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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 January 2025 (16.01.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/014758 A1

(51) International Patent Classification:
G02B 21/36 (2006.01) *G01N 15/14* (2024.01)

(21) International Application Number:
PCT/US2024/036755

(22) International Filing Date:
03 July 2024 (03.07.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/525,540 07 July 2023 (07.07.2023) US

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

(72) Inventors: **SKRYNNYK, Oleksiy**; 6230 Stoneridge Mall Road, Pleasanton, CA 94588-3260 (US). **SMIECHOWSKI, Adam Henryk**; 6230 Stoneridge Mall Road, Pleasanton, CA 94588-3260 (US). **MONKOWSKI, Adam Joseph**; 6230 Stoneridge Mall Road, Pleasanton, CA 94588-3260 (US). **HEIBERG, Andrew**; 6230 Stoneridge Mall Road, Pleasanton, CA 94588-3260 (US). **HOFFMAN, David**; 6230 Stoneridge Mall Road, Pleasanton, CA 94588-3260 (US).

(74) Agent: **HUESTIS, Erik A.** et al.; Foley Hoag LLP, 155 Seaport Boulevard, Boston, MA 02210-2600 (US).

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

WO 2025/014758 A1

(54) Title: SYSTEMS AND METHODS FOR TISSUE BOUNDS DETECTION

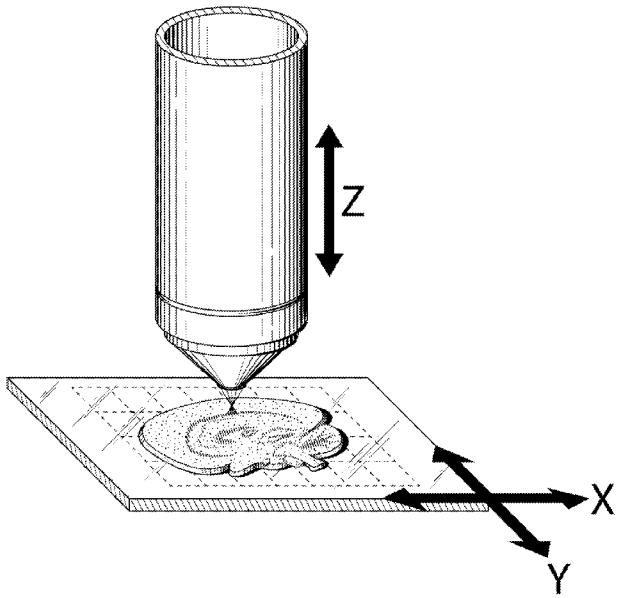


FIG. 1

(57) Abstract: Provided herein are methods for minimizing Z-image acquisition comprising receiving a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

[Continued on next page]

TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

SYSTEMS AND METHODS FOR TISSUE BOUNDS DETECTION

Cross-Reference to Related Applications

[0001] This application claims the benefit of U.S. Provisional Application No. 63/525,540, filed July 7, 2023, which is hereby incorporated by reference in its entirety.

Field of the Invention

[0002] The present disclosure is directed to imaging techniques for samples, *e.g.*, biological samples. More specifically, the present disclosure describes methods for identifying the Z-bounds of a sample based on a signal of interest detected within the sample, and minimizing an imageable volume of said sample to reduce imaging time and/or maximize throughput while retaining data quality.

Background

[0003] *In situ* detection and analysis methods are emerging from the rapidly developing field of spatial transcriptomics. The key objectives in spatial transcriptomics are to detect, quantify, and map gene activity to specific regions in a tissue sample at cellular or sub-cellular resolution. These techniques allow one to study the subcellular distribution of gene activity (as evidenced, *e.g.*, by expressed gene transcripts), and have the potential to provide crucial insights in the fields of developmental biology, oncology, immunology, histology, *etc.*

[0004] *In situ* decoding is a process comprising a plurality of decoding cycles in each of which a different set of barcode probes (*e.g.*, fluorescently-labeled oligonucleotides) is contacted with target analytes (*e.g.*, mRNA sequences) or with target barcodes (*e.g.*, nucleic acid barcodes) associated with the target analytes present in a sample (*e.g.*, a tissue sample) under conditions that promote hybridization. One or more images (*e.g.*, fluorescence images) are acquired in each decoding cycle, and the images are processed to detect the presence and locations of one or more barcode probes in each cycle. The presence and locations of one or more target analyte sequences or associated barcode sequences are then inferred from corresponding code words that are determined based on the set of, *e.g.*, fluorescence signals detected in each decoding cycle of the decoding process.

[0005] Accordingly, there exists a need for fast and accurate Z-dimension image acquisition and analysis methods for use in automated and high-throughput imaging systems.

Summary

[0006] In some aspects, provided herein are methods comprising receiving a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

[0007] In some aspects of the invention as provided herein, is a computer program product comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a processor to cause the processor to perform a method comprising receiving a first set of 3D positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

[0008] Further aspects of the invention include a system comprising a database; and a computing node comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a processor to cause the processor to perform a method comprising receiving, from the database, a first set of 3D positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

[0009] In certain aspects of the invention, provided herein are methods comprising imaging a first imaging volume to determine a first set of 3D positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and imaging the second imaging volume in at least one subsequent probing cycle of the sample.

Brief Description of the Drawings

[0010] FIG. 1 depicts an overview of a volumetric sample imaging system and illustrates a Field of View (FOV) grid bounding the sample (*e.g.*, hydrogel, tissue section, one or more cells, *etc.*) as projected onto the surface of a solid substrate supporting the sample.

[0011] FIG. 2 depicts the XZ cross-sectional view and illustrates tissue non-uniformity in the Z dimension, where the full (non-reduced) imaging volume is oversampled in the Z dimension. The objective lens focal point is positioned to acquire an image at every Z-slice in a Z-stack. An XZ image of signal distribution (bottom) demonstrates a non-uniform distribution of detected signal within the imaging volume.

[0012] FIG. 3 depicts a 3-D view of an individual Z-stack within an imaging volume. The desired effect of thickness profiling is illustrated, where the information/signal-containing slices are imaged based on a measured distribution of signal.

[0013] FIG. 4 is an example workflow of analysis of a biological sample (*e.g.*, a cell or tissue sample) using an opto-fluidic instrument, according to various embodiments.

[0014] FIGS. 5A-5B illustrate cross-sectional views of an optics module in an imaging system, according to some embodiments.

[0015] FIG. 6 depicts a computing node according to some embodiments disclosed herein.

[0016] FIG. 7 is a flowchart illustrating a method of tissue bounds detection according to embodiments of the present disclosure, wherein a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample is received and used to direct imaging volume acquisition in subsequent cycles.

[0017] FIG. 8 is a flowchart illustrating a method of tissue bounds detection according to embodiments of the present disclosure, comprising imaging a first imaging volume to

determine a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, and is used to direct imaging volume acquisition in subsequent cycles.

[0018] FIG. 9 depicts XZ view of an individual Z-stack. The multi-channel nature of the optical acquisition system and how data is acquired and composited (or overlayed) together to result in the total information/signal content within the stack is illustrated, and shows the difference in distributions of signal between channels as objects (RCPs or protein) are labeled with a unique color fluorophore (i.e., C1, C2, C3, and C4).

[0019] FIG. 10 shows imaging of 30 μm "thick" volume to support variation in tissue section thickness and to have margins that mitigate risk of tissue drifting outside of the imaged volume due to thermomechanical deformation, non-orthogonality, and various accumulated position control errors.

[0020] FIG. 11 shows the XZ or YZ view of a cross section of a generic tissue sample, illustrating the distribution of signal within the volume depending upon the target molecule/structure. As provided herein, the illustrated relationship between distributions of signal given a target type can be leveraged for the purposes of estimating Z-distribution of signal in other target types.

[0021] FIG. 12 shows one implementation of fitting a curve using a focus metric (e.g. Discrete Cosine Transform, Tenengrad, Normalized Variance) and determining tissue bounds using a heuristic (e.g. 2nd derivative argmax).

[0022] FIG. 13 shows the XZ view of a cross section of an imaging volume. In particular, the top image represents one implementation where at most the nominal tissue thickness is known, and generous padding is added on top and/or bottom to account for all types of unknown dynamics. The bottom image represents the thickness-profiled implementation, where the distribution of signal is measured (with a large buffer as described above), then reducing the total volume imaged based on the measured signals, and adding smaller padding layers to account for the now known/expected dynamics given the measured thickness, for example.

[0023] FIGS. 14A-14D illustrate the implementation of blob-finding-based image acquisition.

[0024] FIG. 15 depicts *in silico* historical run data analysis assessing the percent of all decoded RCPs with respect to imaged Z-slices, and illustrating that a significant reduction in Z dimension is possible while retaining >99% of decoded RCPs.

[0025] FIG. 16 depicts drift margins for drift compensation, including tissue-based drift compensation.

[0026] FIG. 17 depicts an exemplary instrument data flow for blob-based image analysis.

[0027] FIG. 18 depicts an exemplary thickness profiling sequence using the instrument data flow of FIG. 17.

[0028] FIG. 19 illustrates per cycle imaging duration reduction from blob-based image analysis.

[0029] FIG. 20 illustrates an exemplary optimized profiling sequence including real-time blob detection while acquiring images, without postponement of subsequent cycles.

[0030] In the figures, elements and steps having the same or similar reference numeral have the same or similar attributes or description, unless explicitly stated otherwise.

Detailed Description

[0031] In volumetric sample imaging systems (*e.g.*, an optofluidic instrument), a z-stack of images is obtained for each Field of View (FOV) of the objective (**FIG. 1**). For such automated, high-throughput tissue imaging applications, automatically identifying relevant regions – those regions that contain target molecules such as nucleic acids or proteins – can be challenging as distribution of tissue is non-uniform in many biological samples (**FIG. 2**). Moreover, volumetric sample imaging systems may oversample in the Z-direction to ensure that valuable information is not missed, In automated volumetric imaging, the thickness is overestimated, partly to have margins that mitigate risk of tissue drifting outside of the imaged volume due to thermomechanical deformation, non-orthogonality, and various accumulated position control errors. On the other hand, manual measurement is limited in that it cannot obtain a high resolution of thickness measurements across a sample volume. Even with manual measurement, tissue samples are complex in that they can vary in thickness over the volume and a single thickness estimate obtained by manual methods will have to significantly overestimate the maximum thickness so that no portion of the volume is missed during volumetric imaging.

[0032] Significantly overestimating tissue thickness in volumetric imaging can have negative consequences – specifically, increased computational resources required to store and/or process (*e.g.*, align, register, blob finding, *etc.*) the volumetric images.

[0033] As explained above, current detection and analysis technologies are insufficient for imaging (*e.g.*, volumetric imaging) of samples in automated and high-throughput imaging systems as the imageable volume is determined arbitrarily (*e.g.*, based on estimated tissue thickness) and a significant buffer is added to ensure that signal is not inadvertently missed during imaging, thus causing increased computational burden.

[0034] The present disclosure resolves the above technical problems by providing systems, methods, and computer program products to automatically determine where signal of interest resides within the axial bounds of a sample (*e.g.*, a tissue sample) and minimizing image acquisition in the Z-dimension (FIG. 3). In general, the systems and methods described herein use any suitable method to generate contrast of a sample against a background (*e.g.*, illumination of a sample via bright field imaging, illumination of a sample via fluorescent imaging, inducing autofluorescence within the sample, adding contrast to the sample with one or more stains, *etc.*)

[0035]*General terminology:*

[0036] Specific terminology is used throughout this disclosure to explain various aspects of the methods, systems, and compositions that are described. Unless otherwise defined, all of the technical terms used herein have the same meaning as commonly understood by one of ordinary skill in the art in the field to which this disclosure belongs.

[0037] As used in this specification and the appended claims, the singular forms “a”, “an”, and “the” include plural references unless the context clearly dictates otherwise. For example, "a" or "an" means "at least one" or "one or more". Any reference to “or” herein is intended to encompass “and/or” unless otherwise stated.

[0038] As used herein, the terms "comprising" (and any form or variant of comprising, such as "comprise" and "comprises"), "having" (and any form or variant of having, such as "have" and "has"), "including" (and any form or variant of including, such as "includes" and "include"), or "containing" (and any form or variant of containing, such as "contains" and "contain"), are inclusive or open-ended and do not exclude additional, un-recited additives, components, integers, elements or method steps.

[0039] As used herein, the term "about" a number refers to that number plus or minus 10% of that number. The term 'about' when used in the context of a range refers to that range minus 10% of its lowest value and plus 10% of its greatest value.

[0040] Throughout this disclosure, various aspects of the claimed subject matter are presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the claimed subject matter. Accordingly, the description of a range should be considered to have specifically disclosed all the possible sub-ranges as well as individual numerical values within that range. For example, where a range of values is provided, it is understood that each intervening value, between the upper and lower limit of that range and any other stated or intervening value in that stated range is encompassed within the claimed subject matter. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the claimed subject matter, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the claimed subject matter. This applies regardless of the breadth of the range.

