the art, including, but not limited to, polymerase chain reaction (PCR).

**[00238]** In some aspects, an identifier oligonucleotide can comprise at least one unique molecular identifier.

**[00239]** An identifier oligonucleotide can be a single-stranded, a double-stranded, or a partially double-stranded nucleic acid molecule. In the aspects in which an identifier oligonucleotide is double-stranded or partially double-stranded, at least one of the two strands can comprise at least two separate nucleic acid molecules which, without being bound by theory, allows for denaturing of the identifier oligonucleotide at lower temperatures.

**[00240]** An identifier oligonucleotide can also comprise at least one 3' end that comprises a single nucleotide overhang.

**[00241]** An identifier oligonucleotide can also comprise a capture probe binding site. A capture probe binding site is a nucleic acid sequence to which a capture probe can bind.

**[00242]** A capture probe of the present disclosure can comprise a nucleic acid sequence complementary to a capture probe binding site. A capture probe can also comprise an affinity molecule.

**[00243]** An identifier oligonucleotide can also comprise a multiplexing probe binding site. A multiplexing probe binding site is a nucleic acid sequence to which a multiplexing probe can bind.

**[00244]** A multiplexing probe of the present disclosure can comprise a nucleic acid sequence complementary to a multiplexing probe binding site. A multiplexing probe can also comprise a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released.

**[00245]** A probe of the present disclosure can include a region which permits the release of an identifier oligonucleotide following the application of a suitable force. In one non-limited example, the region is a cleavable motif (*e.g.*, a restriction enzyme site or cleavable linker). The cleavable motif allows release of an identifier oligonucleotide from a bound target nucleic acid or protein and the identifier oligonucleotide can then be collected and detected. The region which permits the release of an identifier oligonucleotide can be positioned between the target-binding domain and the identifier oligonucleotide, allowing for the release of the identifier oligonucleotide from the target binding domain. An identifier oligonucleotide is said to be releasable when it can be separated (*i.e.*, cleaved and released) from the remainder of the probe. Examples of cleavable motives include but are not limited to photo-cleavable linkers. Photo-cleavable linkers can be cleaved by light provided by a suitable coherent light source (*e.g.*, a laser and a UV light source) or a suitable incoherent light source (*e.g.*, an arc-lamp and a light-emitting diode (LED)).

**[00246]** In some aspects, the identifier oligonucleotide is collected from a solution proximal to, *e.g.*, at least immediately above or surrounding, the point at which the identifier oligonucleotide is released or the at least one cell. The proximal solution may be collected by aspirating, *e.g.*, via a pipette, a capillary tube, a microarray pin, a flow cell comprising holes, or another suitable aspirating system known in the art or any combination thereof. The capillary tube may comprise an optical device capable of transmitting a light force, *e.g.*, UV light, to the at least one cell. The pipette or a microarray pin may be attached to an array comprising a plurality of pipettes or microarray pins. The proximal solution may comprise an anionic polymer, *e.g.*, dextran sulfate, and/or salmon sperm DNA and/or the collected signal oligonucleotide may be added to a solution comprising an anionic polymer, *e.g.*, dextran sulfate, and/or salmon sperm DNA. Other non-specific blocking agents known in the art in addition to or instead of salmon sperm DNA may be used.

**[00247]** In some aspects, the identifier oligonucleotide is collected from a tissue, at least one cell or proximal to the point at which the identifier oligonucleotide is released via liquid laminar,

turbulent, or transitional flow. The flow may be via a channel, e.g., having 25 to 500  $\mu$ m depth between the tissue and a fluidic device or impermeable barrier placed over the tissue.

**[00248]** In aspects where the target-binding domain of a probe is an antibody, the probe can be prepared using a cysteine bioconjugation method that is stable, site-specific to, preferably, the antibody's hinge-region heavy-chain. This preparation method provides relatively controllable identifier oligonucleotides to antibody stoichiometric ratios. A probe can comprise a plurality (i.e., more than one, e.g., 2, 3, 4, 5, or more) identifier oligonucleotides per antibody. Generally, "heavier" probes, which comprise 3 or 4 identifier oligonucleotides per antibody, are significantly less sensitive than antibodies lacking an identifier oligonucleotide or "lighter" probes, which comprise 1 or 2 identifier oligonucleotides per antibody.

**[00249]** In aspects, probes are provided to a sample at concentrations typically less than that used for immunohistochemistry (IHC) or for in situ hybridization (ISH). Alternately, the concentration may be significantly less than that used for IHC or ISH. For example, the probe concentration may be 2 fold less, 5 fold less, 10 fold less, 20 fold less, 25 fold less, 30 fold less, 50 fold less, 60 fold less, 70 fold less, 80 fold less, 90 fold less, 100 fold less, 200 fold less, 300 fold less, 400 fold less, 500 fold less, 600 fold less, 700 fold less, 800 fold less, 900 fold less, 1000 fold less, 2000 fold less, or less and any number in between. In aspects, probes are provided at a concentration of 100 nM, 70 nM, 60 nM, 50 nM, 40 nM, 30 nM, 20 nM, 10 nM, 9 nM, 8 nM, 7 nM, 6 nM, 5 nM, 4 nM, 3 nM, 2 nM, 1 nM, 0.9 nM, 0.8 nM, 0.7 nM, 0.6 nM, 0.5 nM, 0.4 nM, 0.3 nM, 0.2 nM, 0.1 nM, 0.09 nM, 0.08 nM, 0.07 nM, 0.06 nM, 0.05 nM, 0.04 nM, 0.03 nM, 0.02 nM, 0.01 nM, and less and any concentration in between.

**[00250]** Background noise, during protein detection, can be reduced by performing a negative purification of the intact probe molecule. This can be done by conducting an affinity purification of the antibody or photo-cleavable linker after collection of eluate from a region of interest. Normally, released signal oligonucleotides will not be pulled out of solution. A protein-G or -O mechanism in a pipet tip, tube, or plate can be employed for this step. Such devices and reagents commercially available.

**[00251]** Background noise, during nucleic acid detection, can be reduced by performing a negative purification of the intact probe molecule. This can be done by conducting an affinity purification of the target binding domain or photo-cleavable linker after collection of eluate from a region of interest. Normally, released signal oligonucleotides will not be pulled out of solution.

To assist in the negative purification, a universal purification sequence may be included in a probe, e.g., in the target binding domain.

**[00252]** Protein-targeting probes and nucleic acid-targeting probes may be applied simultaneously as long as conditions allow for binding of both a protein target and a nucleic acid target.

Alternately, protein-targeting probes and nucleic acid-targeting probes may be applied sequentially when conditions allowing for binding of both a protein target and a nucleic acid target are not possible.

**[00253]** A set of probes is synonymous with a composition of probes. A set of probes includes at least one species of probes, *i.e.*, directed to one target. A set of probes preferably includes at least two, e.g., 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more species of probes. A probe set may include one or multiple copies of each species of probe.

**[00254]** A first set of probes only may be applied to a sample. Alternately, a second set (or higher number) of probes may be later applied to the sample. The first set and second (or higher number) may target only nucleic acids, only proteins, or a combination thereof.

**[00255]** In the present disclosure, two or more targets (*i.e.*, proteins, nucleic acids, or a combination thereof) are detected; 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more targets, and any number there between, are detected.

**[00256]** A set of probes may be pre-defined based upon the cell type or tissue type to be targeted. For example, if the tissue is a breast cancer, then the set of probes will include probes directed to proteins relevant to breast cancer cells (e.g., Her2, EGFR, and PR) and/or probes directed to proteins relevant to normal breast tissues. Additionally, the set of probes may be pre-defined based upon developmental status of a cell or tissue to be targeted. Alternately, the set of probes may be pre-defined based upon subcellular localizations of interest, e.g., nucleus, cytoplasm, and membrane. For example, antibodies directed to Foxp3, Histone H3, or P-S6 label the nucleus, antibodies directed to CD3, CD4, PD-1, or CD45RO label the cytoplasm, and antibodies directed to PD-L1 label membranes.

**[00257]** A probe may be chemically synthesized or may be produced biologically using a vector into which a nucleic acid encoding the probe has been cloned.

[00258] Any probe or set of probes described herein may be used in methods and kits of the present disclosure.

[00259] For the herein-described probes, association of a unique nucleic acid sequence to a specific target nucleic acid or target protein is not fixed.

[00260] As described in the preceding, probes of the present disclosure can be used to detect a target nucleic acid or target protein present in any sample, *e.g.*, a biological sample. As will be appreciated by those in the art, the sample may comprise any number of things, including, but not limited to: cells (including both primary cells and cultured cell lines) and tissues (including cultured or explanted). In aspects, a tissue sample (fixed or unfixed) is embedded, serially sectioned, and immobilized onto a microscope slide. As is well known, a pair of serial sections will include at least one cell that is present in both serial sections. Structures and cell types, located on a first serial section will have a similar location on an adjacent serial section. The sample can be cultured cells or dissociated cells (fixed or unfixed) that have been immobilized onto a slide. A sample can be a formalin-fixed paraffin-embedded (FFPE) tissue sample.

[00261] In aspects, a tissue sample is a biopsied tumor or a portion thereof, *i.e.*, a clinically-relevant tissue sample. For example, the tumor may be from a breast cancer. The sample may be an excised lymph node.

[00262] The sample can be obtained from virtually any organism including multicellular organisms, *e.g.*, of the plant, fungus, and animal kingdoms; preferably, the sample is obtained from an animal, *e.g.*, a mammal. Human samples are particularly preferred.

[00263] In some aspects, the probes, compositions, methods, and kits described herein are used in the diagnosis of a condition. As used herein the term diagnose or diagnosis of a condition includes predicting or diagnosing the condition, determining predisposition to the condition, monitoring treatment of the condition, diagnosing a therapeutic response of the disease, and prognosis of the condition, condition progression, and response to particular treatment of the condition. For example, a tissue sample can be assayed according to any of the probes, methods, or kits described herein to determine the presence and/or quantity of markers of a disease or malignant cell type in the sample (relative to the non-diseased condition), thereby diagnosing or staging a disease or a cancer.

[00264] In general, samples attached to a slide can be first imaged using fluorescence (*e.g.*, fluorescent antibodies or fluorescent stains (*e.g.*, DAPI)) to identify morphology, regions of

interest, cell types of interest, and single cells and then expression of proteins and/or nucleic acids can be digitally counted from the sample on the same slide.

**[00265]** Compositions and kits of the present disclosure can include probes and other reagents, for example, buffers and other reagents known in the art to facilitate binding of a protein and/or a nucleic acid in a sample, *i.e.*, for performing hybridization reactions.

**[00266]** A kit also will include instructions for using the components of the kit, including, but not limited to, information necessary to hybridize labeled oligonucleotides to a probe, to hybridize a probe to a target-specific oligonucleotide, to hybridize a target-specific oligonucleotide to a target nucleic acid and/or to hybridize a probe to target protein.

**[00267]** A region of interest may be a tissue type present in a sample, a cell type, a cell, or a subcellular structure within a cell.

**[00268]** Together, a comparison of the identity and abundance of the target proteins and/or target nucleic acids present in a first region of interest (*e.g.*, tissue type, a cell type (including normal and abnormal cells), and a subcellular structure within a cell) and the identity and abundance of the target proteins and/or target nucleic acids present in second region of interest or more regions of interest can be made using the methods of the present disclosure.

**[00269]** As described in the preceding, the products produced by the methods of the present disclosure can be used for nucleic acid amplification. In a preferred aspect, the nucleic acid amplification can be solid-phase nucleic acid amplification. Thus, in further aspects the invention provides a method of solid-phase nucleic acid amplification of template polynucleotide molecules which comprises: preparing a library of template polynucleotide molecules which have common sequences at their 5' and 3' ends using the methods of the present disclosure and carrying out a solid-phase nucleic acid amplification reaction wherein said template polynucleotide molecules are amplified. Compositions and methods for nucleic acid amplification and sequencing have been described in, *e.g.*, US9376678, which is incorporated herein by reference in its entirety.