[0041] Use of ordinal terms such as "first", "second", "third", *etc.*, in the claims to modify a claim element does not by itself connote any priority, precedence, or order of one claim element over another or the temporal order in which acts of a method are performed, but are used merely as labels to distinguish one claim element having a certain name from another element having a same name (but for use of the ordinal term) to distinguish the claim elements. Similarly, use of a), b), *etc.*, or i), ii), *etc.* does not by itself connote any priority, precedence, or order of steps in the claims. Similarly, the use of these terms in the specification does not by itself connote any required priority, precedence, or order.

[0042] The term "platform" (or "system") may refer to an ensemble of: (i) instruments (*e.g.*, imaging instruments, fluid controllers, temperature controllers, motion controllers and translation stages, *etc.*), (ii) devices (*e.g.*, specimen slides, substrates, flow cells, microfluidic devices, *etc.*, which may comprise fixed and/or removable or disposable components of the platform), (iii) reagents and/or reagent kits, and (iv) software, or any combination thereof, which allows a user to perform one or more bioassay methods (*e.g.*, analyte detection, *in situ* detection or sequencing, and/or nucleic acid detection or sequencing) depending on the particular combination of instruments, devices, reagents, reagent kits, and/or software utilized.

[0043] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

Barcoding and decoding terminology:

[0044] A "barcode" is a label, or identifier, that conveys or is capable of conveying information (*e.g.*, information about an analyte in a sample, a cell, a bead, a location, a sample, and/or a capture probe). The term "barcode" may refer either to a physical barcode molecule (*e.g.*, a nucleic acid barcode molecule) or to its representation in a computer-readable, digital format (*e.g.*, as a string of characters representing the sequence of bases in a nucleic acid barcode molecule).

[0045] The phrase "barcode diversity" refers to the total number of unique barcode sequences that may be represented by a given set of barcodes.

[0046] A physical barcode molecule (*e.g.*, a nucleic acid barcode molecule) that forms a label or identifier as described above. In some instances, a barcode can be part of an analyte, can be independent of an analyte, can be attached to an analyte, or can be attached to or part of a probe that targets the analyte. In some instances, a particular barcode can be unique relative to other barcodes.

[0047] Physical barcodes can have a variety of different formats. For example, barcodes can include polynucleotide barcodes, random nucleic acid and/or amino acid sequences, and synthetic nucleic acid and/or amino acid sequences. A physical barcode can be attached to an analyte, or to another moiety or structure, in a reversible or irreversible manner. A physical barcode can be added to, for example, a fragment of a deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) sample before or during sequencing of the sample. In some instances, barcodes can allow for identification and/or quantification of individual sequencing-reads in sequencing-based methods (*e.g.*, a barcode can be or can include a unique molecular identifier or "UMI"). Barcodes can be used to detect and spatially-resolve molecular components found in biological samples, for example, at single-cell resolution (*e.g.*, a barcode can be, or can include, a molecular barcode, a spatial barcode, a unique molecular identifier (UMI), *etc.*).

[0048] In some instances, barcodes may comprise a series of two or more segments or sub-barcodes (*e.g.*, corresponding to "letters" or "code words" in a decoded barcode), each of which may comprise one or more of the subunits or building blocks used to synthesize the physical (*e.g.*, nucleic acid) barcode molecules. For example, a nucleic acid barcode molecule may comprise two or more barcode segments, each of which comprises one or more

nucleotides. In some instances, a barcode may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10 segments. In some instances, each segment of a barcode molecule may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, or more than 20 subunits or building blocks. For example, each segment of a nucleic acid barcode molecule may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, or more than 20 nucleotides. In some instances, two or more of the segments of a barcode may be separated by non-barcode segments, *i.e.*, the segments of a barcode molecule need not be contiguous.

[0049] A "digital barcode" (or "digital barcode sequence") is a representation of a corresponding physical barcode (or target analyte sequence) in a computer-readable, digital format as described above. A digital barcode may comprise one or more "letters" (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, or more than 20 letters) or one or more "code words" (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10 code words), where a "code word" comprises, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, or more than 20 letters. In some instances, the sequence of letters or code words in a digital barcode sequence may correspond directly with the sequence of building blocks (*e.g.*, nucleotides) in a physical barcode. In some instances, the sequence of letters or code words in a digital barcode sequence may not correspond directly with the sequence of building blocks in a physical barcode, but rather may comprise, *e.g.*, arbitrary code words that each correspond to a segment of a physical barcode. For example, in some instances, the disclosed methods for decoding and error correction may be applied directly to detecting target analyte sequences (*e.g.*, mRNA sequences) as opposed to detecting target barcodes, and the barcode probes used to detect the target analyte sequences may correspond to letters or code words that have been assigned to specific target analyte sequences but that do not directly correspond to the target analyte sequences.

[0050] A "designed barcode" (or "designed barcode sequence") is a barcode (or its digital equivalent; in some instances a designed barcode may comprise a series of code words that can be assigned to gene transcripts and subsequently decoded into a decoded barcode) that meets a specified set of design criteria as required for a specific application. In some instances, a set of designed barcodes may comprise at least 2, at least 5, at least 10, at least 20, at least 40, at least 60, at least 80, at least 100, at least 200, at least 400, at least 600, at least 800, at least 1,000, at least 2,000, at least 4,000, at least 6,000, at least 8,000, at least 10,000, at least 20,000, at least 40,000, at least 60,000, at least 80,000, at least 100,000, at least 200,000, at least 400,000, at least 600,000, at least 800,000, at least 1,000,000, at least 2 x 10^6, at least 3 x 10^6, at least 4 x 10^6, at least 5 x 10^6, at least 6 x 10^6, at least 7 x 10^6, at least

8×10^6, at least 9×10^6, at least 10^7, at least 10^8, at least 10^9, or more than 10^9 unique barcodes. In some instances, a set of designed barcodes may comprise any number of designed barcodes within the range of values in this paragraph, *e.g.*, 1,225 unique barcodes or 2.38×10^6 unique barcodes. As noted above for barcodes in general, in some instances designed barcodes may comprise two or more segments (corresponding to two or more code words in a decode barcode). In those cases, the specified set of design criteria may be applied to the designed barcodes as a whole, or to one or more segments (or positions) within the designed barcodes.

[0051] A "decoded barcode" (or "decoded barcode sequence") is a digital barcode sequence generated via a decoding process that ideally matches a designed barcode sequence, but that may include errors arising from noise in the synthesis process used to create barcodes and/or noise in the decoding process itself. As noted above, in some instances, the disclosed methods for decoding and error correction may be applied directly to detecting target analytes (*e.g.*, mRNA sequences) as opposed to detecting target barcodes, and the barcode probes used to detect the target analytes may correspond to letters or code words that have been assigned to specific target analytes but that do not directly correspond to the target analytes. In these instances, a decoded barcode (*i.e.*, a series of letters or code words) may serve as a proxy for the target analyte.

[0052] A "corrected barcode" (or "corrected barcode sequence") is a digital barcode sequence derived from a decoded barcode sequence by applying one or more error correction methods.

Probe terminology:

[0053] The term "probe" may refer either to a physical probe molecule (*e.g.*, a nucleic acid probe molecule) or to its representation in a computer-readable, digital format (*e.g.*, as a string of characters representing the sequence of bases in a nucleic acid probe molecule). A "probe" may be, for example, a molecule designed to recognize (and bind or hybridize to) another molecule, *e.g.*, a target analyte, another probe molecule, *etc.*

[0054] In some instances, a physical probe molecule may comprise one or more of the following: (i) a target recognition element (*e.g.*, an antibody capable of recognizing and binding to a target peptide, protein, or small molecule; an oligonucleotide sequence that is complementary to a target gene sequence or gene transcript; or a poly-T oligonucleotide sequence that is complementary to the poly-A tails on messenger RNA molecules), (ii) a barcode element (*e.g.*, a molecular barcode, a cell barcode, a spatial barcode, and/or a unique

molecular identifier (UMI)), (iii) an amplification and/or sequencing primer binding site, (iv) one or more linker regions, (v) one or more detectable tags (*e.g.*, fluorophores), or any combination thereof. In some instances, each component of a probe molecule may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, or more than 20 subunits or building blocks. For example, in some instances, each component of a nucleic acid probe molecule may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, or more than 20 nucleotides.

[0055] In some instances, physical probes may bind or hybridize directly to their target. In some instances, physical probes may bind or hybridize indirectly to their target. For example, in some instances, a secondary probe may bind or hybridize to a primary probe, where the primary probe binds or hybridizes directly to the target analyte. In some instances, a tertiary probe may bind or hybridize to a secondary probe, where the secondary probe binds or hybridizes to a primary probe, and where the primary probe binds or hybridizes directly to the target analyte.

[0056] Examples of "probes" and their applications include, but are not limited to, primary probes (*e.g.*, molecules designed to recognize and bind or hybridize to target analyte), intermediate probes (*e.g.*, molecules designed to recognize and bind or hybridize to another molecule and provide a hybridization or binding site for another probe (*e.g.*, a detection probe), detection probes (*e.g.*, molecules designed to recognize and bind or hybridize to another molecule, detection probes may be labeled with a fluorophore or other detectable tag). In some instances, a probe may be designed to recognize and bind (or hybridize) to a physical barcode sequence (or segments thereof). In some instances, a probe may be used to detect and decode a barcode, *e.g.*, a nucleic acid barcode. In some instances, a probe may bind or hybridize directly to a target barcode. In some instances, a probe may bind or hybridize indirectly to a target barcode (*e.g.*, by binding or hybridizing to other probe molecules which itself is bound or hybridized to the target barcode).

Nucleic acid molecule and nucleotide terminology:

[0057] The terms "nucleic acid" (or "nucleic acid molecule") and "nucleotide" are intended to be consistent with their use in the art and to include naturally-occurring species or functional analogs thereof. Particularly useful functional analogs of nucleic acids are capable of hybridizing to a nucleic acid in a sequence-specific fashion (*e.g.*, capable of hybridizing to two nucleic acids such that ligation can occur between the two hybridized nucleic acids) or are capable of being used as a template for replication of a particular nucleotide sequence.

Naturally-occurring nucleic acids generally have a backbone containing phosphodiester bonds. An analog structure can have an alternate backbone linkage including any of a variety of those known in the art. Naturally-occurring nucleic acids generally have a deoxyribose sugar (*e.g.*, found in deoxyribonucleic acid (DNA)) or a ribose sugar (*e.g.* found in ribonucleic acid (RNA)).

[0058] A nucleic acid can contain nucleotides having any of a variety of analogs of these sugar moieties that are known in the art. A nucleic acid can include natural or non-natural nucleotides. In this regard, a naturally-occurring deoxyribonucleic acid can have one or more bases selected from the group consisting of adenine (A), thymine (T), cytosine (C), or guanine (G), and a ribonucleic acid can have one or more bases selected from the group consisting of uracil (U), adenine (A), cytosine (C), or guanine (G). Useful non-natural bases that can be included in a nucleic acid or nucleotide are known in the art. See, for example, Appella (2009), "Non-Natural Nucleic Acids for Synthetic Biology", *Curr Opin Chem Biol.* 13(5-6): 687–696; and Duffy, *et al.* (2020), "Modified Nucleic Acids: Replication, Evolution, and Next-Generation Therapeutics", *BMC Biology* 18:112.

Samples:

[0059] A sample disclosed herein can be or derived from any biological sample. Methods and compositions disclosed herein may be used for analyzing a biological sample, which may be obtained from a subject using any of a variety of techniques including, but not limited to, biopsy, surgery, and laser capture microscopy (LCM), and generally includes cells and/or other biological material from the subject. In addition to the subjects described above, a biological sample can be obtained from a prokaryote such as a bacterium, an archaea, a virus, or a viroid. A biological sample can also be obtained from non-mammalian organisms (*e.g.*, a plant, an insect, an arachnid, a nematode, a fungus, or an amphibian). A biological sample can also be obtained from a eukaryote, such as a tissue sample, a patient derived organoid (PDO) or patient derived xenograft (PDX). A biological sample from an organism may comprise one or more other organisms or components therefrom. For example, a mammalian tissue section may comprise a prion, a viroid, a virus, a bacterium, a fungus, or components from other organisms, in addition to mammalian cells and non-cellular tissue components. Subjects from which biological samples can be obtained can be healthy or asymptomatic individuals, individuals that have or are suspected of having a disease (*e.g.*, a patient with a disease such as cancer) or a pre-disposition to a disease, and/or individuals in need of therapy or suspected of needing therapy.

[0060] The biological sample can include any number of macromolecules, for example, cellular macromolecules and organelles (*e.g.*, mitochondria and nuclei). The biological sample can be a nucleic acid sample and/or protein sample. The biological sample can be a carbohydrate sample or a lipid sample. The biological sample can be obtained as a tissue sample, such as a tissue section, biopsy, a core biopsy, needle aspirate, or fine needle aspirate. The sample can be a fluid sample, such as a blood sample, urine sample, or saliva sample. The sample can be a skin sample, a colon sample, a cheek swab, a histology sample, a histopathology sample, a plasma or serum sample, a tumor sample, living cells, cultured cells, a clinical sample such as, for example, whole blood or blood-derived products, blood cells, or cultured tissues or cells, including cell suspensions. In some instances, the biological sample may comprise cells which are deposited on a surface.

[0061] Cell-free biological samples can include extracellular macromolecules, *e.g.*, polynucleotides. Extracellular polynucleotides can be isolated from a bodily sample, *e.g.*, blood, plasma, serum, urine, saliva, mucosal excretions, sputum, stool, and tears.

[0062] Biological samples can be derived from a homogeneous culture or population of the subjects or organisms mentioned herein or alternatively from a collection of several different organisms, for example, in a community or ecosystem.

[0063] Biological samples can include one or more diseased cells. A diseased cell can have altered metabolic properties, gene expression, protein expression, and/or morphologic features. Examples of diseases include inflammatory disorders, metabolic disorders, nervous system disorders, and cancer. Cancer cells can be derived from solid tumors, hematological malignancies, cell lines, or obtained as circulating tumor cells. Biological samples can also include fetal cells and immune cells.

[0064] In some instances, a substrate herein can be any support that is insoluble in aqueous liquid and which allows for positioning of biological samples, analytes, features, and/or reagents (*e.g.*, probes) on the support. In some instances, a biological sample can be attached to a substrate. Attachment of the biological sample can be irreversible or reversible, depending upon the nature of the sample and subsequent steps in the analytical method. In certain instances, the sample can be attached to the substrate reversibly by applying a suitable polymer coating to the substrate, and contacting the sample to the polymer coating. The sample can then be detached from the substrate, *e.g.*, using an organic solvent that at least partially dissolves the polymer coating. In some instances, the substrate can be coated or functionalized with one or more substances to facilitate attachment of the sample to the

substrate. Suitable substances that can be used to coat or functionalize the substrate include, but are not limited to, lectins, poly-lysine, antibodies, and polysaccharides.

[0065] A variety of steps can be performed to prepare or process a biological sample for and/or during an assay. Except where indicated otherwise, the preparative or processing steps described below can generally be combined in any manner and in any order to appropriately prepare or process a particular sample for and/or analysis.

Endogenous analytes:

[0066] In some instances, an analyte herein is endogenous to a biological sample and can include nucleic acid analytes and non-nucleic acid analytes. Methods and compositions disclosed herein can be used to analyze nucleic acid analytes (*e.g.*, using a nucleic acid probe or probe set that directly or indirectly hybridizes to a nucleic acid analyte) and/or non-nucleic acid analytes (*e.g.*, using a labelling agent that comprises a reporter oligonucleotide and binds directly or indirectly to a non-nucleic acid analyte) in any suitable combination.

[0067] Examples of non-nucleic acid analytes include, but are not limited to, lipids, carbohydrates, peptides, proteins, glycoproteins (N-linked or O-linked), lipoproteins, phosphoproteins, specific phosphorylated or acetylated variants of proteins, amidation variants of proteins, hydroxylation variants of proteins, methylation variants of proteins, ubiquitylation variants of proteins, sulfation variants of proteins, viral coat proteins, extracellular and intracellular proteins, antibodies, and antigen binding fragments. In some instances, the analyte is inside a cell or on a cell surface, such as a transmembrane analyte or one that is attached to the cell membrane. In some instances, the analyte can be an organelle (*e.g.*, nuclei or mitochondria). In some instances, the analyte is an extracellular analyte, such as a secreted analyte. Exemplary analytes include, but are not limited to, a receptor, an antigen, a surface protein, a transmembrane protein, a cluster of differentiation protein, a protein channel, a protein pump, a carrier protein, a phospholipid, a glycoprotein, a glycolipid, a cell-cell interaction protein complex, an antigen-presenting complex, a major histocompatibility complex, an engineered T-cell receptor, a T-cell receptor, a B-cell receptor, a chimeric antigen receptor, an extracellular matrix protein, a posttranslational modification (*e.g.*, phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation or lipidation) state of a cell surface protein, a gap junction, and an adherens junction.

[0068] Examples of nucleic acid analytes include DNA analytes such as single-stranded DNA (ssDNA), double-stranded DNA (dsDNA), genomic DNA, methylated DNA, specific

methylated DNA sequences, fragmented DNA, mitochondrial DNA, *in situ* synthesized PCR products, and RNA/DNA hybrids. The DNA analyte can be a transcript of another nucleic acid molecule (*e.g.*, DNA or RNA such as mRNA) present in a tissue sample.

[0069] Examples of nucleic acid analytes also include RNA analytes such as various types of coding and non-coding RNA. Examples of the different types of RNA analytes include messenger RNA (mRNA), including a nascent RNA, a pre-mRNA, a primary-transcript RNA, and a processed RNA, such as a capped mRNA (*e.g.*, with a 5' 7-methyl guanosine cap), a polyadenylated mRNA (poly-A tail at the 3' end), and a spliced mRNA in which one or more introns have been removed. Also included in the analytes disclosed herein are non-capped mRNA, a non-polyadenylated mRNA, and a non-spliced mRNA. The RNA analyte can be a transcript of another nucleic acid molecule (*e.g.*, DNA or RNA such as viral RNA) present in a tissue sample. Examples of a non-coding RNAs (ncRNA) that is not translated into a protein include transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs), as well as small non-coding RNAs such as microRNA (miRNA), small interfering RNA (siRNA), Piwi-interacting RNA (piRNA), small nucleolar RNA (snoRNA), small nuclear RNA (snRNA), extracellular RNA (exRNA), small Cajal body-specific RNAs (scaRNAs), and the long ncRNAs such as Xist and HOTAIR. The RNA can be small (*e.g.*, less than 200 nucleic acid bases in length) or large (*e.g.*, RNA greater than 200 nucleic acid bases in length). Examples of small RNAs include 5.8S ribosomal RNA (rRNA), 5S rRNA, tRNA, miRNA, siRNA, snoRNAs, piRNA, tRNA-derived small RNA (tsRNA), and small rDNA-derived RNA (srRNA). The RNA can be double-stranded RNA or single-stranded RNA. The RNA can be circular RNA. The RNA can be a bacterial rRNA (*e.g.*, 16s rRNA or 23s rRNA).

[0070] In some instances described herein, an analyte may be a denatured nucleic acid, wherein the resulting denatured nucleic acid is single-stranded. The nucleic acid may be denatured, for example, optionally using formamide, heat, or both formamide and heat. In some instances, the nucleic acid is not denatured for use in a method disclosed herein.

[0071] In certain instances, an analyte can be extracted from a live cell. Processing conditions can be adjusted to ensure that a biological sample remains live during analysis, and analytes are extracted from (or released from) live cells of the sample. Live cell-derived analytes can be obtained only once from the sample, or can be obtained at intervals from a sample that continues to remain in viable condition.

[0072] Methods and compositions disclosed herein can be used to analyze any number of analytes. For example, the number of analytes that are analyzed can be at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about

8, at least about 9, at least about 10, at least about 11, at least about 12, at least about 13, at least about 14, at least about 15, at least about 20, at least about 25, at least about 30, at least about 40, at least about 50, at least about 100, at least about 1,000, at least about 10,000, at least about 100,000 or more different analytes present in a region of the sample or within an individual feature of the substrate.

[0073] In any implementation described herein, the analyte comprises a target sequence. In some instances, the target sequence may be endogenous to the sample, generated in the sample, added to the sample, or associated with an analyte in the sample. In some instances, the target sequence is a single-stranded target sequence (*e.g.*, a sequence in a rolling circle amplification product). In some instances, the analytes comprise one or more single-stranded target sequences. In one aspect, a first single-stranded target sequence is not identical to a second single-stranded target sequence. In another aspect, a first single-stranded target sequence is identical to one or more second single-stranded target sequence. In some instances, the one or more second single-stranded target sequence is comprised in the same analyte (*e.g.*, nucleic acid) as the first single-stranded target sequence. Alternatively, the one or more second single-stranded target sequence is comprised in a different analyte (*e.g.*, nucleic acid) from the first single-stranded target sequence.

Labelling agents:

[0074] In some instances, provided herein are methods and compositions for analyzing endogenous analytes (*e.g.*, RNA, ssDNA, and cell surface or intracellular proteins and/or metabolites) in a sample using one or more labelling agents. In some instances, an analyte labelling agent may include an agent that interacts with an analyte (*e.g.*, an endogenous analyte in a sample). In some instances, the labelling agents can comprise a reporter oligonucleotide that is indicative of the analyte or portion thereof interacting with the labelling agent. For example, the reporter oligonucleotide may comprise a barcode sequence that permits identification of the labelling agent. In some cases, the sample contacted by the labelling agent can be further contacted with a probe (*e.g.*, a single-stranded probe sequence), that hybridizes to a reporter oligonucleotide of the labelling agent, in order to identify the analyte associated with the labelling agent. In some instances, the analyte labelling agent comprises an analyte binding moiety and a labelling agent barcode domain comprising one or more barcode sequences, *e.g.*, a barcode sequence that corresponds to the analyte binding moiety and/or the analyte. An analyte binding moiety barcode includes to a barcode that is associated with or otherwise identifies the analyte binding moiety. In some instances, by

identifying an analyte binding moiety by identifying its associated analyte binding moiety barcode, the analyte to which the analyte binding moiety binds can also be identified. An analyte binding moiety barcode can be a nucleic acid sequence of a given length and/or sequence that is associated with the analyte binding moiety. An analyte binding moiety barcode can generally include any of the variety of aspects of barcodes described herein.

[0075] In some instances, the method comprises one or more post-fixing (also referred to as post-fixation) steps after contacting the sample with one or more labelling agents.

[0076] In the methods and systems described herein, one or more labelling agents capable of binding to or otherwise coupling to one or more features may be used to characterize analytes, cells and/or cell features. In some instances, cell features include cell surface features. Analytes may include, but are not limited to, a protein, a receptor, an antigen, a surface protein, a transmembrane protein, a cluster of differentiation protein, a protein channel, a protein pump, a carrier protein, a phospholipid, a glycoprotein, a glycolipid, a cell-cell interaction protein complex, an antigen-presenting complex, a major histocompatibility complex, an engineered T-cell receptor, a T-cell receptor, a B-cell receptor, a chimeric antigen receptor, a gap junction, an adherens junction, or any combination thereof. In some instances, cell features may include intracellular analytes, such as proteins, protein modifications (*e.g.*, phosphorylation status or other post-translational modifications), nuclear proteins, nuclear membrane proteins, or any combination thereof.

[0077] In some instances, an analyte binding moiety may include any molecule or moiety capable of binding to an analyte (*e.g.*, a biological analyte, *e.g.*, a macromolecular constituent). A labelling agent may include, but is not limited to, a protein, a peptide, an antibody (or an epitope binding fragment thereof), a lipophilic moiety (such as cholesterol), a cell surface receptor binding molecule, a receptor ligand, a small molecule, a bi-specific antibody, a bi-specific T-cell engager, a T-cell receptor engager, a B-cell receptor engager, a pro-body, an aptamer, a monobody, an affimer, a darpin, and a protein scaffold, or any combination thereof. The labelling agents can include (*e.g.*, are attached to) a reporter oligonucleotide that is indicative of the cell surface feature to which the binding group binds. For example, the reporter oligonucleotide may comprise a barcode sequence that permits identification of the labelling agent. For example, a labelling agent that is specific to one type of cell feature (*e.g.*, a first cell surface feature) may have coupled thereto a first reporter oligonucleotide, while a labelling agent that is specific to a different cell feature (*e.g.*, a second cell surface feature) may have a different reporter oligonucleotide coupled thereto. For a description of exemplary labelling agents, reporter oligonucleotides, and methods of

use, see, *e.g.*, U.S. Pat. 10,550,429; U.S. Pat. Pub. 20190177800; and U.S. Pat. Pub. 20190367969, which are each incorporated by reference herein in their entirety.

[0078] In some instances, an analyte binding moiety includes one or more antibodies or antigen binding fragments thereof. The antibodies or antigen binding fragments including the analyte binding moiety can specifically bind to a target analyte. In some instances, the analyte is a protein (*e.g.*, a protein on a surface of the biological sample (*e.g.*, a cell) or an intracellular protein). In some instances, a plurality of analyte labelling agents comprising a plurality of analyte binding moieties bind a plurality of analytes present in a biological sample. In some instances, the plurality of analytes includes a single species of analyte (*e.g.*, a single species of polypeptide). In some instances in which the plurality of analytes includes a single species of analyte, the analyte binding moieties of the plurality of analyte labelling agents are the same. In some instances in which the plurality of analytes includes a single species of analyte, the analyte binding moieties of the plurality of analyte labelling agents are the different (*e.g.*, members of the plurality of analyte labelling agents can have two or more species of analyte binding moieties, wherein each of the two or more species of analyte binding moieties binds a single species of analyte, *e.g.*, at different binding sites). In some instances, the plurality of analytes includes multiple different species of analyte (*e.g.*, multiple different species of polypeptides).