**[00270]** The term "solid-phase amplification" as used herein refers to any nucleic acid amplification reaction carried out on or in association with a solid support such that all or a portion of the amplified products are immobilized on the solid support as they are formed. In particular, the term encompasses solid-phase polymerase chain reaction (solid-phase PCR),

which is a reaction analogous to standard solution phase PCR, except that one or both of the forward and reverse amplification primers is/are immobilized on the solid support.

**[00271]** Although the invention encompasses "solid-phase" amplification methods in which only one amplification primer is immobilized (the other primer usually being present in free solution), it is preferred for the solid support to be provided with both the forward and the reverse primers immobilized. In practice, there will be a "plurality" of identical forward primers and/or a "plurality" of identical reverse primers immobilized on the solid support, since the PCR process requires an excess of primers to sustain amplification. References herein to forward and reverse primers are to be interpreted accordingly as encompassing a "plurality" of such primers unless the context indicates otherwise.

**[00272]** As will be appreciated by the skilled reader, any given PCR reaction requires at least one type of forward primer and at least one type of reverse primer specific for the template to be amplified. However, in certain aspects the forward and reverse primers may comprise template-specific portions of identical sequence, and may have entirely identical nucleotide sequence and structure (including any non-nucleotide modifications). In other words, it is possible to carry out solid-phase amplification using only one type of primer, and such single-primer methods are encompassed within the scope of the invention. Other aspects may use forward and reverse primers which contain identical template-specific sequences but which differ in some other structural features. For example one type of primer may contain a non-nucleotide modification which is not present in the other.

**[00273]** In other aspects of the invention the forward and reverse primers may contain template-specific portions of different sequence.

**[00274]** Amplification primers for solid-phase PCR are preferably immobilized by covalent attachment to the solid support at or near the 5' end of the primer, leaving the template-specific portion of the primer free for annealing to its cognate template and the 3' hydroxyl group free for primer extension. Any suitable covalent attachment means known in the art may be used for this purpose. The chosen attachment chemistry will depend on the nature of the solid support, and any derivatization or functionalization applied to it. The primer itself may include a moiety, which may be a non-nucleotide chemical modification, to facilitate attachment. In one particularly preferred aspect the primer may include a sulphur-containing nucleophile, such as phosphorothioate or thiophosphate, at the 5' end. In the case of solid-supported polyacrylamide

hydrogels (as described below), this nucleophile will bind to a "C" group present in the hydrogel. The most preferred means of attaching primers and templates to a solid support is via 5' phosphorothioate attachment to a hydrogel comprised of polymerised acrylamide and N-(5-bromoacetamidylpentyl)acrylamide (BRAPA).

**[00275]** The terms "cluster" and "colony" are used interchangeably herein to refer to a discrete site on a solid support comprised of a plurality of identical immobilized nucleic acid strands and a plurality of identical immobilized complementary nucleic acid strands. The term "clustered array" refers to an array formed from such clusters or colonies. In this context the term "array" is not to be understood as requiring an ordered arrangement of clusters.

**[00276]** The invention also encompasses methods of sequencing the amplified nucleic acids generated by solid-phase amplification. Thus, the invention provides a method of nucleic acid sequencing comprising amplifying a library of nucleic acid templates by the methods of the present disclosure described above, using solid-phase amplification as described above to amplify this library on a solid support, and carrying out a nucleic acid sequencing reaction to determine the sequence of the whole or a part of at least one amplified nucleic acid strand produced in the solid-phase amplification reaction.

**[00277]** Sequencing, as referred to herein, can be carried out using any suitable "sequencing-by-synthesis" technique, wherein nucleotides are added successively to a free 3' hydroxyl group, resulting in synthesis of a polynucleotide chain in the 5' to 3' direction. The nature of the nucleotide added is preferably determined after each nucleotide addition.

**[00278]** The initiation point for the sequencing reaction may be provided by annealing of a sequencing primer to a product of the whole genome or solid-phase amplification reaction. In this connection, one or both of the adapters added during formation of the template library may include a nucleotide sequence which permits annealing of a sequencing primer to amplified products derived by whole genome or solid-phase amplification of the template library.

**[00279]** The products of solid-phase amplification reactions wherein both forward and reverse amplification primers are covalently immobilized on the solid surface are so-called "bridged" structures formed by annealing of pairs of immobilized polynucleotide strands and immobilized complementary strands, both strands being attached to the solid support at the 5' end. Arrays comprised of such bridged structures provide inefficient templates for nucleic acid sequencing, since hybridization of a conventional sequencing primer to one of the immobilized strands is not

favorable compared to annealing of this strand to its immobilized complementary strand under standard conditions for hybridization.

**[00280]** In order to provide more suitable templates for nucleic acid sequencing it is preferred to remove substantially all or at least a portion of one of the immobilized strands in the "bridged" structure in order to generate a template which is at least partially single-stranded. The portion of the template which is single-stranded will thus be available for hybridization to a sequencing primer. The process of removing all or a portion of one immobilized strand in a "bridged" double-stranded nucleic acid structure may be referred to herein as "linearization".

**[00281]** Bridged template structures may be linearized by cleavage of one or both strands with a restriction endonuclease or by cleavage of one strand with a nicking endonuclease. Other methods of cleavage can be used as an alternative to restriction enzymes or nicking enzymes, including inter alia chemical cleavage (e.g. cleavage of a diol linkage with periodate), cleavage of abasic sites by cleavage with endonuclease, or by exposure to heat or alkali, cleavage of ribonucleotides incorporated into amplification products otherwise comprised of deoxyribonucleotides, photochemical cleavage or cleavage of a peptide linker.

**[00282]** It will be appreciated that a linearization step may not be essential if the solid-phase amplification reaction is performed with only one primer covalently immobilized and the other in free solution.

**[00283]** In order to generate a linearized template suitable for sequencing it is necessary to remove "unequal" amounts of the complementary strands in the bridged structure formed by amplification so as to leave behind a linearized template for sequencing which is fully or partially single stranded. Most preferably one strand of the bridged structure is substantially or completely removed.

**[00284]** Following the cleavage step, regardless of the method used for cleavage, the product of the cleavage reaction may be subjected to denaturing conditions in order to remove the portion(s) of the cleaved strand(s) that are not attached to the solid support. Suitable denaturing conditions will be apparent to the skilled reader with reference to standard molecular biology protocols.

**[00285]** Denaturation (and subsequent re-annealing of the cleaved strands) results in the production of a sequencing template which is partially or substantially single-stranded. A sequencing reaction may then be initiated by hybridization of a sequencing primer to the single-stranded portion of the template.

**[00286]** Thus, the nucleic acid sequencing reaction may comprise hybridizing a sequencing primer to a single-stranded region of a linearized amplification product, sequentially incorporating one or more nucleotides into a polynucleotide strand complementary to the region of amplified template strand to be sequenced, identifying the base present in one or more of the incorporated nucleotide(s) and thereby determining the sequence of a region of the template strand.

**[00287]** One preferred sequencing method which can be used in accordance with the invention relies on the use of modified nucleotides that can act as chain terminators. Once the modified nucleotide has been incorporated into the growing polynucleotide chain complementary to the region of the template being sequenced there is no free 3'-OH group available to direct further sequence extension and therefore the polymerase cannot add further nucleotides. Once the nature of the base incorporated into the growing chain has been determined, the 3' block may be removed to allow addition of the next successive nucleotide. By ordering the products derived using these modified nucleotides it is possible to deduce the DNA sequence of the DNA template. Such reactions can be done in a single experiment if each of the modified nucleotides has attached a different label, known to correspond to the particular base, to facilitate discrimination between the bases added at each incorporation step. Alternatively, a separate reaction may be carried out containing each of the modified nucleotides separately.

**[00288]** The modified nucleotides may carry a label to facilitate their detection. Preferably this is a fluorescent label. Each nucleotide type may carry a different fluorescent label. However the detectable label need not be a fluorescent label. Any label can be used which allows the detection of an incorporated nucleotide.

**[00289]** One method for detecting fluorescently labelled nucleotides comprises using laser light of a wavelength specific for the labelled nucleotides, or the use of other suitable sources of illumination. The fluorescence from the label on the nucleotide may be detected by a CCD camera or other suitable detection means.

**[00290]** The invention is not intended to be limited to use of the sequencing method outlined above, as essentially any sequencing methodology which relies on successive incorporation of nucleotides into a polynucleotide chain can be used. Suitable alternative techniques include, for example, Pyrosequencing, FISSEQ (fluorescent in situ sequencing), MPSS (massively parallel signature sequencing) and sequencing by ligation-based methods.

**[00291]** In methods of the present disclosure, the unique nucleic acid sequence present in the identifier oligonucleotide of a probe which identifies the target analyte bound to the target binding domain of the probe can comprise between about 5 nucleotides and about 50 nucleotides. Preferably, the sequence comprises between about 20 nucleotides and about 40 nucleotides. Even more preferably, the sequence comprises about 35 nucleotides. In some preferred aspects, the sequence comprises 10 nucleotides.

**[00292]** In methods of the present disclosure, the nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released comprises between about 6 nucleotides and about 15 nucleotides. Preferably, the sequence comprises about 12 nucleotides.

**[00293]** In methods of the present disclosure, an amplification primer binding sites comprises between about 18 nucleotides and about 40 nucleotides. Preferably, an amplification primer binding sites comprises about 32 nucleotides.

**[00294]** In some aspects of the methods of the present disclosure, an amplification primer binding site can comprise an i7 sequence, wherein the i7 sequence comprises the sequence set forth in SEQ ID NO: 1.

**[00295]** In some aspects of the methods of the present disclosure, an amplification primer binding site can comprise an i5 sequence, wherein the i5 sequence comprises the sequence set forth in SEQ ID NO: 2.

**[00296]** In some aspects of the methods of the present disclosure, an amplification primer can comprise a flow cell adapter sequence, wherein the flow cell adapter sequence is suitable for sequencing. Preferably, at least one amplification primer used in the methods of the present disclosure comprises a P5 flow cell adapter sequence, wherein the P5 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 3. Preferably still, at least one amplification primer used in the methods of the present disclosure comprises a P7 flow cell adapter sequence, wherein the P7 flow cell adapter sequence comprises the sequence set forth in SEQ ID NO: 4.

**[00297]** In methods of the present disclosure, a unique molecular identifier can comprise between about 6 nucleotides and about 30 nucleotides. Preferably, a unique molecular identifier can comprise about 15 nucleotides. The terms unique molecular identifier and random molecular tags are used interchangeably herein. Using methods known in that art, unique molecular identifiers are random sequences that can be used to correct for biases in amplification prior to sequencing.

**[00298]** In methods of the present disclosure, a constant nucleic acid sequence to minimize ligation bias comprises between about 1 nucleotide and about 15 nucleotides. Preferably, the constant sequence comprises about 8 nucleotides.

**[00299]** In some aspects, a flow cell binding site can comprise between about 15 to about 40 nucleotides. A flow cell binding site can comprise about 29 nucleotides. A flow cell binding site can comprise about 24 nucleotides.

**[00300]** In some aspects, a target binding domain can comprise between about 10 to about 70 nucleotides. A target binding domain can comprise between about 30 to about 55 nucleotides. A target binding domain can comprise between about 35 to about 50 nucleotides.

**[00301]** In some aspects, a unique nucleic acid sequence which identifies the target analyte bound to the target binding domain can comprise between about 20 to about 40. A unique nucleic acid sequence which identifies the target analyte bound to the target binding domain can comprise about 25 nucleotides, or about 35 nucleotides, or about 12 nucleotides.

**[00302]** In some aspects, an amplification primer binding site can comprise between about 20 to about 50 nucleotides. An amplification primer binding site can comprise about 33 nucleotides, or about 34 nucleotides.

**[00303]** In some aspects, a nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can comprise between about 1 to about 20 nucleotides. A nucleic acid sequence which identifies the specific location of the tissue sample from which the identifier oligonucleotide was released can comprise about 8 nucleotides.

**[00304]** In some aspects, a nucleic acid sequence comprising a unique molecular identifier can comprise between about 5 to about 20 nucleotides. A nucleic acid sequence comprising a unique molecular identifier can comprise about 14 nucleotides.