[0079] In other instances, *e.g.*, to facilitate sample multiplexing, a labelling agent that is specific to a particular cell feature may have a first plurality of the labelling agent (*e.g.*, an antibody or lipophilic moiety) coupled to a first reporter oligonucleotide and a second plurality of the labelling agent coupled to a second reporter oligonucleotide.

[0080] In some aspects, these reporter oligonucleotides may comprise nucleic acid barcode sequences that permit identification of the labelling agent which the reporter oligonucleotide is coupled to. The selection of oligonucleotides as the reporter may provide advantages of being able to generate significant diversity in terms of sequence, while also being readily attachable to most biomolecules, *e.g.*, antibodies, *etc.*, as well as being readily detected.

[0081] Attachment (coupling) of the reporter oligonucleotides to the labelling agents may be achieved through any of a variety of direct or indirect, covalent or non-covalent associations or attachments. For example, oligonucleotides may be covalently attached to a portion of a labelling agent (such a protein, *e.g.*, an antibody or antibody fragment) using chemical conjugation techniques (*e.g.*, Lightning-Link® antibody labelling kits available from Innova Biosciences), as well as other non-covalent attachment mechanisms, *e.g.*, using biotinylated antibodies and oligonucleotides (or beads that include one or more biotinylated linker,

coupled to oligonucleotides) with an avidin or streptavidin linker. Antibody and oligonucleotide biotinylation techniques are available. See, *e.g.*, Fang, *et al.*, "Fluoride-Cleavable Biotinylation Phosphoramidite for 5′-end-Labelling and Affinity Purification of Synthetic Oligonucleotides," Nucleic Acids Res. Jan. 15, 2003; 31(2):708-715, which is entirely incorporated herein by reference for all purposes. Likewise, protein and peptide biotinylation techniques have been developed and are readily available. See, *e.g.*, U.S. Pat. No. 6,265,552, which is entirely incorporated herein by reference for all purposes. Furthermore, click reaction chemistry may be used to couple reporter oligonucleotides to labelling agents. Commercially available kits, such as those from Thunderlink and Abcam, and techniques common in the art may be used to couple reporter oligonucleotides to labelling agents as appropriate. In another example, a labelling agent is indirectly (*e.g.*, via hybridization) coupled to a reporter oligonucleotide comprising a barcode sequence that identifies the label agent. For instance, the labelling agent may be directly coupled (*e.g.*, covalently bound) to a hybridization oligonucleotide that comprises a sequence that hybridizes with a sequence of the reporter oligonucleotide. Hybridization of the hybridization oligonucleotide to the reporter oligonucleotide couples the labelling agent to the reporter oligonucleotide. In some instances, the reporter oligonucleotides are releasable from the labelling agent, such as upon application of a stimulus. For example, the reporter oligonucleotide may be attached to the labelling agent through a labile bond (*e.g.*, chemically labile, photolabile, thermally labile, *etc.*) as generally described for releasing molecules from supports elsewhere herein. In some instances, the reporter oligonucleotides described herein may include one or more functional sequences that can be used in subsequent processing, such as an adapter sequence, a unique molecular identifier (UMI) sequence, a sequencer specific flow cell attachment sequence (such as an P5, P7, or partial P5 or P7 sequence), a primer or primer binding sequence, a sequencing primer or primer binding sequence (such as an R1, R2, or partial R1 or R2 sequence).

[0082] In some cases, the labelling agent can comprise a reporter oligonucleotide and a label. A label can be fluorophore, a radioisotope, a molecule capable of a colorimetric reaction, a magnetic particle, or any other suitable molecule or compound capable of detection. The label can be conjugated to a labelling agent (or reporter oligonucleotide) either directly or indirectly (*e.g.*, the label can be conjugated to a molecule that can bind to the labelling agent or reporter oligonucleotide). In some cases, a label is conjugated to a first oligonucleotide that is complementary (*e.g.*, hybridizes) to a sequence of the reporter oligonucleotide.

[0083] In some instances, multiple different species of analytes (*e.g.*, polypeptides) from the biological sample can be subsequently associated with the one or more physical properties of the biological sample. For example, the multiple different species of analytes can be associated with locations of the analytes in the biological sample. Such information (*e.g.*, proteomic information when the analyte binding moiety(ies) recognizes a polypeptide(s)) can be used in association with other spatial information (*e.g.*, genetic information from the biological sample, such as DNA sequence information, transcriptome information (*i.e.*, sequences of transcripts), or both). For example, a cell surface protein of a cell can be associated with one or more physical properties of the cell (*e.g.*, a shape, size, activity, or a type of the cell). The one or more physical properties can be characterized by imaging the cell. The cell can be bound by an analyte labelling agent comprising an analyte binding moiety that binds to the cell surface protein and an analyte binding moiety barcode that identifies that analyte binding moiety. Results of protein analysis in a sample (*e.g.*, a tissue sample or a cell) can be associated with DNA and/or RNA analysis in the sample.

Assays for in situ detection and analysis:

[0084] Objectives for *in situ* detection and analysis methods include detecting, quantifying, and/or mapping analytes (*e.g.*, gene activity) to specific regions in a biological sample (*e.g.*, a tissue sample or cells deposited on a surface) at cellular or sub-cellular resolution. Methods for performing *in situ* studies include a variety of techniques, *e.g.*, *in situ* hybridization and *in situ* sequencing techniques. These techniques allow one to study the subcellular distribution of target analytes (*e.g.*, gene activity as evidenced, *e.g.*, by expressed gene transcripts), and have the potential to provide crucial insights in the fields of developmental biology, oncology, immunology, histology, *etc.*

[0085] Various methods can be used for *in situ* detection and analysis of target analytes, *e.g.*, sequencing by synthesis (SBS), sequencing by ligation (SBL), sequencing by hybridization (SBH). Non-limiting examples of *in situ* hybridization techniques include single molecule fluorescence *in situ* hybridization (smFISH) and multiplexed error-robust fluorescence *in situ* hybridization (MERFISH). smFISH enables *in situ* detection and quantification of gene transcripts in tissue samples at the locations where they reside by making use of libraries of multiple short oligonucleotide probes (*e.g.*, approximately 20 base pairs (bp) in length), each labeled with a fluorophore. The probes are sequentially hybridized to gene sequences (*e.g.*, DNA) or gene transcript sequences (*e.g.*, mRNA) sequences, and visualized as diffraction-limited spots by fluorescence microscopy (Levsky, *et al.* (2003) "Fluorescence *In situ*

Hybridization: Past, Present and Future", Journal of Cell Science 116(14):2833-2838; Raj, *et al.* (2008) "Imaging Individual mRNA Molecules Using Multiple Singly Labeled Probes", Nat Methods 5(10): 877–879; Moor, *et al.* (2016), ibid.). Variations on the smFISH method include, for example, the use of combinatorial labelling schemes to improve multiplexing capability (Levsky, *et al.* (2003), ibid.), the use of smFISH in combination with super-resolution microscopy (Lubeck, *et al.* (2014) "Single-Cell *In situ* RNA Profiling by Sequential Hybridization", Nature Methods 11(4):360–361).

[0086] MERFISH addresses two of the limitations of earlier *in situ* hybridization approaches, namely the limited number of target sequences that could be simultaneously identified and the robustness of the approach to readout errors caused by the stochastic nature of the hybridization process (Moor, *et al.* (2016), ibid.). MERFISH utilizes a binary barcoding scheme in which the probed target mRNA sequences are either fluorescence positive or fluorescence negative for any given imaging cycle (Ke, *et al.* (2016), ibid.; Moffitt, *et al.* (2016) "RNA Imaging with Multiplexed Error Robust Fluorescence *In situ* Hybridization", *Methods Enzymol.* 572:1–49). The encoding probes that contain a combination of target-specific hybridization sequence regions and barcoded readout sequence regions are first hybridized to the target mRNA sequences. In each imaging cycle, a subset of fluorophore-conjugated readout probes is hybridized to a subset of encoding probes. Target mRNA sequences that fluoresce in a given cycle are assigned a value of "1" and the remaining target mRNA sequences are assigned a value of "0". Between imaging cycles, the fluorescent probes from the previous cycle are photobleached. After, *e.g.*, 14 or 16 rounds of readout probe hybridization and imaging, unique combinations of the detected fluorescence signals generate a 14-bit or 16-bit code that identifies the different gene transcripts. To address the increased error rate for correctly calling the readout codes increases as the number of hybridization and imaging cycles increases, the method may also entail the use of Hamming distances for barcode design and correction of decoding errors (see., *e.g.*, Buschmann, *et al.* (2013) "Levenshtein Error-Correcting Barcodes for Multiplexed DNA Sequencing", *Bioinformatics* 14:272), thereby resulting in an error-robust barcoding scheme.

[0087] Some *in situ* sequencing techniques generally comprise both *in situ* target capture (*e.g.*, of mRNA sequences) and *in situ* sequencing. Non-limiting examples of *in situ* sequencing techniques include *in situ* sequencing with padlock probes (ISS-PLP), fluorescent *in situ* sequencing (FISSEQ), barcode *in situ* targeted sequencing (Barista-Seq), and spatially-resolved transcript amplicon readout mapping (STARmap) (see, *e.g.*, Ke, *et al.* (2016), ibid., Asp, *et al.* (2020), ibid.).

[0088] Some methods for *in situ* detection and analysis of analytes utilize a probe (*e.g.*, padlock or circular probe) that detects specific target analytes. The *in situ* sequencing using padlock probes (ISS-PLP) method, for example, combines padlock probing to target specific gene transcripts, rolling-circle amplification (RCA), and sequencing by ligation (SBL) chemistry. Within intact tissue sections, reverse transcription primers are hybridized to target sequence (*e.g.*, mRNA sequences) and reverse transcription is performed to create cDNA to which a padlock probe (a single-stranded DNA molecule comprising regions that are complementary to the target cDNA) can bind (see, *e.g.*, Asp, *et al.* (2020), ibid.). In one variation of the method, the padlock probe binds to the cDNA target with a gap remaining between the ends which is then filled in using a DNA polymerization reaction. In another variation of the method, the ends of the bound padlock probe are adjacent to each other. The ends are then ligated to create a circular DNA molecule. Target amplification using rolling-circle amplification (RCA) results in micrometer-sized RCA products (RCPs), containing a plurality of concatenated repeats of the probe sequence. In some examples, RCPs are then subjected to, *e.g.*, sequencing-by-ligation (SBL) or sequencing-by-hybridization (SBH). In some cases, the method allows for a barcode located within the probe to be decoded.

Products of endogenous analytes and/or labelling agents:

[0089] In some instances, provided herein are methods and compositions for analyzing one or more products of an endogenous analyte and/or a labelling agent in a biological sample. In some instances, an endogenous analyte (*e.g.*, a viral or cellular DNA or RNA) or a product (*e.g.*, a hybridization product, a ligation product, an extension product (*e.g.*, by a DNA or RNA polymerase), a replication product, a transcription/reverse transcription product, and/or an amplification product such as a rolling circle amplification (RCA) product) thereof is analyzed. In some instances, a labelling agent that directly or indirectly binds to an analyte in the biological sample is analyzed. In some instances, a product (*e.g.*, a hybridization product, a ligation product, an extension product (*e.g.*, by a DNA or RNA polymerase), a replication product, a transcription/reverse transcription product, and/or an amplification product such as a rolling circle amplification (RCA) product) of a labelling agent that directly or indirectly binds to an analyte in the biological sample is analyzed.

[0090] In some instances, the analyzing comprises using primary probes which comprise a target binding region (*e.g.*, a region that binds to a target such as RNA transcripts) and the primary probes may contain one or more barcodes (*e.g.*, primary barcode). In some instances, the barcodes are bound by detection primary probes, which do not need to be

fluorescent, but that include a target-binding portion (*e.g.*, for hybridizing to one or more primary probes) and one or more barcodes (*e.g.*, secondary barcodes). In some instances, the detection primary probe comprises an overhang that does not hybridize to the target nucleic acid but hybridizes to another probe. In some examples, the overhang comprises the barcode(s). In some instances, the barcodes of the detection primary probes are targeted by detectably labeled detection oligonucleotides, such as fluorescently labeled oligos. In some instances, one or more decoding schemes are used to decode the signals, such as fluorescence, for sequence determination. Various probes and probe sets can be used to hybridize to and detect an endogenous analyte and/or a sequence associated with a labelling agent. In some instances, these assays may enable multiplexed detection, signal amplification, combinatorial decoding, and error correction schemes. Exemplary barcoded probes or probe sets may be based on a padlock probe, a gapped padlock probe, a SNAIL (Splint Nucleotide Assisted Intramolecular Ligation) probe set, a PLAYR (Proximity Ligation Assay for RNA) probe set, a PLISH (Proximity Ligation *in situ* Hybridization) probe set. The specific probe or probe set design can vary.

Hybridization and ligation:

[0091] Various probes and probe sets can be hybridized to an endogenous analyte and/or a labelling agent and each probe may comprise one or more barcode sequences. The specific probe or probe set design can vary. In some instances, the hybridization of a primary probe or probe set (*e.g.*, a circularizable probe or probe set) to a target nucleic acid analyte and may lead to the generation of a rolling circle amplification (RCA) template. In some instances, the assay uses or generates a circular nucleic acid molecule which can be the RCA template.

[0092] In some instances, a product of an endogenous analyte and/or a labelling agent is a ligation product. In some instances, the ligation product is formed from circularization of a circularizable probe or probe set upon hybridization to a target sequence. In some instances, the ligation product is formed between two or more endogenous analytes. In some instances, the ligation product is formed between an endogenous analyte and a labelling agent. In some instances, the ligation product is formed between two or more labelling agent. In some instances, the ligation product is an intramolecular ligation of an endogenous analyte. In some instances, the ligation product is an intramolecular ligation of a labelling agent, for example, the circularization of a circularizable probe or probe set upon hybridization to a target sequence. The target sequence can be comprised in an endogenous analyte (*e.g.*, nucleic acid such as a genomic DNA or mRNA) or a product thereof (*e.g.*, cDNA from a

cellular mRNA transcript), or in a labelling agent (*e.g.*, the reporter oligonucleotide) or a product thereof.

[0093] In some instances, provided herein is a probe or probe set capable of DNA-templated ligation, such as from a cDNA molecule. See, *e.g.*, U.S. Pat. 8,551,710, which is hereby incorporated by reference in its entirety. In some instances, provided herein is a probe or probe set capable of RNA-templated ligation. See, *e.g.*, U.S. Pat. Pub. 2020/0224244 which is hereby incorporated by reference in its entirety. In some instances, the probe set is a SNAIL probe set. See, *e.g.*, U.S. Pat. Pub. 20190055594, which is hereby incorporated by reference in its entirety. In some instances, provided herein is a multiplexed proximity ligation assay. See, *e.g.*, U.S. Pat. Pub. 20140194311 which is hereby incorporated by reference in its entirety. In some instances, provided herein is a probe or probe set capable of proximity ligation, for instance a proximity ligation assay for RNA (*e.g.*, PLAYR) probe set. See, *e.g.*, U.S. Pat. Pub. 20160108458, which is hereby incorporated by reference in its entirety. In some instances, a circular probe can be indirectly hybridized to the target nucleic acid. In some instances, the circular construct is formed from a probe set capable of proximity ligation, for instance a proximity ligation *in situ* hybridization (PLISH) probe set. See, *e.g.*, U.S. Pat. Pub. 2020/0224243 which is hereby incorporated by reference in its entirety.

[0094] In some instances, the ligation involves chemical ligation. In some instances, the ligation involves template dependent ligation. In some instances, the ligation involves template independent ligation. In some instances, the ligation involves enzymatic ligation.

[0095] In some instances, the enzymatic ligation involves use of a ligase. In some aspects, the ligase used herein comprises an enzyme that is commonly used to join polynucleotides together or to join the ends of a single polynucleotide. An RNA ligase, a DNA ligase, or another variety of ligase can be used to ligate two nucleotide sequences together. Ligases comprise ATP-dependent double-strand polynucleotide ligases, NAD-i-dependent double-strand DNA or RNA ligases and single-strand polynucleotide ligases, for example any of the ligases described in EC 6.5.1.1 (ATP-dependent ligases), EC 6.5.1.2 (NAD+-dependent ligases), EC 6.5.1.3 (RNA ligases). Specific examples of ligases comprise bacterial ligases such as E. coli DNA ligase, Tth DNA ligase, Thermococcus sp. (strain 9° N) DNA ligase (9°N™ DNA ligase, New England Biolabs), Taq DNA ligase, Ampligase™ (Epicentre Biotechnologies) and phage ligases such as T3 DNA ligase, T4 DNA ligase and T7 DNA ligase and mutants thereof. In some instances, the ligase is a T4 RNA ligase. In some instances, the ligase is a splintR ligase. In some instances, the ligase is a single stranded DNA ligase. In some instances, the ligase is a T4 DNA ligase. In some instances, the ligase is a

ligase that has an DNA-splinted DNA ligase activity. In some instances, the ligase is a ligase that has an RNA-splinted DNA ligase activity.

[0096] In some instances, the ligation herein is a direct ligation. In some instances, the ligation herein is an indirect ligation. "Direct ligation" means that the ends of the polynucleotides hybridize immediately adjacently to one another to form a substrate for a ligase enzyme resulting in their ligation to each other (intramolecular ligation). Alternatively, "indirect" means that the ends of the polynucleotides hybridize non-adjacently to one another, *i.e.*, separated by one or more intervening nucleotides or "gaps". In some instances, said ends are not ligated directly to each other, but instead occurs either via the intermediacy of one or more intervening (so-called "gap" or "gap-filling" (oligo)nucleotides) or by the extension of the 3' end of a probe to "fill" the "gap" corresponding to said intervening nucleotides (intermolecular ligation). In some cases, the gap of one or more nucleotides between the hybridized ends of the polynucleotides may be "filled" by one or more "gap" (oligo)nucleotide(s) which are complementary to a splint, padlock probe, or target nucleic acid. The gap may be a gap of 1 to 60 nucleotides or a gap of 1 to 40 nucleotides or a gap of 3 to 40 nucleotides. In specific implementations, the gap may be a gap of about 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 or more nucleotides, of any integer (or range of integers) of nucleotides in between the indicated values. In some instances, the gap between said terminal regions may be filled by a gap oligonucleotide or by extending the 3' end of a polynucleotide. In some cases, ligation involves ligating the ends of the probe to at least one gap (oligo)nucleotide, such that the gap (oligo)nucleotide becomes incorporated into the resulting polynucleotide. In some instances, the ligation herein is preceded by gap filling. In other implementations, the ligation herein does not require gap filling.

[0097] In some instances, ligation of the polynucleotides produces polynucleotides with melting temperature higher than that of un-ligated polynucleotides. Thus, in some aspects, ligation stabilizes the hybridization complex containing the ligated polynucleotides prior to subsequent steps, comprising amplification and detection.

[0098] In some aspects, a high fidelity ligase, such as a thermostable DNA ligase (*e.g.*, a Taq DNA ligase), is used. Thermostable DNA ligases are active at elevated temperatures, allowing further discrimination by incubating the ligation at a temperature near the melting temperature (Tm) of the DNA strands. This selectively reduces the concentration of annealed mismatched substrates (expected to have a slightly lower Tm around the mismatch) over annealed fully base-paired substrates. Thus, high-fidelity ligation can be achieved through a

combination of the intrinsic selectivity of the ligase active site and balanced conditions to reduce the incidence of annealed mismatched dsDNA.

[0099] In some instances, the ligation herein is a proximity ligation of ligating two (or more) nucleic acid sequences that are in proximity with each other, *e.g.*, through enzymatic means (*e.g.*, a ligase). In some instances, proximity ligation can include a "gap-filling" step that involves incorporation of one or more nucleic acids by a polymerase, based on the nucleic acid sequence of a template nucleic acid molecule, spanning a distance between the two nucleic acid molecules of interest (see, *e.g.*, U.S. Patent No. 7,264,929, the entire contents of which are incorporated herein by reference). A wide variety of different methods can be used for proximity ligating nucleic acid molecules, including (but not limited to) "sticky-end" and "blunt-end" ligations. Additionally, single-stranded ligation can be used to perform proximity ligation on a single-stranded nucleic acid molecule. Sticky-end proximity ligations involve the hybridization of complementary single-stranded sequences between the two nucleic acid molecules to be joined, prior to the ligation event itself. Blunt-end proximity ligations generally do not include hybridization of complementary regions from each nucleic acid molecule because both nucleic acid molecules lack a single-stranded overhang at the site of ligation.

Primer extension and amplification:

[0100] In some instances, the hybridization of a primary probe or probe set (*e.g.* a circularizable probe or probe set) to a target analyte and may lead to the generation of an extension or amplification product. In some instances, a product is a primer extension product of an analyte, a labelling agent, a probe or probe set bound to the analyte (*e.g.*, a circularizable probe bound to genomic DNA, mRNA, or cDNA), or a probe or probe set bound to the labelling agent (*e.g.*, a circularizable probe bound to one or more reporter oligonucleotides from the same or different labelling agents.

[0101] A primer is generally a single-stranded nucleic acid sequence having a 3' end that can be used as a substrate for a nucleic acid polymerase in a nucleic acid extension reaction. RNA primers are formed of RNA nucleotides, and are used in RNA synthesis, while DNA primers are formed of DNA nucleotides and used in DNA synthesis. Primers can also include both RNA nucleotides and DNA nucleotides (*e.g.*, in a random or designed pattern). Primers can also include other natural or synthetic nucleotides described herein that can have additional functionality. In some examples, DNA primers can be used to prime RNA synthesis and vice versa (*e.g.*, RNA primers can be used to prime DNA synthesis). Primers can vary in length.

For example, primers can be about 6 bases to about 120 bases. For example, primers can include up to about 25 bases. A primer, may in some cases, refer to a primer binding sequence. A primer extension reaction generally refers to any method where two nucleic acid sequences become linked (*e.g.*, hybridized) by an overlap of their respective terminal complementary nucleic acid sequences (*i.e.*, for example, 3' termini). Such linking can be followed by nucleic acid extension (*e.g.*, an enzymatic extension) of one, or both termini using the other nucleic acid sequence as a template for extension. Enzymatic extension can be performed by an enzyme including, but not limited to, a polymerase and/or a reverse transcriptase.

[0102] In some instances, a product of an endogenous analyte and/or a labelling agent is an amplification product of one or more polynucleotides, for instance, a circular probe or circularizable probe or probe set. In some instances, the disclosed methods may comprise the use of a rolling circle amplification (RCA) technique to amplify signal. Rolling circle amplification is an isothermal, DNA polymerase-mediated process in which long single-stranded DNA molecules are synthesized on a short circular single-stranded DNA template using a single DNA primer (Zhao, *et al.* (2008), "Rolling Circle Amplification: Applications in Nanotechnology and Biodetection with Functional Nucleic Acids", *Angew Chem Int Ed Engl.* 47(34):6330-6337; Ali, *et al.* (2014), "Rolling Circle Amplification: A Versatile Tool for Chemical Biology, Materials Science and Medicine", *Chem Soc Rev.* 43(10):3324-3341). The RCA product is a concatemer containing tens to hundreds of tandem repeats that are complementary to the circular template, and may be used to develop sensitive techniques for the detection of a variety of targets, including nucleic acids (DNA, RNA), small molecules, proteins, and cells (Ali, *et al.* (2014), ibid.). In some implementations, a primer that hybridizes to the circular probe or circularized probe is added and used as such for amplification. In some instances, the RCA comprises a linear RCA, a branched RCA, a dendritic RCA, or any combination thereof.

[0103] In some instances, the amplification is performed at a temperature between or between about 20°C and about 60°C. In some instances, the amplification is performed at a temperature between or between about 30°C and about 40°C. In some aspects, the amplification step, such as the rolling circle amplification (RCA) is performed at a temperature between at or about 25°C and at or about 50°C, such as at or about 25°C, 27°C, 29°C, 31°C, 33°C, 35°C, 37°C, 39°C, 41°C, 43°C, 45°C, 47°C, or 49°C.

[0104] In some instances, upon addition of a DNA polymerase in the presence of appropriate dNTP precursors and other cofactors, a primer is elongated to produce multiple copies of the

circular template. This amplification step can utilize isothermal amplification or non-isothermal amplification. In some instances, after the formation of the hybridization complex and association of the amplification probe, the hybridization complex is rolling-circle amplified to generate a cDNA nanoball (*i.e.*, amplicon) containing multiple copies of the cDNA. Techniques for rolling circle amplification (RCA) are known in the art such as linear RCA, a branched RCA, a dendritic RCA, or any combination thereof. (See, *e.g.*, Baner et al, Nucleic Acids Research, 26:5073-5078, 1998; Lizardi et al, Nature Genetics 19:226, 1998; Mohsen *et al.*, Acc Chem Res. 2016 November 15; 49(11): 2540–2550; Schweitzer *et al.* Proc. Natl Acad. Sci. USA 97:101 13- 1 19, 2000; Faruqi et al, BMC Genomics 2:4, 2000; Nallur et al, Nucl. Acids Res. 29:el 18, 2001; Dean *et al.* Genome Res. 1 1 :l095- 1099, 2001; Schweitzer et al, Nature Biotech. 20:359-365, 2002; U.S. Patent Nos. 6,054,274, 6,291,187, 6,323,009, 6,344,329 and 6,368,801). Exemplary polymerases for use in RCA comprise DNA polymerase such phi29 (φ29) polymerase, Klenow fragment, *Bacillus stearothermophilus* DNA polymerase (BST), T4 DNA polymerase, T7 DNA polymerase, or DNA polymerase I. In some aspects, DNA polymerases that have been engineered or mutated to have desirable characteristics can be employed. In some instances, the polymerase is phi29 DNA polymerase.