**[00305]** As used herein, the terms “region of interest” and “ROI” are used in their broadest sense to refer to a specific location within a sample that is to be analyzed using the methods of the present disclosure.

**[00306]** As used herein, the term “adjacent” can mean within about 1 nucleotide, or within about 2 nucleotides, or within about 3 nucleotides, or within about 4 nucleotides, or within about 5 nucleotides, or within about 6 nucleotides, or within about 7 nucleotides, or within about 8 nucleotides, or within about 9 nucleotides, or within about 10 nucleotides, or within about 11

nucleotides, or within about 12 nucleotides, or within about 13 nucleotides, or within about 14 nucleotides, or within about 15 nucleotides, or within about 16 nucleotides, or within about 17 nucleotides, or within about 18 nucleotides, or within about 19 nucleotides, or within about 20 nucleotides, or within about 21 nucleotides, or within about 22 nucleotides, or within about 23 nucleotides, or within about 24 nucleotides, or within about 25 nucleotides, or within about 26 nucleotides, or within about 27 nucleotides, or within about 28 nucleotides, or within about 29 nucleotides, or within about 30 nucleotides, or within about 40 nucleotides, or within about nucleotides, or within about 50 nucleotides, or within about 60 nucleotides, or within about 70 nucleotides, or within about 80 nucleotides, or within about 90 nucleotides, or within about 100 nucleotides.

**[00307]** As used herein, the term “spatially detecting” is used in its broadest sense to refer to the identification of the presence of a specific target analyte within a specific region of interest in a sample. Spatially detecting can comprise quantifying the amount of a specific target analyte present within a specific region of interest in a sample. Spatially detecting can further comprise quantifying the relative amount of a first target analyte within a specific region of interest in a sample as compared to the amount of at least a second target analyte within a specific region of interest in a sample. Spatially detecting can also comprise quantifying the relative amount of a specific target analyte within a first region of interest in a sample compared to the amount of the same target analyte in at least a second region of interest in the same sample or different sample.

**[00308]** In some aspects of the methods and compositions of the present disclosure, a target analyte can be any molecule within a sample that is to be spatially detected. Target analytes include, but are not limited to, nucleic acid molecules and protein molecules. When the target analyte is a protein, the protein can be referred to as a target protein. When the target analyte is a nucleic acid, the nucleic acid can be referred to as a target nucleic acid. Target nucleic acids can include, but are not limited to, mRNA molecules, micro RNA (miRNA) molecules, tRNA molecules, rRNA molecules, gDNA or any other nucleic acid present within a sample.

**[00309]** In some aspects of the methods and compositions of the present disclosure, the term target binding domain is used in its broadest sense to refer to a portion of a probe of the present disclosure that binds to, either directly or indirectly, a target analyte located in a sample. A target binding domain can comprise nucleic acid, protein, at least one antibody, an aptamer, or any combination

thereof. A target binding domain can comprise DNA, RNA or any combination thereof. A target binding domain can comprise any number of modified nucleotides and/or nucleic acid analogues.

**[00310]** In the aspect that the target analyte to be spatially detected is a target protein, a target binding domain can be a protein-target binding domain. A protein-target binding domain can comprise an antibody or antibody fragment that binds to the target protein.

**[00311]** In the aspect that the target analyte to be spatially detected is a target nucleic acid, a target binding domain can be a target nucleic acid-binding region. A target nucleic acid-binding region can comprise a nucleic acid that is complementary to the target nucleic acid to be spatially detected. A target nucleic acid-binding region can comprise a nucleic acid that hybridizes to the target nucleic acid to be detected.

**[00312]** As used herein, the term “hybridize” is used in its broadest sense to mean the formation of a stable nucleic acid duplex. In one aspect, “stable duplex” means that a duplex structure is not destroyed by a stringent wash under conditions such as, for example, a temperature of either about 5 °C below or about 5 °C above the  $T_m$  of a strand of the duplex and low monovalent salt concentration, *e.g.*, less than 0.2 M, or less than 0.1 M or salt concentrations known to those of skill in the art. A duplex can be “perfectly matched”, such that the polynucleotide and/or oligonucleotide strands making up the duplex form a double stranded structure with one another such that every nucleotide in each strand undergoes Watson-Crick base pairing with a nucleotide in the other strand. The term “duplex” comprises, but is not limited to, the pairing of nucleoside analogs, such as deoxyinosine, nucleosides with 2-aminopurine bases, PNAs, and the like, that can be employed. A duplex can comprise at least one mismatch, wherein the term “mismatch” means that a pair of nucleotides in the duplex fail to undergo Watson-Crick bonding.

**[00313]** As used herein, the term “hybridization conditions,” will typically include salt concentrations of less than about 1 M, more usually less than about 500 mM and even more usually less than about 200 mM. Hybridization temperatures can be as low as 5 °C, but are typically greater than 22 °C, more typically greater than about 30 °C, and often in excess of about 37 °C. Hybridizations are usually performed under stringent conditions, *e.g.*, conditions under which a probe will specifically hybridize to its target analyte. Stringent conditions are sequence-dependent and are different in different circumstances. Longer fragments can require higher hybridization temperatures for specific hybridization. As other factors can affect the stringency of hybridization, including base composition and length of the complementary strands, presence of organic solvents

and extent of base mismatching, the combination of parameters is more important than the absolute measure of any one alone. Certain hybridization conditions will promote the formation of a duplex between the entire length of a target binding domain and the target analyte. Other hybridization conditions will promote the formation of a duplex only along certain portions of the target binding domain.

**[00314]** In some aspects of the methods and compositions of the present disclosure, a probe can comprise a target binding domain directly, or indirectly linked to an identifier oligonucleotide. In the context of a probe, an identifier oligonucleotide is a polynucleotide that comprises a nucleic acid sequence that identifies the target analyte bound to the target binding domain of that probe. That is to say, the identifier oligonucleotide comprises a specific nucleic acid sequence that is *a priori* assigned to the specific target analyte bound to the target binding to which the identifier oligonucleotide is attached. In a non-limiting example, a probe designated as “probe *X*” designed to spatially detect “target analyte *X*” comprises a target binding domain designated “target binding domain *X*” linked to an identifier oligonucleotide designated “identifier oligonucleotide *X*”. Target binding domain *X* binds to target analyte *X* and identifier oligonucleotide *X* comprises a nucleic acid sequence, designated as “nucleic acid sequence *X*”, which corresponds to target analyte *X*. Thus, if a skilled artisan practicing the methods of the present disclosure were to collect identifier oligonucleotides released from a region of interest in sample and obtain nucleic acid sequence *X* after sequencing, the skilled artisan would understand that to indicate that target analyte *X* was present in that region of interest. The amount, or number of sequencing reads, of nucleic acid sequence *X* can be used to determine the quantify, in absolute or relative terms, the amount of target analyte *X* within the region of interest.

**[00315]** As used herein, the term “amplification primer binding site” is used in its broadest sense to refer to a nucleic acid sequence that is complementary to, or at least partially complementary to at least one amplification primer, wherein the amplification primer is a short single-stranded or partially single-stranded oligonucleotide that is sufficient to prime DNA and/or RNA synthesis, for example, by PCR.

**[00316]** In some aspects of the methods and compositions of the present disclosure, a target binding domain can be linked to an identifier oligonucleotide by a cleavable linker. Suitable cleavable linkers include, but are not limited to, chemically cleavable linkers (*e.g.* a linker that is cleaved when exposed to a particular chemical, combination of chemicals or reaction conditions), a photo-

cleavable linker (*e.g.* a linker that is cleaved when exposed to light of a sufficient wavelength or light comprising a sufficient range of wavelengths), or an enzymatically cleavable linker (*e.g.* a linker that is cleaved by a specific enzyme or class of enzymes). Thus, as used herein the phrase “providing a force to a location of the sample sufficient to release an identifier oligonucleotide” is used in its broadest sense to describe changing the conditions within a certain region of interest in a sample such that, for any probe bound to a target analyte within that region of interest, the linker between the target binding domain of the probe and the identifier oligonucleotide of the probe is cleaved, thereby separating the identifier oligonucleotide from the target binding domain so that the identifier oligonucleotide can be subsequently collected from solution. For example, in aspects wherein a probe comprises a chemically cleavable linker between the target binding domain and the identifier oligonucleotide, providing a force to a location of the sample sufficient to release an identifier oligonucleotide can comprise exposing that location of the sample to the specific chemical, combination of chemicals or reaction conditions that catalyze the cleavage of the linker. In another non-limiting example, in aspects wherein a probe comprises a photo-cleavable linker between the target binding domain and the identifier oligonucleotide, providing a force to a location of the sample sufficient to release an identifier oligonucleotide can comprise exposing/exciting that location of the sample with light of a sufficient wavelength capable of cleaving the photo-cleavable linker. In another non-limiting example, in aspects wherein a probe comprises an enzymatically cleavable linker between the target binding domain and the identifier oligonucleotide, providing a force to a location of the sample sufficient to release an identifier oligonucleotide can comprise exposing that location of the sample to an amount of enzyme sufficient to catalyze the cleavage of the linker.

**[00317]** Providing a force to a location of the sample sufficient to release an identifier oligonucleotide can result in at least about 10%, or at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 99% of probes bound to target analytes within that location of the sample to undergo cleavage of the linker connecting the target binding domain and the identifier oligonucleotide.

**[00318]** As would be appreciated by one skilled in the art, the term “unique molecular identifier” or “UMI” refer to short nucleic acid sequences that are used to quantify and reduce quantitative bias caused by nucleic acid amplification prior to sequencing reactions.

**[00319]** In some aspects of the methods and compositions of the present disclosure, an affinity moiety can comprise biotin, avidin, streptavidin, nucleic acid, or any combination thereof.

**[00320]** In some aspects of the methods and compositions of the present disclosure, a probe can comprise at least at least about 5, about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

**[00321]** In some aspects of the methods and compositions of the present disclosure, a target binding domain can comprise at least at least about 5, about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

**[00322]** In some aspects of the methods and compositions of the present disclosure, an identifier oligonucleotide can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

**[00323]** In some aspects of the methods and compositions of the present disclosure, an amplification primer can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about

150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

**[00324]** In some aspects of the methods and compositions of the present disclosure, a nucleic acid probe can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 190 or at least about 200 nucleotides.

**[00325]** In some aspects of the methods and compositions of the present disclosure, a nucleic acid complementary to a portion of a identifier oligonucleotide can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95 or at least about 100 nucleotides.

**[00326]** In some aspects of the methods and compositions of the present disclosure, a nucleic acid sequence comprising a molecular identifier can comprise at least about 5, or at least about 10 nucleotides, or at least about 15, or at least about 20, or at least about 25, or at least about 30, or at least about 35, or at least about 40, or at least about 45, or at least about 50 nucleotides.

**[00327]** In some aspects of the methods and compositions of the present disclosure, an amplification primer binding site can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65 or at least about 70 nucleotides.

**[00328]** In some aspects of the methods and compositions of the present disclosure, a flow cell adapter sequence suitable for sequencing can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35 at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95 or at least about 100 nucleotides.

**[00329]** In some aspects of the methods and compositions of the present disclosure, a flow cell binding site can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 90, at least about 95 or at least about 100 nucleotides.

**[00330]** In some aspects of the methods and compositions of the present disclosure, a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released can comprise at least about 5, or at least about 10 nucleotides, or at least about 15, or at least about 20, or at least about 25, or at least about 30, or at least about 35, or at least about 40, or at least about 45, or at least about 50 nucleotides.

**[00331]** In some aspects of the methods and compositions of the present disclosure, a unique nucleic acid sequence which identifies the target analyte bound to a target binding domain can comprise at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 90, at least about 95 or at least about 100 nucleotides.