[0105] In some aspects, during the amplification step, modified nucleotides can be added to the reaction to incorporate the modified nucleotides in the amplification product (*e.g.*, nanoball). Exemplary of the modified nucleotides comprise amine-modified nucleotides. In some aspects of the methods, for example, for anchoring or cross-linking of the generated amplification product (*e.g.*, nanoball) to a scaffold, to cellular structures and/or to other amplification products (*e.g.*, other nanoballs). In some aspects, the amplification products comprises a modified nucleotide, such as an amine-modified nucleotide. In some instances, the amine-modified nucleotide comprises an acrylic acid N- hydroxysuccinimide moiety modification. Examples of other amine-modified nucleotides comprise, but are not limited to, a 5-Aminoallyl-dUTP moiety modification, a 5-Propargylamino-dCTP moiety modification, a N6-6-Aminohexyl-dATP moiety modification, or a 7-Deaza-7-Propargylamino-dATP moiety modification.

[0106] In some instances, the RCA template may comprise the target analyte, or a part thereof, where the target analyte is a nucleic acid, or it may be provided or generated as a proxy, or a marker, for the analyte. In some instances, the RCA template may comprise a sequence of the probes and probe sets hybridized to an endogenous analyte and/or a labelling agent. In some instances, the amplification product can be generated as a proxy, or a marker,

for the analyte. As noted above, many assays are known for the detection of numerous different analytes, which use a RCA-based detection system, *e.g.*, where the signal is provided by generating a RCP from a circular RCA template which is provided or generated in the assay, and the RCP is detected to detect the analyte. The RCP may thus be regarded as a reporter which is detected to detect the target analyte. However, the RCA template may also be regarded as a reporter for the target analyte; the RCP is generated based on the RCA template, and comprises complementary copies of the RCA template. The RCA template determines the signal which is detected, and is thus indicative of the target analyte. As will be described in more detail below, the RCA template may be a probe, or a part or component of a probe, or may be generated from a probe, or it may be a component of a detection assay (*i.e.* a reagent in a detection assay), which is used as a reporter for the assay, or a part of a reporter, or signal-generation system. The RCA template used to generate the RCP may thus be a circular (*e.g.* circularized) reporter nucleic acid molecule, namely from any RCA-based detection assay which uses or generates a circular nucleic acid molecule as a reporter for the assay. Since the RCA template generates the RCP reporter, it may be viewed as part of the reporter system for the assay.

[0107] In some instances, an assay may detect a product herein that includes a molecule or a complex generated in a series of reactions, *e.g.*, hybridization, ligation, extension, replication, transcription/reverse transcription, and/or amplification (*e.g.*, rolling circle amplification), in any suitable combination. For example, a product comprising a target sequence for a probe disclosed herein (*e.g.*, a bridge probe or L-probe) may be a hybridization complex formed of a cellular nucleic acid in a sample and an exogenously added nucleic acid probe. The exogenously added nucleic acid probe may comprise an overhang that does not hybridize to the cellular nucleic acid but hybridizes to another probe (*e.g.*, a detection probe). The exogenously added nucleic acid probe may be optionally ligated to a cellular nucleic acid molecule or another exogenous nucleic acid molecule. In other examples, a product comprising a target sequence for a probe disclosed herein (*e.g.*, an anchor probe) may be an RCP of a circularizable probe or probe set which hybridizes to a cellular nucleic acid molecule (*e.g.*, genomic DNA or mRNA) or product thereof (*e.g.*, a transcript such as cDNA, a DNA-templated ligation product of two probes, or an RNA-templated ligation product of two probes). In other examples, a product comprising a target sequence for a probe disclosed herein (*e.g.*, a bridge probe or L-probe) may be a probe hybridizing to an RCP. The probe may comprise an overhang that does not hybridize to the RCP but hybridizes to another probe (*e.g.*, a detection probe).

Signal amplification methods:

[0108] In some instances, a method disclosed herein may also comprise one or more signal amplification components and detecting such signals. In some instances, the present disclosure relates to the detection of nucleic acid sequences *in situ* using probe hybridization and generation of amplified signals associated with the probes. In some instances, the target nucleic acid of a nucleic acid probe comprises multiple target sequences for nucleic acid probe hybridization, such that the signal corresponding to a barcode sequence of the nucleic acid probe is amplified by the presence of multiple nucleic acid probes hybridized to the target nucleic acid. For example, multiple sequences can be selected from a target nucleic acid such as an mRNA, such that a group of nucleic acid probes (*e.g.*, 20-50 nucleic acid probes) hybridize to the mRNA in a tiled fashion. In another example, the target nucleic acid can be an amplification product (*e.g.*, an RCA product) comprising multiple copies of a target sequence (*e.g.*, a barcode sequence of the RCA product).

[0109] Alternatively or additionally, amplification of a signal associated with a barcode sequence of a nucleic acid probe can be amplified using one or more signal amplification strategies off of an oligonucleotide probe that hybridizes to the barcode sequence. In some aspects, amplification of the signal associated with the oligonucleotide probe can reduce the number of nucleic acid probes needed to hybridize to the target nucleic acid to obtain a sufficient signal-to-noise ratio. For example, the number of nucleic acid probes to tile a target nucleic acid such as an mRNA can be reduced. In some aspects, reducing the number of nucleic acid probes tiling a target nucleic acid enables detection of shorter target nucleic acids, such as shorter mRNAs. In some instances, no more than one, two, three, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18. 19, or 20 nucleic acid probes may be hybridized to the target nucleic acid. In instances wherein the target nucleic acid is an amplification product, signal amplification off of the oligonucleotide probes may reduce the number of target sequences required for detection (*e.g.*, the length of the RCA product can be reduced).

[0110] Exemplary signal amplification methods include targeted deposition of detectable reactive molecules around the site of probe hybridization, targeted assembly of branched structures (*e.g.*, bDNA or branched assay using locked nucleic acid (LNA)), programmed *in situ* growth of concatemers by enzymatic rolling circle amplification (RCA) (*e.g.*, as described in US 2019/0055594 incorporated herein by reference), hybridization chain reaction, assembly of topologically catenated DNA structures using serial rounds of chemical ligation (clampFISH), signal amplification via hairpin-mediated concatemerization (*e.g.*, as

described in US 2020/0362398 incorporated herein by reference), *e.g.*, primer exchange reactions such as signal amplification by exchange reaction (SABER) or SABER with DNA-Exchange (Exchange-SABER). In some instances, a non-enzymatic signal amplification method may be used.

[0111] The detectable reactive molecules may comprise tyramide, such as used in tyramide signal amplification (TSA) or multiplexed catalyzed reporter deposition (CARD)-FISH. In some instances, the detectable reactive molecule may be releasable and/or cleavable from a detectable label such as a fluorophore. In some instances, a method disclosed herein comprises multiplexed analysis of a biological sample comprising consecutive cycles of probe hybridization, fluorescence imaging, and signal removal, where the signal removal comprises removing the fluorophore from a fluorophore-labeled reactive molecule (*e.g.*, tyramide). Exemplary detectable reactive reagents and methods are described in US 6,828,109, US 2019/0376956, WO 2019/236841, WO 2020/102094, WO 2020/163397, and WO 2021/067475, all of which are incorporated herein by reference in their entireties.

[0112] In some instances, hybridization chain reaction (HCR) can be used for signal amplification. HCR is an enzyme-free nucleic acid amplification based on a triggered chain of hybridization of nucleic acid molecules starting from HCR monomers, which hybridize to one another to form a nicked nucleic acid polymer. This polymer is the product of the HCR reaction which is ultimately detected in order to indicate the presence of the target analyte. HCR is described in detail in Dirks and Pierce, 2004, PNAS, 101(43), 15275-15278 and in US 7,632,641 and US 7,721,721 (see also US 2006/00234261; Chemeris et al, 2008 Doklady Biochemistry and Biophysics, 419, 53-55; Niu et al, 2010, 46, 3089-3091; Choi et al, 2010, Nat. Biotechnol. 28(11), 1208-1212; and Song et al, 2012, Analyst, 137, 1396-1401). HCR monomers typically comprise a hairpin, or other metastable nucleic acid structure. In the simplest form of HCR, two different types of stable hairpin monomer, referred to here as first and second HCR monomers, undergo a chain reaction of hybridization events to form a long nicked double-stranded DNA molecule when an "initiator" nucleic acid molecule is introduced. The HCR monomers have a hairpin structure comprising a double stranded stem region, a loop region connecting the two strands of the stem region, and a single stranded region at one end of the double stranded stem region. The single stranded region which is exposed (and which is thus available for hybridization to another molecule, *e.g.* initiator or other HCR monomer) when the monomers are in the hairpin structure may be known as the "toehold region" (or "input domain"). The first HCR monomers each further comprise a sequence which is complementary to a sequence in the exposed toehold region of the second

HCR monomers. This sequence of complementarity in the first HCR monomers may be known as the "interacting region" (or "output domain"). Similarly, the second HCR monomers each comprise an interacting region (output domain), *e.g.* a sequence which is complementary to the exposed toehold region (input domain) of the first HCR monomers. In the absence of the HCR initiator, these interacting regions are protected by the secondary structure (*e.g.* they are not exposed), and thus the hairpin monomers are stable or kinetically trapped (also referred to as "metastable"), and remain as monomers (*e.g.* preventing the system from rapidly equilibrating), because the first and second sets of HCR monomers cannot hybridize to each other. However, once the initiator is introduced, it is able to hybridize to the exposed toehold region of a first HCR monomer, and invade it, causing it to open up. This exposes the interacting region of the first HCR monomer (*e.g.* the sequence of complementarity to the toehold region of the second HCR monomers), allowing it to hybridize to and invade a second HCR monomer at the toehold region. This hybridization and invasion in turn opens up the second HCR monomer, exposing its interacting region (which is complementary to the toehold region of the first HCR monomers), and allowing it to hybridize to and invade another first HCR monomer. The reaction continues in this manner until all of the HCR monomers are exhausted (*e.g.* all of the HCR monomers are incorporated into a polymeric chain). Ultimately, this chain reaction leads to the formation of a nicked chain of alternating units of the first and second monomer species. The presence of the HCR initiator is thus required in order to trigger the HCR reaction by hybridization to and invasion of a first HCR monomer. The first and second HCR monomers are designed to hybridize to one another are thus may be defined as cognate to one another. They are also cognate to a given HCR initiator sequence. HCR monomers which interact with one another (hybridize) may be described as a set of HCR monomers or an HCR monomer, or hairpin, system.

[0113] An HCR reaction could be carried out with more than two species or types of HCR monomers. For example, a system involving three HCR monomers could be used. In such a system, each first HCR monomer may comprise an interacting region which binds to the toehold region of a second HCR monomer; each second HCR may comprise an interacting region which binds to the toehold region of a third HCR monomer; and each third HCR monomer may comprise an interacting region which binds to the toehold region of a first HCR monomer. The HCR polymerization reaction would then proceed as described above, except that the resulting product would be a polymer having a repeating unit of first, second and third monomers consecutively. Corresponding systems with larger numbers of sets of HCR monomers could readily be conceived. Branching HCR systems have also been devised

and described (see, *e.g.*, WO 2020/123742 incorporated herein by reference), and may be used in the methods herein.

[0114] In some instances, similar to HCR reactions that use hairpin monomers, linear oligo hybridization chain reaction (LO-HCR) can also be used for signal amplification. In some instances, provided herein is a method of detecting an analyte in a sample comprising: (i) performing a linear oligo hybridization chain reaction (LO-HCR), wherein an initiator is contacted with a plurality of LO-HCR monomers of at least a first and a second species to generate a polymeric LO-HCR product hybridized to a target nucleic acid molecule, wherein the first species comprises a first hybridization region complementary to the initiator and a second hybridization region complementary to the second species, wherein the first species and the second species are linear, single-stranded nucleic acid molecules; wherein the initiator is provided in one or more parts, and hybridizes directly or indirectly to or is comprised in the target nucleic acid molecule; and (ii) detecting the polymeric product, thereby detecting the analyte. In some instances, the first species and/or the second species may not comprise a hairpin structure. In some instances, the plurality of LO-HCR monomers may not comprise a metastable secondary structure. In some instances, the LO-HCR polymer may not comprise a branched structure. In some instances, performing the linear oligo hybridization chain reaction comprises contacting the target nucleic acid molecule with the initiator to provide the initiator hybridized to the target nucleic acid molecule. In any of the instances herein, the target nucleic acid molecule and/or the analyte can be an RCA product.

[0115] In some instances, detection of nucleic acids sequences *in situ* includes combination of the sequential decoding methods described herein with an assembly for branched signal amplification. In some instances, the assembly complex comprises an amplifier hybridized directly or indirectly (via one or more oligonucleotides) to a sequence of an oligonucleotide probe described herein. In some instances, the assembly includes one or more amplifiers each including an amplifier repeating sequence. In some aspects, the one or more amplifiers is labeled. Described herein is a method of using the aforementioned assembly, including for example, using the assembly in multiplexed error-robust fluorescent *in situ* hybridization (MERFISH) applications, with branched DNA amplification for signal readout. In some instances, the amplifier repeating sequence is about 5-30 nucleotides, and is repeated N times in the amplifier. In some instances, the amplifier repeating sequence is about 20 nucleotides, and is repeated at least two times in the amplifier. In some aspects, the one or more amplifier repeating sequence is labeled. For exemplary branched signal amplification, see *e.g.*, U.S. Pat. Pub. No. US20200399689A1 and Xia *et al.*, Multiplexed Detection of RNA using

MERFISH and branched DNA amplification. Scientific Reports (2019), each of which is fully incorporated by reference herein.

[0116] In some instances, an oligonucleotide probe described herein can be associated with an amplified signal by a method that comprises signal amplification by performing a primer exchange reaction (PER). In various instances, a primer with domain on its 3' end binds to a catalytic hairpin, and is extended with a new domain by a strand displacing polymerase. For example, a primer with domain 1 on its 3' ends binds to a catalytic hairpin, and is extended with a new domain 1 by a strand displacing polymerase, with repeated cycles generating a concatemer of repeated domain 1 sequences. In various instances, the strand displacing polymerase is Bst. In various instances, the catalytic hairpin includes a stopper which releases the strand displacing polymerase. In various instances, branch migration displaces the extended primer, which can then dissociate. In various instances, the primer undergoes repeated cycles to form a concatemer primer (see *e.g.*, U.S. Pat. Pub. No. US20190106733, which is incorporated herein by reference, for exemplary molecules and PER reaction components).

Barcoded analytes and detection:

[0117] A target sequence for a probe disclosed herein may be comprised in any analyte disclose herein, including an endogenous analyte (*e.g.*, a viral or cellular nucleic acid), a labelling agent, or a product generated in the biological sample using an endogenous analyte and/or a labelling agent.

[0118] In some aspects, one or more of the target sequences includes or is associated with one or more barcode(s), *e.g.*, at least two, three, four, five, six, seven, eight, nine, ten, or more barcodes. Barcodes can spatially-resolve molecular components found in biological samples, for example, within a cell or a tissue sample. A barcode can be attached to an analyte or to another moiety or structure in a reversible or irreversible manner. A barcode can be added to, for example, a fragment of a deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) sample before or during sequencing of the sample. Barcodes can allow for identification and/or quantification of individual sequencing-reads (*e.g.*, a barcode can be or can include a unique molecular identifier or "UMI"). In some aspects, a barcode comprises about 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more than 30 nucleotides.

[0119] In some instances, a barcode includes two or more sub-barcodes that together function as a single barcode. For example, a polynucleotide barcode can include two or more

polynucleotide sequences (*e.g.*, sub-barcodes) that are separated by one or more non-barcode sequences. In some instances, the one or more barcode(s) can also provide a platform for targeting functionalities, such as oligonucleotides, oligonucleotide-antibody conjugates, oligonucleotide-streptavidin conjugates, modified oligonucleotides, affinity purification, detectable moieties, enzymes, enzymes for detection assays or other functionalities, and/or for detection and identification of the polynucleotide.

[0120] In any of the preceding implementations, barcodes (*e.g.*, primary and/or secondary barcode sequences) can be analyzed (*e.g.*, detected or sequenced) using any suitable method or technique, including those described herein, such as sequencing by synthesis (SBS), sequencing by ligation (SBL), or sequencing by hybridization (SBH). In some instances, barcoding schemes and/or barcode detection schemes as described in RNA sequential probing of targets (RNA SPOTs), single-molecule fluorescent *in situ* hybridization (smFISH), multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) or sequential fluorescence *in situ* hybridization (seqFISH+) can be used. In any of the preceding implementations, the methods provided herein can include analyzing the barcodes by sequential hybridization and detection with a plurality of labelled probes (*e.g.*, detection probes (*e.g.*, detection oligos) or barcode probes). In some instances, the barcode detection steps can be performed as described in hybridization-based *in situ* sequencing (HybISS). In some instances, probes can be detected and analyzed (*e.g.*, detected or sequenced) as performed in fluorescent *in situ* sequencing (FISSEQ), or as performed in the detection steps of the spatially-resolved transcript amplicon readout mapping (STARmap) method. In some instances, signals associated with an analyte can be detected as performed in sequential fluorescent *in situ* hybridization (seqFISH).

[0121] In some instances, in a barcode-based detection method, barcode sequences are detected for identification of other molecules including nucleic acid molecules (DNA or RNA) longer than the barcode sequences themselves, as opposed to direct sequencing of the longer nucleic acid molecules. In some instances, a N-mer barcode sequence comprises 4N complexity given a sequencing read of N bases, and a much shorter sequencing read may be required for molecular identification compared to non-barcode sequencing methods such as direct sequencing. For example, 1024 molecular species may be identified using a 5-nucleotide barcode sequence (4^5=1024), whereas 8 nucleotide barcodes can be used to identify up to 65,536 molecular species, a number greater than the total number of distinct genes in the human genome. In some instances, the barcode sequences contained in the probes or RCPs are detected, rather than endogenous sequences, which can be an efficient

read-out in terms of information per cycle of sequencing. Because the barcode sequences are pre-determined, they can also be designed to feature error detection and correction mechanisms, see, *e.g.*, U.S. Pat. Pub. 20190055594 and WO2019199579A1, which are hereby incorporated by reference in their entirety.

Sequential hybridization:

[0122] In some instances, the present disclosure relates to methods and compositions for encoding and detecting analytes in a temporally sequential manner for *in situ* analysis of an analyte in a biological sample, *e.g.*, a target nucleic acid in a cell in an intact tissue. In some aspects, provided herein is a method for detecting the detectably-labeled probes, thereby generating a signal signature. In some instances, the signal signature corresponds to an analyte of the plurality of analytes. In some instances, the methods described herein are based, in part, on the development of a multiplexed biological assay and readout, in which a sample is first contacted with a plurality of nucleic acid probes comprising one or more probe types (*e.g.*, labelling agent, circularizable probe, circular probe, *etc.*), allowing the probes to directly or indirectly bind target analytes, which may then be optically detected (*e.g.*, by detectably-labeled probes) in a temporally-sequential manner. In some instances, the probes or probe sets comprising various probe types may be applied to a sample simultaneously. In some instances, the probes or probe sets comprising various probe types may be applied to a sample sequentially. In some aspects, the method comprises sequential hybridization of labelled probes to create a spatiotemporal signal signature or code that identifies the analyte.

[0123] In some aspects, provided herein is a method involving a multiplexed biological assay and readout, in which a sample is first contacted with a plurality of nucleic acid probes, allowing the probes to directly or indirectly bind target analytes, which may then be optically detected (*e.g.*, by detectably-labeled probes) in a temporally sequential manner. The plurality of nucleic acid probes themselves may be detectably-labeled and detected; in other words, the nucleic acid probes themselves serve as the detection probes. In other implementations, a nucleic acid probe itself is not directly detectably-labeled (*e.g.*, the probe itself is not conjugated to a detectable label); rather, in addition to a target binding sequence (*e.g.*, a sequence binding to a barcode sequence in an RCA product), the nucleic acid probe further comprises a sequence for detection which can be recognized by one or more detectably-labeled detection probes. In some instances, the probes or probe sets comprising various probe types may be applied to a sample simultaneously. In some instances, the probes or

probe sets comprising various probe types may be applied to a sample sequentially. In some instances, the method comprises detecting a plurality of analytes in a sample.

[0124] In some instances, the method presented herein comprises contacting the sample with a plurality of probes comprising one or more probes having distinct labels and detecting signals from the plurality of probes in a temporally sequential manner, wherein said detection generates signal signatures each comprising a temporal order of signal or absence thereof, and the signal signatures correspond to said plurality of probes that identify the corresponding analytes. In some instances, the temporal order of the signals or absence thereof corresponding to the analytes can be unique for each different analyte of interest in the sample. In some instances, the plurality of probes hybridize to an endogenous molecule in the sample, such as a cellular nucleic acid molecule, *e.g.*, genomic DNA, RNA (*e.g.*, mRNA), or cDNA. In some instances, the plurality of probes hybridize to a product of an endogenous molecule in the sample (*e.g.*, directly or indirectly via an intermediate probe). In some instances, the plurality of probes hybridize to labelling agent that binds directly or indirectly to an endogenous molecule in the sample or a product thereof. In some instances, the plurality of probes hybridize to a product (*e.g.*, an RCA product) of a labelling agent that binds directly or indirectly to an endogenous molecule in the sample or a product thereof.

[0125] In any of the implementations disclosed herein, the detection of signals can be performed sequentially in cycles, one for each distinct label. In any of the implementations disclosed herein, signals or absence thereof from detectably-labeled probes targeting an analyte in a particular location in the sample can be recorded in a first cycle for detecting a first label, and signals or absence thereof from detectably-labeled probes targeting the analyte in the particular location can be recorded in a second cycle for detecting a second label distinct from the first label. In any of the implementations disclosed herein, a unique signal signature can be generated for each analyte of the plurality of analytes. In any of the implementations disclosed herein, one or more molecules comprising the same analyte or a portion thereof can be associated with the same signal signature.

[0126] In some instances, the *in situ* assays employ strategies for optically encoding the spatial location of target analytes (*e.g.*, mRNAs) in a sample using sequential rounds of fluorescent hybridization. Microcopy may be used to analyze 4 or 5 fluorescent colors indicative of the spatial localization of a target, followed by various rounds of hybridization and stripping, in order to generate a large set of unique optical signal signatures assigned to different analytes. These methods often require a large number of hybridization rounds, and a large number of microscope lasers (*e.g.*, detection channels) to detect a large number of

fluorophores, resulting in a one to one mapping of the lasers to the fluorophores. Specifically, each detectably-labeled probe comprises one detectable moiety, *e.g.*, a fluorophore.

[0127] In some aspects, provided herein is a method for analyzing a sample using a detectably-labeled set of probes. In some instances, the method comprises contacting the sample with a first plurality of detectably-labeled probes for targeting a plurality of analytes; performing a first detection round comprising detecting signals from the first plurality of detectably-labeled probes; contacting the sample with a second plurality of detectably-labeled probes for targeting the plurality of analytes; performing a second detection round of detecting signals from the second plurality of detectably-labeled probes, thereby generating a signal signature comprising a plurality of signals detected from the first detection round and second detection round, wherein the signal signature corresponds to an analyte of the plurality of analytes.

[0128] In some instances, detection of an optical signal signature comprises several rounds of detectably-labeled probe hybridization (*e.g.*, contacting a sample with detectably-labeled probes), detectably-labeled probe detection, and detectably-labeled probe removal. In some instances, a sample is contacted with plurality first detectably-labeled probes, and said probes are hybridized to a plurality of nucleic acid analytes within the sample in decoding hybridization round 1. In some instances, a first detection round is performed following detectably-labeled probe hybridization. After hybridization and detection of a first plurality of detectably-labeled probes, probes are removed, and a sample may be contacted with a second plurality round of detectably-labeled probes targeting the analytes targeted in decoding hybridization round 1. The second plurality of detectably-labeled probes may hybridize to the same nucleic acid(s) as the first plurality of detectably-labeled probes (*e.g.*, hybridize to an identical or hybridize to new nucleic acid sequence within the same nucleic acid), or the second plurality of detectably-labeled probes may hybridize to different nucleic acid(s) compared to the first plurality of detectably-labeled probes. Following m rounds of contacting a sample with a plurality of detectably-labeled probes, probe detection, and probe removal, ultimately a unique signal signature to each nucleic acid is produced that may be used to identify and quantify said nucleic acids and the corresponding analytes (*e.g.*, if the nucleic acids themselves are not the analytes of interest and each is used as part of a labelling agent for one or more other analytes such as protein analytes and/or other nucleic acid analytes).

[0129] In some instances, after hybridization of a detectably-labeled probes (*e.g.*, fluorescently labeled oligonucleotide) that detects a sequence (*e.g.*, barcode sequence on a

secondary probe or a primary probe), and optionally washing away the unbound molecules of the detectably-labeled probe, the sample is imaged and the detection oligonucleotide or detectable label is inactivated and/or removed. In some instances, removal of the signal associated with the hybridization between rounds can be performed by washing, heating, stripping, enzymatic digestion, photo-bleaching, displacement (*e.g.*, displacement of detectably-labeled probes with another reagent or nucleic acid sequence), cleavage, quenching, chemical degradation, bleaching, oxidation, or any combinations thereof.