**[00332]** In some aspects of the methods and compositions of the present disclosure, a probe, a target binding domain, an identifier oligonucleotide, an amplification primer, a nucleic acid probe, a nucleic acid complementary to a portion of a identifier oligonucleotide, a nucleic acid sequence comprising a molecular identifier, an amplification primer binding site, a flow cell adapter sequence, a flow cell binding site, a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released, a unique nucleic acid sequence which identifies the target analyte bound to a target binding domain or any combination thereof can comprise at least one natural base, can comprise no natural bases, can comprise at least one modified nucleotide or nucleic acid analog, can comprise no modified nucleotides or nucleic acid analogs, can comprise at least one universal base, can comprise no universal bases, can comprise at least one degenerate base or can comprise no degenerate bases.

**[00333]** In some aspects of the methods and compositions of the present disclosure, a probe, a target binding domain, an identifier oligonucleotide, an amplification primer, a nucleic acid probe, a nucleic acid complementary to a portion of a identifier oligonucleotide, a nucleic acid sequence comprising a molecular identifier, an amplification primer binding site, a flow cell

adapter sequence, a flow cell binding site, a nucleic acid sequence which identifies the specific location of the tissue sample from which an identifier oligonucleotide was released, a unique nucleic acid sequence which identifies the target analyte bound to a target binding domain or any combination thereof can comprise any combination natural bases (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more natural bases), modified nucleotides or nucleic acid analogs (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more modified or analog nucleotides), universal bases (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more universal bases), or degenerate bases (*e.g.* 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more degenerative bases). When present in a combination, the natural bases, modified nucleotides or nucleic acid analogs, universal bases and degenerate bases can be arranged in any order.

**[00334]** The terms “modified nucleotides” or “nucleic acid analogues” include, but are not limited to, locked nucleic acids (LNA), bridged nucleic acids (BNA), propyne-modified nucleic acids, zip nucleic acids (ZNA<sup>®</sup>), isoguanine and isocytosine. Preferably, the modified nucleotides or nucleic acid analogues are locked nucleic acids (LNAs).

**[00335]** The term “locked nucleic acids (LNA)” as used herein includes, but is not limited to, a modified RNA nucleotide in which the ribose moiety comprises a methylene bridge connecting the 2' oxygen and the 4' carbon. This methylene bridge locks the ribose in the 3'-endo confirmation, also known as the north confirmation, that is found in A-form RNA duplexes. The term inaccessible RNA can be used interchangeably with LNA. The term “bridged nucleic acids (BNA)” as used herein includes, but is not limited to, modified RNA molecules that comprise a five-membered or six-membered bridged structure with a fixed 3'-endo confirmation, also known as the north confirmation. The bridged structure connects the 2' oxygen of the ribose to the 4' carbon of the ribose. Various different bridge structures are possible containing carbon, nitrogen, and hydrogen atoms. The term “propyne-modified nucleic acids” as used herein includes, but is not limited to, pyrimidines, namely cytosine and thymine/uracil, that comprise a propyne modification at the C5 position of the nucleic acid base. The term “zip nucleic acids (ZNA<sup>®</sup>)” as used herein includes, but is not limited to, oligonucleotides that are conjugated with cationic spermine moieties.

**[00336]** The term “universal base” as used herein includes, but is not limited to, a nucleotide base does not follow Watson-Crick base pair rules but rather can bind to any of the four canonical bases (A, T/U, C, G) located on the target nucleic acid. The term “degenerate base” as used herein includes, but is not limited to, a nucleotide base that does not follow Watson-Crick base

pair rules but rather can bind to at least two of the four canonical bases A, T/U, C, G), but not all four. A degenerate base can also be termed a Wobble base; these terms are used interchangeably herein.

**[00337]** As used in this Specification and the appended claims, the singular forms “a,” “an” and “the” include plural referents unless the context clearly dictates otherwise.

**[00338]** Unless specifically stated or obvious from context, as used herein, the term “or” is understood to be inclusive and covers both “or” and “and”.

**[00339]** Unless specifically stated or obvious from context, as used herein, the term “about” is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term “about.”

**[00340]** Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although other probes, compositions, methods, and kits similar, or equivalent, to those described herein can be used in the practice of the present disclosure, the preferred materials and methods are described herein. It is to be understood that the terminology used herein is for the purpose of describing particular aspects only, and is not intended to be limiting.

### EXAMPLES

**[00341]** Example 1—two-ended adapter ligation method for 96 multiplexed samples

**[00342]** In this example, a two-ended adapter ligation method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. The nucleic acid adapters used in this experiment were partially double-stranded. The nucleic acid adapters comprised a first strand and a second strand. The first strand comprised a 5' phosphate moiety for ligation. The first strand also comprised a constant nucleic acid sequence to minimize ligation bias (GCGTAGTG), a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a first amplification primer binding site (SEQ ID NO: 2). The second strand comprised a single overhanging thymine nucleotide at the 3' end, a sequence complementary to the constant nucleic acid sequence to minimize ligation bias present

in the first strand, a sequence complementary to the unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released present in the first strand and a second amplification primer binding site (SEQ ID NO: 1).

**[00343]** To form the partially double stranded nucleic acid adapters, first strand oligonucleotides and second strand oligonucleotides were combined in equimolar proportion for a final total oligonucleotide concentration of 28  $\mu\text{M}$  in buffer comprising 50 mM NaCl. The oligonucleotide mixture was heated at 95°C for 2 minutes and cooled at ambient temperature for 30 minutes, thereby annealing the first strand and second strand oligonucleotides together to form the partially double-stranded nucleic acid adapters. Annealed nucleic acid adapters were diluted to final concentration ranging between 0.02  $\mu\text{M}$  to 0.002  $\mu\text{M}$  in a solution of 10 mM Tris pH 8 and 0.05% Tween20.

**[00344]** Collected identifier oligonucleotides were end repaired and A-tailed using NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) with a modified protocol. End repair/A-tail master mix was prepared by combining the following: 627.8  $\mu\text{L}$  of PCR-grade H<sub>2</sub>O, 143.9  $\mu\text{L}$  of NEBNext Ultra II End Prep Reaction Buffer, and 61.7  $\mu\text{L}$  of NEBNext Ultra II End Prep Enzyme Mix. 8.3  $\mu\text{L}$  of end repair/A-tail master mix was added to 4  $\mu\text{L}$  of each sample of identifier oligonucleotides. The reaction was incubated for 30 minutes at 20°C with a heated lid of >75°C, followed by a second incubation for 30 minutes at 65°C. The repaired/A-tailed identifier oligonucleotide mixtures were then stored at 4°C.

**[00345]** Following end repair and A-tailing, the nucleic acid adapters were ligated to the repaired/A-tailed identifier oligonucleotides by adding 6.4  $\mu\text{L}$  of NEBNext Ultra II Ligation Master Mix, 0.2  $\mu\text{L}$  of NEBNext Ligation Enhancer, and 1  $\mu\text{L}$  of the nucleic acid adapter dilution to each repaired/A-tailed identifier oligonucleotide mixture. These reactions were incubated for 15 minutes at 20°C with the heated lid off and subsequently quenched with 1  $\mu\text{L}$  0.5M EDTA. All of the reactions were then pooled into a single 15 mL conical tube to form a pooled adapter-ligated sample.

**[00346]** The pooled adaptor-ligated sample was purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350  $\mu\text{L}$  of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). AMPure XP bead cleanup was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200  $\mu\text{L}$

of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. AMPure XP bead cleanup was then repeated with 400  $\mu$ L of AMPure XP beads and eluted in 20  $\mu$ L of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20 to obtain purified adapter-ligated samples.

[00347] Following AMPure XP cleanup, PCR reactions with purified adaptor-ligated sample were prepared to amplify the adapter-ligated identifier oligonucleotides. To 6  $\mu$ L of the purified adapter-ligated sample, 10  $\mu$ L of NEBNext Ultra II Q5 Master Mix, 0.2  $\mu$ L of 100  $\mu$ M forward and reverse primers, and 3.6  $\mu$ L of PCR-grade H<sub>2</sub>O was added. The forward primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the first strand of the nucleic acid adapter. Table 1 provides the sequences of the forward primers used. Table 1. Forward primers for two-ended adapter ligation.

| Primer Sequence                                                                    | SEQ ID NO |
|------------------------------------------------------------------------------------|-----------|
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 5         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTT<br>TAAGACGACGTCGCTATGGCCTCTCC | 6         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 7         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 8         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTT<br>TAAGACGACGTCGCTATGGCCTCTCC | 9         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 10        |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTT<br>TAAGACGACGTCGCTATGGCCTCTCC | 11        |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 12        |

[00348] The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second amplification primer binding site located on the second strand of the nucleic acid adapter. Table 2 provides the sequences of the reverse primers used.

Table 2. Reverse primers for two-ended adapter ligation

| Primer Sequence                                                     | SEQ ID NO |
|---------------------------------------------------------------------|-----------|
| CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 13        |
| CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 14        |
| CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 15        |
| CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 16        |
| CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 17        |
| CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 18        |
| CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 19        |
| CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 20        |
| CAAGCAGAAGACGGCATACGAGATCATAGCCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 21        |
| CAAGCAGAAGACGGCATACGAGATTTTCGCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT | 22        |
| CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 23        |
| CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 24        |

[00349] The optimal number of PCR cycles was determined empirically with triplicate PCR reactions. Alternately, the optimal number of PCR cycles could have been determined using real-time/qPCR. The PCR program used comprised the following steps:

- (1) 30 seconds at 98°C
- (2) 10 seconds at 98°C
- (3) 1 minute at 65°C
- (4) Repeating steps (2) and (3) nine times
- (5) 5 minutes at 65°C

[00350] The amplified products were purified using 18 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00351] The amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with a custom spike-in primer comprising the nucleotide sequence  
ACACTCTTTAAGACGACGTCGCTATGGCCTCTCC (SEQ ID NO: 25).

**[00352]** Example 2—one-ended adapter ligation method for 96 multiplexed samples

[00353] In this example, a one-ended adapter ligation method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. The nucleic acid adapters used in this experiment were partially double-stranded. The nucleic acid adapters comprised a first strand and a second strand. The first strand comprised a 5' phosphate moiety for ligation. The first strand also comprised a constant nucleic acid sequence to minimize ligation bias (CACTACGC), a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a first amplification primer binding site (SEQ ID NO: 1). The second strand comprised a single overhanging thymine nucleotide at the 3' end, a sequence complementary to the constant nucleic acid sequence to minimize ligation bias present in the first strand and a sequence complementary to the unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released present in the first strand.

[00354] To form the partially double stranded nucleic acid adapters, first strand oligonucleotides and second strand oligonucleotides were combined in equimolar proportion for a final total oligonucleotide concentration of 28 μM in 50 mM NaCl. The oligonucleotide mixture was heated to 95°C for 2 minutes and cooled at ambient temperature for 30 minutes, thereby annealing the first strand and second strand oligonucleotides together to form the partially double-stranded nucleic acid adapters. Annealed nucleic acid adapters were diluted to a final concentration ranging between 0.02 μM to 0.002 μM in a solution of 10 mM Tris pH 8 and 0.05% Tween20.

[00355] Nucleic acid adapters were ligated to the collected identifier oligonucleotides by addition of 10 μL of 2X rapid ligation buffer (Enzymatics), 1 μL of T4 DNA Rapid Ligase (Enzymatics), and 1 μL of annealed nucleic acid adapter dilutions to each sample of collected identifier oligonucleotides. Samples were incubated for 15 minutes at 20°C with the heated lid off and

subsequently quenched with 1  $\mu$ L 0.5M EDTA. All of the reactions were then pooled into a single 15 mL conical tube to form a pooled adaptor-ligated sample.

**[00356]** The pooled adaptor-ligated sample was purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350  $\mu$ L of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). AMPure XP bead cleanup was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200  $\mu$ L of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. AMPure XP bead cleanup was then repeated with 400  $\mu$ L of AMPure XP beads and eluted in 20  $\mu$ L of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20 to obtain purified adaptor-ligated samples.

**[00357]** Following AMPure XP cleanup, PCR reactions with purified adaptor-ligated sample were prepared to amplify the adaptor-ligated identifier oligonucleotides. To 6  $\mu$ L of the purified adaptor-ligated sample, 10  $\mu$ L of NEBNext Ultra II Q5 Master Mix, 0.2  $\mu$ L of 100  $\mu$ M forward and reverse primers, and 3.6  $\mu$ L of PCR-grade H<sub>2</sub>O was added. The forward primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the first strand of the nucleic acid adapter. Table 3 provides the sequences of the forward primers used.

Table 3. Forward primers for two-ended adapter ligation.

| Primer Sequence                                                                    | SEQ ID NO |
|------------------------------------------------------------------------------------|-----------|
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 5         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 6         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 7         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 8         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 9         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 10        |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 11        |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 12        |

**[00358]** The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second

amplification primer binding site located on the second strand of the nucleic acid adapter. Table 4 provides the sequences of the reverse primers used.

Table 4. Reverse primers for two-ended adapter ligation.

| Primer Sequence                                                     | SEQ ID NO |
|---------------------------------------------------------------------|-----------|
| CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 13        |
| CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 14        |
| CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 15        |
| CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 16        |
| CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 17        |
| CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 18        |
| CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 19        |
| CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 20        |
| CAAGCAGAAGACGGCATACGAGATCATAGCCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 21        |
| CAAGCAGAAGACGGCATACGAGATTTTCGCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT | 22        |
| CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 23        |
| CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT  | 24        |

[00359] The optimal number of PCR cycles was determined empirically with triplicate PCR reactions. Alternately, the optimal number of PCR cycles could have been determined using real-time/qPCR. The PCR program used comprised the following steps:

- (1) 30 seconds at 98°C
- (2) 10 seconds at 98°C
- (3) 1 minute at 65°C
- (4) Repeating steps (2) and (3) nine times
- (5) 5 minutes at 65°C

[00360] The amplified products were purified using 18 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00361] The amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with a custom spike-in primer comprising the nucleotide sequence  
ACACTCTTTAAGACGACGTCGCTATGGCCTCTCC (SEQ ID NO: 25).

[00362] Example 3—templated-primer extension method for 96 multiplexed samples

[00363] In this example, a templated-primer extension method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. The single stranded nucleic acid templates used in this example comprised a 3' biotin moiety, a region complementary to the unique nucleic acid sequences present in the collected identifier oligonucleotides, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding sequence

(GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT, SEQ ID NO: 26). Table 5 provides the sequences of the single-stranded nucleic acid templates used in this example.

Table 5. Single stranded nucleic acid templates for templated-primer extension method

| Single-stranded nucleic acid templates                                                    | SEQ ID NO |
|-------------------------------------------------------------------------------------------|-----------|
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTTGA<br>AGCACACCGTTTTTCTTCTTCTTTCACGG   | 27        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNACCC<br>ACAGGTTATACGGGATTATCCGGTTATCCA  | 28        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCGAC<br>ACCGAGTTCGACCGTTATGTTGGTAGGATC  | 29        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCGGT<br>GTGTAAGCGTAACGATGTTGGTGTCTGCTCT | 30        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCAGA<br>CACTGCGACAACACGATCATGACACAGA    | 31        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNATAT<br>TCTGTACTCAGTGCCTATCCACCTAATAGG  | 32        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTTCA<br>GTTATAATGTGTCCAGCAGAAGCAGGAATT  | 33        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNGTCC<br>TTTGTGGGCGGACCGTAATGAGGAATTTG   | 34        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNGATG<br>AGACTTCTACATGTCCGATGTTTTTGTGCT  | 35        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNACTC<br>ACACATAGTACTGACACGTAAGATAGGATG  | 36        |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTTAC<br>CCTATCTCGTCTATGTACGTCAGGCTGAAT  | 37        |

|                                                                                           |    |
|-------------------------------------------------------------------------------------------|----|
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNATCA<br>ACGTAGGGTAAGGTCATATTTTACCTTAC   | 38 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTTCC<br>CTCTTTCTCCGCTTATGGATGAAAGGACAG  | 39 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCCTG<br>CACAGTGAGTTTCTTCACTCTAACTCTCT   | 40 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTGTC<br>GCTCTAGTGTGACTTTTCCACCTCGCATCT  | 41 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNATAT<br>CTTTCTCGGGTAAAGATTAGGCGTCCGATA  | 42 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCGAT<br>TAGCCGTAGACGCAACTCATTGCCGAAGAT  | 43 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTGTC<br>AGCATTTCAGTACGAGTGATGCAGATAAAC  | 44 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTATA<br>GTTACCAAGTACTATGGGTTGGTGGGAAGCC | 45 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCCAA<br>TTATACTGTCTGTTATGTTCTCGGATAAGC  | 46 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTCAG<br>GTGCTTGTAGGCTCATGATAGGGGTAATGC  | 47 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCTCT<br>GCTGTAATCTCAGCTCCACTTGTTTCTAAG  | 48 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNGTGC<br>ATATTGCAGCTGAGCCAGCTCAATTTGAAG  | 49 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCCGT<br>TGATTTACGCAACAGCGGCTTATATAGCTC  | 50 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCATC<br>ATCGACAGTTTCGACGCCCTATAACATGATA | 51 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNATCG<br>CAGGATGGTACAGCATCATACATGATGAGC  | 52 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCTGA<br>TAAGTCGTAGGAATGTCGCTTAATACGGAT  | 53 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNATGG<br>CGGTTTCGGGTCCTGCACTATTCCTAATAA  | 54 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNCCAG<br>TACGGGTACTAATAAGTGTCATATCTATTG  | 55 |
| GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNNTGTT<br>GGAGAGGTTAGAGGTGAGGAGGCGAAGATA  | 56 |

[00364] Single stranded nucleic acid templates were ordered from Integrated DNA Technologies, Inc. and quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific). Individual single stranded nucleic acid templates were normalized to a standard concentration and then pooled to be equimolar. The pool of single stranded nucleic acid templates was diluted to 0.83 nM in a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00365] The collected identifier oligonucleotides were hybridized to the single stranded nucleic acid templates and extended by addition of 10  $\mu$ L of NEBNext Ultra II Q5 Master Mix (New England Biolabs), 4  $\mu$ L of the diluted single stranded nucleic acid template pool and 4  $\mu$ L of H<sub>2</sub>O to 2  $\mu$ L of each sample of identifier oligonucleotides. The following PCR program was used to extend the identifier oligonucleotides:

- (1) 30 seconds at 98°C, 10×
- (2) 1 minutes at 98°C,
- (3) 1 minutes at 68°C
- (4) 1 minutes at 72°C
- (5) Repeating steps (2)-(4) ten times
- (6) 2 minutes at 72°C

[00366] The extension products were stored at 4°C. Magnetic streptavidin beads (MyOne Streptavidin C1 beads, Thermo Fisher Scientific) were washed in 1X Binding and Washing Buffer (5 mM Tris-HCl, 0.5 mM EDTA, 1M NaCl), and 5  $\mu$ L of streptavidin beads were added to each extension product sample. The extension product samples were incubated with the beads on an orbital mixer for a minimum of 15 minutes. Following incubation, the samples were heated to 95°C for 3 minutes and transferred to a magnetic plate. Supernatant was extracted immediately after sufficient bead pelleting to yield the purified extension product samples.

[00367] The purified extension product samples were amplified by adding to 7.5  $\mu$ L of each purified extension product sample, 12.5  $\mu$ L of NEBNext Ultra II Q5 Master Mix, 0.25  $\mu$ L of 100  $\mu$ M forward primer, 1  $\mu$ L of 25  $\mu$ M reverse primer and 3.8  $\mu$ L of PCR-grade H<sub>2</sub>O. The forward primer comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the identifier oligonucleotide. Table 6 provides the sequences of the forward primers used in this example.

Table 6. Forward primers for templated-primer extension method.

| Primer Sequence                                                                    | SEQ ID NO |
|------------------------------------------------------------------------------------|-----------|
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 5         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 6         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 7         |

|                                                                                    |    |
|------------------------------------------------------------------------------------|----|
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 8  |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 9  |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 10 |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 11 |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 12 |

[00368] The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second amplification primer binding site located single-stranded nucleic acid template. Table 7 provides the sequences of the reverse primers used.

Table 7. Reverse primers for templated-primer extension method.

| Primer Sequence                                                        | SEQ ID NO |
|------------------------------------------------------------------------|-----------|
| CAAGCAGAAGACGGCATACGAGATGTCGGTAAAGTACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 57        |
| CAAGCAGAAGACGGCATACGAGATAGGTCACTGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 58        |
| CAAGCAGAAGACGGCATACGAGATGAATCCGAGTACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 59        |
| CAAGCAGAAGACGGCATACGAGATGTACCTTGGTACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT   | 60        |
| CAAGCAGAAGACGGCATACGAGATCATGAGGAGTACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 61        |
| CAAGCAGAAGACGGCATACGAGATTGACTGACGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 62        |
| CAAGCAGAAGACGGCATACGAGATCGTATTCGGTACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT   | 63        |
| CAAGCAGAAGACGGCATACGAGATCTCCTAGAGTACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT   | 64        |
| CAAGCAGAAGACGGCATACGAGATTAGTTGCGGTACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT   | 65        |
| CAAGCAGAAGACGGCATACGAGATGAGATACGGTACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 66        |
| CAAGCAGAAGACGGCATACGAGATAGGTGTACGTACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 67        |
| CAAGCAGAAGACGGCATACGAGATTAATGCCGGTACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT   | 68        |
| CAAGCAGAAGACGGCATACGAGATTCAGACGAGTACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 69        |
| CAAGCAGAAGACGGCATACGAGATGATAGGCTGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 70        |

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| CAAGCAGAAGACGGCATAACGAGATTGGTACAGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 71 |
| CAAGCAGAAGACGGCATAACGAGATCAAGGTCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 72 |
| CAAGCAGAAGACGGCATAACGAGATGCTATCCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 73 |
| CAAGCAGAAGACGGCATAACGAGATATGGAAGGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 74 |
| CAAGCAGAAGACGGCATAACGAGATTCAAGGACGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 75 |
| CAAGCAGAAGACGGCATAACGAGATGTTACGCAGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 76 |
| CAAGCAGAAGACGGCATAACGAGATAGTCTGTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 77 |
| CAAGCAGAAGACGGCATAACGAGATGCACGTAAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 78 |
| CAAGCAGAAGACGGCATAACGAGATAACCTTGGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 79 |
| CAAGCAGAAGACGGCATAACGAGATATTGCGTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 80 |
| CAAGCAGAAGACGGCATAACGAGATACCTGGAAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 81 |
| CAAGCAGAAGACGGCATAACGAGATGGAGATGAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 82 |
| CAAGCAGAAGACGGCATAACGAGATGTACTCTCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 83 |
| CAAGCAGAAGACGGCATAACGAGATGTAACGACGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 84 |
| CAAGCAGAAGACGGCATAACGAGATATTCCTCCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 85 |
| CAAGCAGAAGACGGCATAACGAGATGTGTTCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 86 |
| CAAGCAGAAGACGGCATAACGAGATAAGCACTGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 87 |
| CAAGCAGAAGACGGCATAACGAGATCTAGCAAGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 88 |
| CAAGCAGAAGACGGCATAACGAGATTGCTTCCAGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 89 |
| CAAGCAGAAGACGGCATAACGAGATGCTTAGCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 90 |
| CAAGCAGAAGACGGCATAACGAGATAACCGTTCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 91 |
| CAAGCAGAAGACGGCATAACGAGATGACATTCCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 92 |
| CAAGCAGAAGACGGCATAACGAGATAGACCGTAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 93 |
| CAAGCAGAAGACGGCATAACGAGATGATACTGGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 94 |
| CAAGCAGAAGACGGCATAACGAGATTGCGTAGAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 95 |

|                                                                          |     |
|--------------------------------------------------------------------------|-----|
| CAAGCAGAAGACGGCATAACGAGATTCCGGTTACGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 96  |
| CAAGCAGAAGACGGCATAACGAGATATGACGTCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 97  |
| CAAGCAGAAGACGGCATAACGAGATGCTGTAAGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 98  |
| CAAGCAGAAGACGGCATAACGAGATGCAATGGAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 99  |
| CAAGCAGAAGACGGCATAACGAGATATCTCGCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 100 |
| CAAGCAGAAGACGGCATAACGAGATGGCTATTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 101 |
| CAAGCAGAAGACGGCATAACGAGATGGTGTCTTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 102 |
| CAAGCAGAAGACGGCATAACGAGATTCAACTGGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 103 |
| CAAGCAGAAGACGGCATAACGAGATCTTACCAGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT   | 104 |
| CAAGCAGAAGACGGCATAACGAGATACGGTCTTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 105 |
| CAAGCAGAAGACGGCATAACGAGATTCTCGCAAGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 106 |
| CAAGCAGAAGACGGCATAACGAGATGGAATTGCGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 107 |
| CAAGCAGAAGACGGCATAACGAGATACGGATTCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 108 |
| CAAGCAGAAGACGGCATAACGAGATTTAAGCGGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 109 |
| CAAGCAGAAGACGGCATAACGAGATTGCAGGTAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 110 |
| CAAGCAGAAGACGGCATAACGAGATCAATCGACGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 111 |
| CAAGCAGAAGACGGCATAACGAGATGTGCCATAGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 112 |
| CAAGCAGAAGACGGCATAACGAGATTGTTTCGAGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 113 |
| CAAGCAGAAGACGGCATAACGAGATTGGAGTTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 114 |
| CAAGCAGAAGACGGCATAACGAGATACGATGACGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 115 |
| CAAGCAGAAGACGGCATAACGAGATTGATGTCCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 116 |
| CAAGCAGAAGACGGCATAACGAGATTGAACCTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 117 |
| CAAGCAGAAGACGGCATAACGAGATCTTCGTTTCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 118 |
| CAAGCAGAAGACGGCATAACGAGATCTTCTGAGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 119 |
| CAAGCAGAAGACGGCATAACGAGATTGCTCATGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 120 |

|                                                                          |     |
|--------------------------------------------------------------------------|-----|
| CAAGCAGAAGACGGCATAACGAGATAGTTCGTCTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT   | 121 |
| CAAGCAGAAGACGGCATAACGAGATTAGCGTCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 122 |
| CAAGCAGAAGACGGCATAACGAGATGGCGTTATGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 123 |
| CAAGCAGAAGACGGCATAACGAGATGGTGATTCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 124 |
| CAAGCAGAAGACGGCATAACGAGATAACTTGCCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 125 |
| CAAGCAGAAGACGGCATAACGAGATGCAAGATCGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 126 |
| CAAGCAGAAGACGGCATAACGAGATTTCGATTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 127 |
| CAAGCAGAAGACGGCATAACGAGATTGTACACCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 128 |
| CAAGCAGAAGACGGCATAACGAGATAGCTCCTAGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 129 |
| CAAGCAGAAGACGGCATAACGAGATGCAATTCGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 130 |
| CAAGCAGAAGACGGCATAACGAGATCTTAGGACGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 131 |
| CAAGCAGAAGACGGCATAACGAGATGTCCTAAGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 132 |
| CAAGCAGAAGACGGCATAACGAGATAACGTGGAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 133 |
| CAAGCAGAAGACGGCATAACGAGATCTGTGTTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 134 |
| CAAGCAGAAGACGGCATAACGAGATGTTAAGGCGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 135 |
| CAAGCAGAAGACGGCATAACGAGATCACCTTACGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 136 |
| CAAGCAGAAGACGGCATAACGAGATTGGTAGCTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 137 |
| CAAGCAGAAGACGGCATAACGAGATCAGTGAAGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 138 |
| CAAGCAGAAGACGGCATAACGAGATGTTCAACCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 139 |
| CAAGCAGAAGACGGCATAACGAGATTGGCTATCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 140 |
| CAAGCAGAAGACGGCATAACGAGATCTGGAGTAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 141 |
| CAAGCAGAAGACGGCATAACGAGATTCTCTTCCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 142 |
| CAAGCAGAAGACGGCATAACGAGATTCTAACGCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 143 |
| CAAGCAGAAGACGGCATAACGAGATGGTCAGATGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT  | 144 |
| CAAGCAGAAGACGGCATAACGAGATCTCTGGTTGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT  | 145 |

|                                                                         |     |
|-------------------------------------------------------------------------|-----|
| CAAGCAGAAGACGGCATAACGAGATTGTGGTACGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 146 |
| CAAGCAGAAGACGGCATAACGAGATCCTATACCGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 147 |
| CAAGCAGAAGACGGCATAACGAGATTTCTCTCGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 148 |
| CAAGCAGAAGACGGCATAACGAGATGTATGCTGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 149 |
| CAAGCAGAAGACGGCATAACGAGATAAGTCGAGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 150 |
| CAAGCAGAAGACGGCATAACGAGATAACCGAAGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 151 |
| CAAGCAGAAGACGGCATAACGAGATTGTTGTGGGTGACTGGAGTTCAGACGTGT<br>GCTCTTCCGATCT | 152 |

[00369] The PCR program used to amplify the purified extension products comprised the following steps:

- (1) 30 seconds at 98°C
- (2) 10 seconds at 98°C
- (3) 30 seconds at 65°C
- (4) 30 seconds at 72°C
- (5) Repeat steps (2)-(4) eighteen times
- (6) 2 minutes at 72°C

[00370] The amplified extension products were stored at 4°C. 4 µL of each PCR reaction were combined into four pools, for 24 samples per pool.

[00371] The pooled PCR reactions were purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 100 µL of AMPure XP beads and 400 µL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). Purification was performed with 76.8 µL of diluted AMPure XP beads and eluted in 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. Beads were retained and cleanup process was repeated with 24 µL of AMPure XP buffer and eluted in 20 µL of buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00372] The purified PCR products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The purified PCR products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with

a custom spike-in primer comprising the nucleotide sequence

ACACTCTTTAAGACGACGTCGCTATGGCCTCTCC (SEQ ID NO: 25).

**[00373]** Example 4—long probe hybridization method for 96 multiplexed samples

**[00374]** In this example, a long probe hybridization method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. In this example, the first nucleic acid probe comprises a 5' phosphate moiety, a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a first amplification primer binding site comprising an i7 sequence (SEQ ID NO: 1), unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released and a P7 flow cell adapter sequence (SEQ ID NO: 4). The second nucleic acid probe comprises a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier, a second amplification primer binding site comprising an i5 sequence (SEQ ID NO: 2) and a P5 flow cell adapter sequence (SEQ ID NO: 3).

**[00375]** The first and second nucleic acid probes were ordered from Integrated DNA Technologies, Inc. and quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific). Individual nucleic acid probes were normalized to a standard concentration, pooled to be equimolar, and diluted to 0.83 nM in a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The nucleic acid probes and the identifier oligonucleotides were hybridized by combining 0.5  $\mu$ L of diluted nucleic acid probe pool with 2  $\mu$ L of a mixture of identifier oligonucleotides collected from a sample solution in a buffer comprising 50 mM NaCl. This mixture was heated for 2 minutes at 95°C and cooled for 30 minutes at ambient temperature to yield an annealed identifier oligonucleotide-nucleic acid probe mixture.

**[00376]** In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were not adjacent and were not overlapping, a gap extension reaction was performed. To 2.5  $\mu$ L of each annealed identifier oligonucleotide-nucleic acid probe mixture, 3.8  $\mu$ L of NEBNext Ultra II Q5 Master mix and 1.3  $\mu$ L of PCR-grade H<sub>2</sub>O was added. The mixture was then subjected to the following Gap extension temperature cycle:

- (1) 30 seconds at 98°C
- (2) 1 minute at 98°C
- (3) 1 minute at 68°C

- (4) 1 minute at 72°C,
- (5) Repeat steps (2)-(4) ten times
- (6) 2 minutes at 72°C

[00377] The gap extension products were then stored at 4°C. The first and second nucleic acid probes were then ligated together by adding to 1 µL of the gap extension product, 10 µL of 2X rapid ligation buffer (Enzymatics), 1 µL of T4 DNA Rapid Ligase (Enzymatics), and 8 µL of PCR-grade H<sub>2</sub>O. These ligation reactions were incubated for 15 minutes at 20°C, subsequently quenched with 1 µL 0.5M EDTA, and pooled into a single 15 mL conical tube.

[00378] In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were adjacent and were not overlapping, nick repair ligation reaction was performed. To 2.5 µL of each annealed identifier oligonucleotide-nucleic acid probe mixture, 10 µL of 2X rapid ligation buffer (Enzymatics), 1 µL of T4 DNA Rapid Ligase (Enzymatics), and 1 µL of PCR-grade H<sub>2</sub>O was added. These ligation reactions were incubated for 15 minutes at 20°C, subsequently quenched with 1 µL of 0.5M EDTA, and pooled into a single 15 mL conical tube.

[00379] The pools of quenched ligation reactions were then purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350 µL of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). The purification was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The AMPure XP bead cleanup was then repeated with 400 µL of AMPure XP beads and eluted in 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00380] To amplify the purified ligation products, PCR reactions with purified ligation products and primers were prepared. To 6 µL of purified ligation product, 10 µL of NEBNext Ultra II Q5 Master Mix, 0.2 µL of 100 µM forward primer (CAAGCAGAAGACGGCATACGA, SEQ ID NO: 153) and reverse primer (AATGATACGGCGACCACCGA, SEQ ID NO: 154) and 3.6 µL of PCR-grade H<sub>2</sub>O was added. The PCR program used to amplify the purified extension products comprised the following steps:

- (1) 30 seconds at 98°C
- (2) 10 seconds at 98°C
- (3) 30 seconds at 65°C

- (4) 30 seconds at 72°C
- (5) Repeat steps (2)-(4) eighteen times
- (6) 2 minutes at 72°C

[00381] The amplified products were stored at 4°C. 4 µL of each PCR reaction were combined into six pools, for 16 samples per pool. The amplified products were further purified using an AMPure XP bead cleanup with 64 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00382] The purified amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The purified amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with either standard sequencing primers or a custom spike-in Read1 primer (SEQ ID NO: 25).

[00383] Example 5—short probe hybridization method for 96 multiplexed samples

[00384] In this example, a short probe hybridization method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. In this example, the first nucleic acid probe comprises a 5' phosphate moiety, a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a first amplification primer binding site comprising an i7 sequence (SEQ ID NO: 1) and a unique nucleic acid sequence which identifies the specific location of the sample from which the identifier oligonucleotide was released. The second nucleic acid probe comprises a nucleic acid sequence complementary to a portion of the identifier oligonucleotide, a nucleic acid sequence comprising a unique molecular identifier and a second amplification primer binding site comprising an i5 sequence (SEQ ID NO: 2).

[00385] The first and second nucleic acid probes were ordered from Integrated DNA Technologies, Inc. and quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific). Individual nucleic acid probes were normalized to a standard concentration, pooled to be equimolar, and diluted to 0.83 nM in a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The nucleic acid probes and the identifier oligonucleotides were hybridized by combining 0.5 µL of diluted nucleic acid probe pool with 2 µL of a mixture of identifier oligonucleotides collected from a sample solution in a buffer comprising 50 mM NaCl. This

mixture was heated for 2 minutes at 95°C and cooled for 30 minutes at ambient temperature to yield an annealed identifier oligonucleotide-nucleic acid probe mixture.

**[00386]** In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were not adjacent and were not overlapping, a gap extension reaction was performed. To 2.5 µL of each annealed identifier oligonucleotide-nucleic acid probe mixture, 3.8 µL of NEBNext Ultra II Q5 Master mix and 1.3 µL of PCR-grade H<sub>2</sub>O was added. The mixture was then subjected to the following Gap extension temperature cycle:

- (1) 30 seconds at 98°C
- (2) 1 minute at 98°C
- (3) 1 minute at 68°C
- (4) 1 minute at 72°C,
- (5) Repeat steps (2)-(4) ten times
- (6) 2 minutes at 72°C

**[00387]** The gap extension products were then stored at 4°C. The first and second nucleic acid probes were then ligated together by adding to 1 µL of the gap extension product, 10 µL of 2X rapid ligation buffer (Enzymatics), 1 µL of T4 DNA Rapid Ligase (Enzymatics), and 8 µL of PCR-grade H<sub>2</sub>O. These ligation reactions were incubated 15 min 20°C, quenched with 1 µL 0.5M EDTA, and pooled into a single 15 mL conical tube.

**[00388]** In the case in which the first and the second nucleic acid probes hybridized to the identifier oligonucleotide such that the first and the second nucleic acid probes were adjacent and were not overlapping, nick repair ligation reaction was performed. To 2.5 µL of each annealed identifier oligonucleotide-nucleic acid probe mixture, 10 µL of 2X rapid ligation buffer (Enzymatics), 1 µL of T4 DNA Rapid Ligase (Enzymatics), and 1 µL of PCR-grade H<sub>2</sub>O was added. These ligation reactions were incubated for 15 minutes at 20°C, subsequently quenched with 1 µL of 0.5M EDTA, and pooled into a single 15 mL conical tube.

**[00389]** The pools of quenched ligation reactions were then purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 350 µL of AMPure XP beads and 3.15 mL of AMPure XP buffer (2.5M NaCl, 20% PEG8000). The purification was performed with 3.5 mL of diluted AMPure XP beads and eluted in 200 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20. The AMPure XP bead

cleanup was then repeated with 400  $\mu$ L of AMPure XP beads and eluted in 20  $\mu$ L of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00390] To amplify the purified ligation products, PCR reactions with purified ligation products and primers were prepared. To 6  $\mu$ L of purified ligation product, 10  $\mu$ L of NEBNext Ultra II Q5 Master Mix, 0.2  $\mu$ L of 100  $\mu$ M forward primer and reverse primer and 3.6  $\mu$ L of PCR-grade H<sub>2</sub>O was added. The forward primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the first amplification primer binding site located on the first strand of the nucleic acid adapter. Table 8 provides the sequences of the forward primers used.

Table 8. Forward primers for short probe hybridization method.

| Primer Sequence                                                        | SEQ ID NO |
|------------------------------------------------------------------------|-----------|
| CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 13        |
| CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 14        |
| CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 15        |
| CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 16        |
| CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 17        |
| CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 18        |
| CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 19        |
| CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 20        |
| CAAGCAGAAGACGGCATACGAGATCATAGCCGGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 21        |
| CAAGCAGAAGACGGCATACGAGATTTGCGGGAGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 22        |
| CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGACTGGAGTTCAGACGTG<br>TGCTCTTCCGATCT | 23        |
| CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGACTGGAGTTCAGACGTG<br>GCTCTTCCGATCT  | 24        |

[00391] The reverse primers comprised a flow cell adapter sequence suitable for sequencing, a unique molecular identifier and a nucleic acid sequence complementary to the second amplification primer binding site located on the second strand of the nucleic acid adapter. Table 9 provides the sequences of the reverse primers used.

Table 9. Reverse primers for short probe hybridization method.

| Primer Sequence                                                            | SEQ ID NO |
|----------------------------------------------------------------------------|-----------|
| AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCTACA<br>CGACGCTCTTCCGATCT | 155       |
| AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCCTAC<br>ACGACGCTCTTCCGATCT | 156       |
| AATGATACGGCGACCACCGAGATCTACACCCTATCCTACACTCTTTCCCTACA<br>CGACGCTCTTCCGATCT | 157       |
| AATGATACGGCGACCACCGAGATCTACACGGCTCTGAACACTCTTTCCCTACA<br>CGACGCTCTTCCGATCT | 158       |
| AATGATACGGCGACCACCGAGATCTACACAGGCGAAGACACTCTTTCCCTAC<br>ACGACGCTCTTCCGATCT | 159       |
| AATGATACGGCGACCACCGAGATCTACACTAATCTTAACACTCTTTCCCTACA<br>CGACGCTCTTCCGATCT | 160       |
| AATGATACGGCGACCACCGAGATCTACACCAGGACGTACACTCTTTCCCTACA<br>CGACGCTCTTCCGATCT | 161       |
| AATGATACGGCGACCACCGAGATCTACACGTACTGACACACTCTTTCCCTACA<br>CGACGCTCTTCCGATCT | 162       |

[00392] The PCR program used to amplify the purified extension products comprised the following steps:

- (1) 30 seconds at 98°C
- (2) 10 seconds at 98°C
- (3) 30 seconds at 65°C
- (4) 30 seconds at 72°C
- (5) Repeat steps (2)-(4) eighteen times
- (6) 2 minutes at 72°C

[00393] The amplified products were stored at 4°C. 4 µL of each PCR reaction were combined into six pools, for 16 samples per pool. The amplified products were further purified using an AMPure XP bead cleanup with 64 µL of AMPure XP beads and eluting with 20 µL of a buffer comprising 10 mM Tris pH 8 and 0.05% Tween20.

[00394] The purified amplified products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics) and KAPA Library Quantification Kit for Illumina Platforms (Kapa Biosystems). The purified amplified products were also diluted to 15 pM for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2×75 bp) with either standard sequencing primers or a custom spike-in Read1 primer (SEQ ID NO: 25).

**[00395]** Example 6—direct PCR method for 96 multiplexed samples

**[00396]** In this example, a direct PCR method of the present disclosure was used to sequence identifier oligonucleotides collected from 96 multiplexed samples. In this example, 8 species of forward amplification primers and 12 species of reverse amplification primers were used. The forward primers comprised a P5 flow cell adapter (SEQ ID NO: 3), a nucleic acid sequence comprising a unique molecular identifier and a region complementary to a first amplification primer binding site present on the identifier oligonucleotide. Table 10 provides the sequences of the forward amplification primers used.

Table 10. Forward amplification primers for direct PCR method.

| Primer Sequence                                                                    | SEQ ID NO |
|------------------------------------------------------------------------------------|-----------|
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATATAGCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 5         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAATAGAGGCACACTCTT<br>TAAGACGACGTCGCTATGGCCTCTCC | 6         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACCTATCCTACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 7         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGGCTCTGAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 8         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAAGGCGAAGACACTCTT<br>TAAGACGACGTCGCTATGGCCTCTCC | 9         |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGATAATCTTAACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 10        |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGACAGGACGTACACTCTT<br>TAAGACGACGTCGCTATGGCCTCTCC | 11        |
| AATGATACGGCGACCACCGAGATCTACACGCTCAGAGTACTGACACACTCTTT<br>AAGACGACGTCGCTATGGCCTCTCC | 12        |

**[00397]** The reverse primers comprised a P7 flow cell adapter (SEQ ID NO: 4), a nucleic acid sequence comprising a unique molecular identifier and a region complementary to a second amplification primer binding site present on the identifier oligonucleotide. Table 11 provides the sequences of the reverse amplification primers used.

Table 11. Reverse amplification primers for direct PCR method.

| Primer Sequence                                                                                           | SEQ ID NO |
|-----------------------------------------------------------------------------------------------------------|-----------|
| CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCT<br>CTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 163       |
| CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCT<br>CTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 164       |
| CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCT<br>CTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 165       |

|                                                                                                        |     |
|--------------------------------------------------------------------------------------------------------|-----|
| CAAGCAGAAGACGGCATAACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 166 |
| CAAGCAGAAGACGGCATAACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 167 |
| CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 168 |
| CAAGCAGAAGACGGCATAACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 169 |
| CAAGCAGAAGACGGCATAACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 170 |
| CAAGCAGAAGACGGCATAACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 171 |
| CAAGCAGAAGACGGCATAACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 172 |
| CAAGCAGAAGACGGCATAACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 173 |
| CAAGCAGAAGACGGCATAACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNNNNNNAACGGACAGGATGCAGCAAAAT | 174 |

**[00398]** In the case of 8 forward amplification primers and 12 reverse amplification primers, when the unique molecular identifiers from a pair of forward and reverse primers are combined, a total of 96 unique combinations can be obtained, allowing for the multiplexing of 96 samples.

**[00399]** To amplify the collected identifier oligonucleotides for sequencing, PCR reactions with collected identifier oligonucleotides and forward and reverse amplification primers were prepared on a 96-well plate with 2  $\mu$ L of each identifier oligonucleotide sample, 10  $\mu$ L of NEBNext Ultra II Q5 Master Mix, 2  $\mu$ L of 10  $\mu$ M forward amplification primer, 2  $\mu$ L of reverse amplification primer, and 4  $\mu$ L of PCR-grade H<sub>2</sub>O. Each well in the 96-well plate contained an identifier oligonucleotide sample and a unique combination of forward and reverse amplification primers. The PCR program used comprised the following steps:

- (1) 30 seconds at 98°C
- (2) 10 seconds at 98°C
- (3) 30 seconds at 65°C
- (4) 30 seconds at 72°C
- (5) Repeat steps (2)-(4) six to ten times
- (6) 2 minutes at 72°C

**[00400]** The amplified products were stored at 4°C. 10  $\mu$ L of each PCR reaction was combined into a single 15 mL conical tube.

[00401] The pooled PCR reactions were purified using diluted Agencourt AMPure XP magnetic beads (Beckman Coulter Genomics Inc.), which were prepared by combining 115.2  $\mu$ L of AMPure XP beads and 1036.8  $\mu$ L of AMPure XP buffer (2.5M NaCl, 20% PEG8000). Purification was performed with 1152  $\mu$ L of diluted AMPure XP beads and eluted in 60  $\mu$ L of a buffer comprising 10 mM Tris pH 8. The purification process was repeated with 60  $\mu$ L of AMPure XP beads and eluted in 70  $\mu$ L of a buffer comprising 10 mM Tris pH 8.

[00402] Following AMPure XP cleanup, PCR reactions with universal primers were prepared with 9  $\mu$ L of pooled direct PCR product, 15  $\mu$ L of NEBNext Ultra II Q5 Master Mix, 3  $\mu$ L of 10  $\mu$ M universal P7 primer (SEQ ID NO: 153) and 2  $\mu$ L of 10  $\mu$ M universal P5 primer (SEQ ID NO: 154). The PCR program used was:

- (1) 30 seconds at 98°C,
- (2) 10 seconds at 98°C
- (3) 30 seconds at 65°C
- (4) 30 seconds at 72°C
- (5) Repeat steps (2)-(4) 15 to 24 times
- (6) 2 minutes at 72°C

[00403] Two rounds of AMPure XP bead cleanup was performed. The first round was performed with 30  $\mu$ L of beads and eluted with 20  $\mu$ L of a buffer comprising 10 mM Tris pH 8 and the second round was performed with 20  $\mu$ L beads and eluted with 11  $\mu$ L of a buffer comprising 10 mM Tris pH 8.

[00404] These purified PCR products were assessed using a High Sensitivity DNA chip on 2100 Bioanalyzer (Agilent Genomics). The purified PCR products were also diluted for sequencing on a MiSeq (Illumina) according to manufacturer's instructions (MiSeq Reagent Kit v3 2 $\times$ 75 bp) with a custom spike-in primer (SEQ ID NO: 25).

**[00405] Example 7—spatially detecting target analytes in a FFPE sample**

[00406] The methods of the present invention were used to spatially detect a plurality of different target analytes, including target proteins and target RNAs, in a sample of inflamed human tonsil tissue FFPE section.

[00407] In one experiment, 30 different target proteins were spatially detected using the methods of the present disclosure in two serial sections cut from the inflamed human tonsil tissue FFPE section. The 30 target proteins are put forth in Table 12. The 30 target proteins included IgG

Rabbit isotype and IgG Mouse isotype as negative controls, as these target proteins should not have been present in the inflamed human tonsil sample and therefore should not have been detected.

Table 12. Target Proteins

| Target Protein       | Target Protein                        |
|----------------------|---------------------------------------|
| AKT                  | FOXP3                                 |
| B7-H3                | GZMB                                  |
| Bcl-2                | Histone H3                            |
| Beta-2-microglobulin | Ki67                                  |
| Beta-catenin         | CD20                                  |
| CD14                 | P-AKT                                 |
| CD19                 | PanCK                                 |
| CD3                  | PD1                                   |
| CD4                  | PD-L1                                 |
| CD44                 | S6                                    |
| CD45                 | STAT3                                 |
| CD45RO               | P-STAT3                               |
| CD56                 | VISTA                                 |
| CD68                 | IgG Rabbit isotype (negative control) |
| CD8A                 | IgG Mouse isotype (control)           |

[00408] To spatially detect the 30 target proteins, 30 different probes of the present disclosure were used. Each probe comprised a target binding domain comprising an antibody that specifically binds to one of the 30 target proteins in Table 12. The two serial sections were contacted with a plurality of the 30 different probes. Ninety-six regions of interest (ROI) were then identified. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a short probe hybridization method of the present disclosure, thereby spatially detecting the 30 target proteins in the two serial sections. As shown in Figure 21, the number of reads per target protein in each ROI for the two serial sections were well correlated, demonstrating that the method yields reproducible results.

[00409] In a second experiment, 20 different target RNAs were spatially detected using the methods of the present disclosure in two different serial sections cut from the inflamed human tonsil tissue FFPE section. The 20 different target RNAs are put forth in Table 13. The 20 target RNAs included 6 negative controls (Negative Probe) that should not have been detected in the sample.

Table 13. Target RNAs

| Target RNA        | Target RNA        | Target RNA        |
|-------------------|-------------------|-------------------|
| CD3E              | CD40              | CTLA4             |
| CD3G              | CD45              | GAPDH             |
| CD4               | CD74              | KRT13             |
| CD20              | CD79A             | PD1               |
| PSA               | RP56              | Negative Probe #1 |
| Negative Probe #2 | Negative Probe #3 | Negative Probe #4 |
| Negative Probe #5 | Negative Probe #6 |                   |

[00410] To spatially detect the 20 Target RNAs, 20 different probes of the present disclosure were used. Each probe comprised a target binding domain comprising a nucleic acid sequence complementary to at least one portion of one of the 20 target RNAs. The two serial sections were contacted with a plurality of the 20 different probes. Ninety-six regions of interest (ROI) were then identified. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a direct PCR method of the present disclosure, thereby spatially detecting the 20 target RNAs in the two serial sections. As shown in Figure 22, the number of reads per target RNA in each ROI for the two serial sections were well correlated, demonstrating that the method yields reproducible results.

**[00411] Example 8—spatially detecting target proteins in a fluorescently stained FFPE sample**

[00412] In another experiment, a 5  $\mu$ m FFPE section of inflamed human tonsil tissue was stained with 4 fluorescent visualization markers: (1) CD3E, a T-cell marker; (2) PanCK, an epithelial cell marker; (3) Ki-67, a proliferation marker; and (4) SYTO83, a DNA stain, as shown in the left panel of Figure 23. The stained FFPE section was then contacted with the probes directed against 30 target proteins, as described in Example 7. As shown in the left panel Figure 23, 96 regions of interest (ROIs) were selected. Each ROI was a circle with a 500  $\mu$ m diameter. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a short probe hybridization method of the present disclosure, thereby spatially detecting the 30 target proteins in the FFPE section. As shown in the right panel of Figure 23, PanCK, CD3E and Ki67 were spatially detected in ROIs that correlated with their

fluorescent visualization markers. Thus, the results generated by the methods of the present disclosure correlate with the results generated using established immunohistochemical method.

**[00413]** Example 9—spatially detecting target RNAs in a FFPE sample

**[00414]** In another experiment, a 5  $\mu$ m section from an inflamed human tonsil tissue FFPE block was contacted with probes directed against 20 target RNAs, as described in Example 7. 96 regions of interest (ROIs) were then selected. Each ROI was a circle with a 500  $\mu$ m diameter. For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a direct PCR method of the present disclosure, thereby spatially detecting the 20 target RNAs in the two serial sections. The total RNA from a 20  $\mu$ m section from the same inflamed human tonsil tissue FFPE block was then isolated. The total RNA was analyzed using the NanoString nCounter® system. Figure 24 shows that the average number of counts for 11 different RNA targets recorded using the methods of the present disclosure were well correlated with the average number of counts for the same 11 different RNA targets recorded using the nCounter® system. Thus, the results generated using the methods of the present disclosure correlate with the results generated using established direct detection methods.

**[00415]** Example 10—spatially detecting target RNAs in specific sub-regions of an ROI

**[00416]** In another experiment, a 5  $\mu$ m section from an inflamed human tonsil tissue FFPE block was contacted with probes directed against 30 target proteins, as described in Example 7. The same 5  $\mu$ m section was also stained with 4 fluorescent visualization markers: (1) CD3E, a T-cell marker; (2) PanCK, an epithelial cell marker; (3) Ki-67, a proliferation marker; and (4) SYTO83, a DNA stain. As shown in Figure 25, 48 regions of interest (ROIs) were identified. For each ROI, two sub-regions were then identified based on the fluorescent staining. Areas of an ROI that were fluorescently stained positive for PanCK (PanCK+) were designated a “tumor” sub-region and the areas of an ROI that lacked PanCK fluorescent staining were designated a “micro-environment” sub-region, as shown in Figure 25. For each ROI, the tumor sub-region and the micro-environment sub-region were separately illuminated with UV light to release the identifier oligonucleotides from the probes bound within each sub-region by creating a custom mask based on the intensity of PanCK fluorescent staining. The released identifier oligonucleotides were also separately collected. The collected identifier oligonucleotides were then analyzed using the short-probe hybridization method of the present disclosure and the NanoString nCounter®

system. As shown in the bottom panel of Figure 25, the results using the NanoString nCounter® system and the short-probe hybridization method of the present disclosure were well correlated. Furthermore, in the tumor sub-regions, PanCK was detected at a significantly higher level as compared to the micro-environment sub-regions. Thus, the spatial detection results provided by the methods of the present disclosure are consistent with established fluorescent immunohistochemical methods and allows for the spatial detection within highly specific regions of a sample.

**[00417]** Example 11—96-plex human immuno-oncology panel

**[00418]** A 96-plex human immuno-oncology panel was designed for use in the direct-PCR methods of the present invention. The panel comprised a plurality of probes that could be used to spatially detect 96 different human target RNAs using the direct-PCR methods of the present disclosure. The 96 target RNAs are shown in Table 14.

Table 14. Target RNAs

| Target | Target  | Target  | Target   | Target | Target | Target   | Target  |
|--------|---------|---------|----------|--------|--------|----------|---------|
| AKT1   | CD3E    | CEACAM8 | FOXP3    | IFNG   | KRT1   | NCAM1    | SOD2    |
| ARG1   | CD4     | CMKLR1  | GZMB     | IFNGR1 | KRT10  | NKG7     | SOX10   |
| B2M    | CD40    | CSF1R   | H3F3A    | IL10   | KRT14  | NT5E     | STAT1   |
| BATF3  | CD40LG  | CTLA4   | HAVCR2   | IL12B  | KRT17  | PDCD1    | STAT2   |
| BCL2   | CD44    | CTNNB1  | HIF1A    | IL15   | KRT18  | PDCD1LG2 | STAT3   |
| BCL2L1 | CD47    | CXCL10  | HLA_DQA1 | IL1B   | KRT19  | PECAM1   | TBX21   |
| CCL5   | CD68    | CXCL9   | HLA_DRB  | IL6    | KRT6A  | PMEL     | TIGIT   |
| CCND1  | CD74    | CSCR6   | HLA_E    | ITGAM  | KRT7   | PSMB10   | TNF     |
| CD14   | CD86    | DKK2    | ICAM1    | ITGAV  | LAG3   | PTEN     | TNFRSF9 |
| CD27   | CD8A    | EPCAM   | ICOSLG   | ITGAX  | LY6E   | PRTPRC   | TNFSF4  |
| CD274  | CEACAM1 | FAS     | IDO1     | ITGB2  | MKI67  | RPS6     | VEGFA   |
| CD276  | CEACAM6 | FASLG   | IFNAR1   | ITGB8  | MS4A1  | S100B    | VSIR    |

**[00419]** In total, the panel comprised 928 different probes. Each of the probes comprised an identifier oligonucleotide comprising a first amplification primer binding site, a nucleic acid sequence comprising a unique molecular identifier, a unique nucleic acid sequence which identified the target RNA bound to the target binding domain and a second amplification primer binding site. Figure 26 shows a schematic of the probes used in the panel. For each of the 96

target RNAs, there was at least one probe within the 928 probe set comprising a target binding domain that directly or indirectly hybridized to that target RNA. For most of the 96 target RNAs, there were 10 different probes that directly or indirectly hybridized to the specific target RNA. These 10 different probes directly or indirectly hybridized to different locations on the target RNA to create a “tiling” effect, as shown in the top panel of Figure 27. Tiling multiple probes onto a target RNA means that each target RNA will be individually detected multiple times, increasing the overall accuracy of the measurement. For example, as shown in the bottom panel of Figure 27, in the case where 10 probes are tiled onto a single target RNA, one of the probes may be incorrectly detected too many times (outlier high count probe), while another probe may be incorrectly detected too few times (outlier low count probe). However, the other 8 probes may be detected at a similar level, indicating that the two outliers should be discarded during analysis and the signals from the 8 probes used to generate a more accurate measurement of the abundance of the target RNA.

[00420] The set of 928 probes also comprised 80 negative control probes. Each of the 80 negative control probes comprised a target binding domain comprising a scrambled, non-specific nucleic acid sequence that was designed using guidelines from the External RNA Controls Consortium such that the target binding domain should not be complementary to RNA molecules present within a human sample. Thus, these 80 negative control probes should not be detected during analysis.

[00421] The 96-plex human immune-oncology panel was used to analyze a tissue microarray. The tissue microarray comprised FFPE samples of 22 common human cell lines, including normal and cancerous cell types. Some of the cell lines are shown in Table 15.

[00422] Table 15. Cell lines

| Cell Line | Cell Line  | Cell Line | Cell Line |
|-----------|------------|-----------|-----------|
| CCRF-CEM  | DAUDI      | H596      | H2228     |
| HT29      | HUT78      | HUH7      | JURKAT    |
| M14       | MDA-MB-468 | MOLT4     | RAJI      |
| SKBR3     | SUDHL1     | SUDHL4    |           |

[00423] The tissue microarray also comprised one mouse cell line (3T3) as a negative control. Each of the FFPE samples on the microarray was contacted with a plurality of the 928 different probes in the immun-oncology panel. As shown in Figure 28, for each of the FFPE samples, at least three circular regions of interest (ROIs) with a diameter of 300  $\mu$ m were selected. As a

negative control, ROIs were also selected on regions of the microarray that did not comprise a FFPE sample (glass negative control). For each ROI, the ROI was illuminated with UV light to release the identifier oligonucleotides from the probes bound within the ROI. The released identifier oligonucleotides were then collected and identified using a direct PCR method of the present disclosure thereby spatially detecting the 96 target RNAs in each of the FFPE samples on the tissue microarray.

[00424] Figure 29 shows that a sufficient read depth was achieved using a MiSeq v3 flowcell. The top panel of Figure 30 shows that none of the target RNAs were spatially detected for the glass negative control ROIs. Likewise, the bottom panel of Figure 30 shows that nearly none of the target RNAs were detected in the negative control mouse 3T3 FFPE sample. Conversely, as shown in Figures 31, 33 and 34, specific target RNAs were successfully detected in the HEK293 (human embryonic kidney) FFPE sample and the Jurkat (human T-cell lymphocyte) FFPE sample. Figures 31, 33 and 34 show that clusters of “tiled” probes were detected for particular target RNAs, including AKT1, B2M, CD3E, HIF1A, PTEN, RPS6, STAT1, STAT2, STAT3, VEGF, PTPRC (CD45), and KRT1/10/18/19. These results indicated that there are certain target RNAs that are differentially transcribed in the two different cells lines. The results of this experiment were also verified using the NanoString nCounter system to identify the collected identifier oligonucleotides. As shown in Figure 27, the results using the direct-PCR method of the present disclosure were consistent with the results obtained using the NanoString nCounter system.