[0130] In some examples, removal of a probe (*e.g.*, un-hybridizing the entire probe), signal modifications (*e.g.*, quenching, masking, photo-bleaching, signal enhancement (*e.g.*, via FRET), signal amplification, *etc.*), signal removal (*e.g.*, cleaving off or permanently inactivating a detectable label) can be performed. Inactivation may be caused by removal of the detectable label (*e.g.*, from the sample, or from the probe, *etc.*), and/or by chemically altering the detectable label in some fashion, *e.g.*, by photobleaching the detectable label, bleaching or chemically altering the structure of the detectable label, *e.g.*, by reduction, *etc.*). In some instances, the fluorescently labeled oligonucleotide and/or the intermediate probe hybridized to the fluorescently labeled oligonucleotide (*e.g.*, bridge probe or L-probe) can be removed. In some instances, a fluorescent detectable label may be inactivated by chemical or optical techniques such as oxidation, photobleaching, chemically bleaching, stringent washing or enzymatic digestion or reaction by exposure to an enzyme, dissociating the detectable label from other components (*e.g.*, a probe), chemical reaction of the detectable label (*e.g.*, to a reactant able to alter the structure of the detectable label) or the like. For instance, bleaching may occur by exposure to oxygen, reducing agents, or the detectable label could be chemically cleaved from the nucleic acid probe and washed away via fluid flow.

[0131] In some instances, removal of a signal comprises displacement of probes with another reagent (*e.g.*, probe) or nucleic acid sequence. For example, a given probe (*e.g.*, detectably-labeled probes and/or the intermediate probe hybridized to the fluorescently labeled oligonucleotide (*e.g.*, bridge probe or L-probe)) may be displaced by a subsequent probe that hybridizes to an overlapping region shared between the binding sites of the probes. In some cases, a displacement reaction can be very efficient, and thus allows for probes to be switched quickly between cycles, without the need for chemical stripping (or any of the damage to the sample that is associated therewith). In some instances, a sequence for hybridizing the subsequent or displacer probe (*i.e.* a toehold sequence) may be common across a plurality of probes capable of hybridizing to a given binding site. In some aspects, a single displacement probe can be used to simultaneously displace detection probes bound to an equivalent

barcode position from all of the RCPs within a given sample simultaneously (with the displacement mediated by the subsequent detection probes). This may further increase efficiency and reduce the cost of the method, as fewer different probes are required.

[0132] After a signal is inactivated and/or removed, then the sample is re-hybridized in a subsequent round with a subsequent fluorescently labeled oligonucleotide, and the oligonucleotide can be labeled with the same color or a different color as the fluorescently labeled oligonucleotide of the previous cycle. In some instances, as the positions of the analytes, probes, and/or products thereof can be fixed (*e.g.*, via fixing and/or crosslinking) in a sample, the fluorescent spot corresponding to an analyte, probe, or product thereof remains in place during multiple rounds of hybridization and can be aligned to read out a string of signals associated with each target analyte.

Decoding:

[0133] A "decoding process" is a process comprising a plurality of decoding cycles in which different sets of barcode probes are contacted with target analytes (*e.g.*, mRNA sequences) or target barcodes (*e.g.*, barcodes associated with target analytes) present in a sample, and used to detect the target sequences or associated target barcodes, or segments thereof. In some instances, the decoding process comprises acquiring one or more images (*e.g.*, fluorescence images) for each decoding cycle. Decoded barcode sequences are then inferred based on a set of physical signals (*e.g.*, fluorescence signals) detected in each decoding cycle of a decoding process. In some instances, the set of physical signals (*e.g.*, fluorescence signals) detected in a series of decoding cycles for a given target barcode (or target analyte sequence) may be considered a "signal signature" for the target barcode (or target analyte sequence). In some instances, a decoding process may comprise, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10 decoding cycles. In some instances, each decoding cycle may comprise contacting a plurality of target sequences or target barcodes with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10 barcode probes (*e.g.*, fluorescently-labeled barcode probes) that are configured to hybridize or bind to specific target sequences or target barcodes, or segments thereof. In some instances, a decoding process may comprise performing a series of *in situ* barcode probe hybridization steps and acquiring images (*e.g.*, fluorescence images) at each step. Systems and methods for performing multiplexed fluorescence *in situ* hybridization and imaging are described in, for example, WO 2021/127019 A1; U.S. Pat. 11,021,737; and PCT/EP2020/065090 (WO2020240025A1), each of which is incorporated herein by reference in its entirety.

Anchor probes:

[0134] In some instances, the present methods may further involve contacting the target analyte, *e.g.*, a nucleic acid molecule, or proxy thereof with an anchor probe. In some instances, the anchor probe comprises a sequence complementary to an anchor probe binding region, which is present in all target nucleic acid molecules (*e.g.*, in primary or secondary probes), and a detectable label. The detection of the anchor probe via the detectable label confirms the presence of the target nucleic acid molecule. The target nucleic acid molecule may be contacted with the anchor probe prior to, concurrently with, or after being contacted with the first set of detection probes. In some instances, the target nucleic acid molecule may be contacted with the anchor probe during multiple decoding cycles. In some instances, multiple different anchor probes comprising different sequences and/or different reporters may be used to confirm the presence of multiple different target nucleic acid molecules. The use of multiple anchor probes is particularly useful when detection of a large number of target nucleic acid molecules is required, as it allows for optical crowding to be reduced and thus for detected target nucleic acid molecules to be more clearly resolved

[0135] Described herein is a method applying the aforementioned techniques to reduce optical crowding in an assay. In some instances, a method can include analyzing a biological sample, by contacting the biological sample with a first number of primary probe(s) configured to hybridize to a first target nucleic acid and a second number of primary probes configured to hybridize to a second target nucleic acid, each primary probe including a target-hybridizing region configured to hybridize to a different target region in the corresponding target nucleic acid, and a barcode region, and the first number is 1 or more, and the second number is greater than the first number, contacting the biological sample with a plurality of detectable probes, wherein each detectable probe is configured to hybridize to a barcode sequence in the barcode regions of the first number of primary probe(s) and/or the second number of primary probes, or a complement of the barcode sequence, detecting a signal associated with the plurality of detectable probes or absence thereof at one or more locations in the biological sample, and contacting the biological sample with a subsequent plurality of detectable probes, wherein each detectable probe in the subsequent plurality is configured to hybridize to a subsequent barcode sequence in the barcode regions of the first number of primary probe(s) and/or the second number of primary probes, or a complement of the subsequent barcode sequence. In some instances, wherein the second number is greater than the first number, the difference is based on absolute or relative numbers of the first target

nucleic acid and second target nucleic acid in a sample. In some instances, the difference between second and first number is linearly or non-linearly proportional to absolute or relative numbers of the first target nucleic acid and second target nucleic acid in a sample. In some instances, the difference is inversely proportional, where the first number is for first target nucleic acid present in greater amounts (*i.e.*, abundant), and the greater second number is for second target nucleic acid present in lesser amounts.

[0136] Reference will now be made in detail to implementations and embodiments of various aspects and variations of systems and methods described herein. Although several exemplary variations of the systems and methods are described herein, other variations of the systems and methods may include aspects of the systems and methods described herein combined in any suitable manner having combinations of all or some of the aspects described.

[0137] Target molecules (*e.g.*, nucleic acids, proteins, antibodies, *etc.*) can be detected in biological samples (*e.g.*, one or more cells or a tissue sample) using an instrument having integrated optics and fluidics modules (an "opto-fluidic instrument" or "opto-fluidic system"). In an opto-fluidic instrument, the fluidics module is configured to deliver one or more reagents (*e.g.*, fluorescent probes) to the biological sample and/or remove spent reagents therefrom. Additionally, the optics module is configured to illuminate the biological sample with light having one or more spectral emission curves (over a range of wavelengths) and subsequently capture one or more images of emitted light signals from the biological sample during one or more probing cycles. In various embodiments, the captured images may be processed in real time and/or at a later time to determine the presence of the one or more target molecules in the biological sample, as well as three-dimensional position information associated with each detected target molecule. Additionally, the opto-fluidics instrument includes a sample module configured to receive (and, optionally, secure) one or more biological samples. In some instances, the sample module includes an X-Y stage configured to move the biological sample along an X-Y plane (*e.g.*, perpendicular to an objective lens of the optics module).

[0138] In various embodiments, the opto-fluidic instrument is configured to analyze one or more target molecules in their naturally occurring place (*i.e.*, *in situ*) within the biological sample. For example, an opto-fluidic instrument may be an *in-situ* analysis system used to analyze a biological sample and detect target molecules including but not limited to DNA, RNA, proteins, antibodies, and/or the like.

[0139] A sample disclosed herein can be or be derived from any biological sample. Biological samples may be obtained from any suitable source using any of a variety of

techniques including, but not limited to, biopsy, surgery, and laser capture microscopy (LCM), and generally includes cells, tissues, and/or other biological material from the subject. A biological sample can be obtained from a prokaryote such as a bacterium, an archaea, a virus, or a viroid. A biological sample can also be obtained from non-mammalian organisms (e.g., a plant, an insect, an arachnid, a nematode, a fungus, or an amphibian). A biological sample can also be obtained from a eukaryote, such as a tissue sample from a mammal. A biological sample from an organism may comprise one or more other organisms or components therefrom. For example, a mammalian tissue section may comprise a prion, a viroid, a virus, a bacterium, a fungus, or components from other organisms, in addition to mammalian cells and non-cellular tissue components. Subjects from which biological samples can be obtained can be healthy or asymptomatic subjects, subjects that have or are suspected of having a disease (e.g., an individual with a disease such as cancer) or a pre-disposition to a disease, and/or subjects in need of therapy or suspected of needing therapy.

[0140] The biological sample can include any number of macromolecules, for example, cellular macromolecules and organelles (e.g., mitochondria and nuclei). The biological sample can be obtained as a tissue sample, such as a tissue section, biopsy, a core biopsy, needle aspirate, or fine needle aspirate. The sample can be a fluid sample, such as a blood sample, urine sample, or saliva sample. The sample can be a skin sample, a colon sample, a cheek swab, a histology sample, a histopathology sample, a plasma or serum sample, a tumor sample, living cells, cultured cells, a clinical sample such as, for example, whole blood or blood-derived products, blood cells, or cultured tissues or cells, including cell suspensions.

[0141] In some embodiments, the biological sample may comprise cells or a tissue sample which are deposited on a substrate. As described herein, a substrate can be any support that is insoluble in aqueous liquid and allows for positioning of biological samples, analytes, features, and/or reagents on the support. In some embodiments, a biological sample is attached to a substrate. In some embodiments, the substrate is optically transparent to facilitate analysis on the opto-fluidic instruments disclosed herein. For example, in some instances, the substrate is a glass substrate (e.g., a microscopy slide, cover slip, or other glass substrate). Attachment of the biological sample can be irreversible or reversible, depending upon the nature of the sample and subsequent steps in the analytical method. In certain embodiments, the sample can be attached to the substrate reversibly by applying a suitable polymer coating to the substrate and contacting the sample to the polymer coating. The sample can then be detached from the substrate, e.g., using an organic solvent that at least partially dissolves the polymer coating. Hydrogels are examples of polymers that are suitable

for this purpose. In some embodiments, the substrate can be coated or functionalized with one or more substances to facilitate attachment of the sample to the substrate. Suitable substances that can be used to coat or functionalize the substrate include, but are not limited to, lectins, poly-lysine, antibodies, and polysaccharides.

[0142] It is to be noted that, although the above discussion relates to an opto-fluidic instrument that can be used for *in situ* target molecule detection via probe hybridization, the discussion herein equally applies to any opto-fluidic instrument that employs any imaging or target molecule detection technique. That is, for example, an opto-fluidic instrument may include a fluidics module that includes fluids needed for establishing the experimental conditions required for the probing of target molecules in the sample. Further, such an opto-fluidic instrument may also include a sample module configured to receive the sample, and an optics module including an imaging system for illuminating (*e.g.*, exciting one or more fluorescent probes within the sample) and/or imaging light signals received from the probed sample. The *in-situ* analysis system may also include other ancillary modules configured to facilitate the operation of the opto-fluidic instrument, such as, but not limited to, cooling systems, motion calibration systems, etc.

[0143] FIG. 4 shows an example workflow of analysis of a biological sample 110 (*e.g.*, cell or tissue sample) using an opto-fluidic instrument 120, according to various embodiments. In various embodiments, the sample 110 can be a biological sample (e.g., a tissue) that includes molecules such as DNA, RNA, proteins, antibodies, etc. For example, the sample 110 can be a sectioned tissue that is treated to access the RNA thereof for labeling with circularizable DNA probes. Ligation of the probes may generate a circular DNA probe which can be enzymatically amplified and bound with fluorescent oligonucleotides, which can create bright signal that is convenient to image and has a high signal-to-noise ratio.

[0144] In various embodiments, the sample 110 may be placed in the opto-fluidic instrument 120 for analysis and detection of the molecules in the sample 110. In various embodiments, the opto-fluidic instrument 120 can be a system configured to facilitate the experimental conditions conducive for the detection of the target molecules. For example, the opto-fluidic instrument 120 can include a fluidics module 140, an optics module 150, a sample module 160, and an ancillary module 170, and these modules may be operated by a system controller 130 to create the experimental conditions for the probing of the molecules in the sample 110 by selected probes (e.g., circularizable DNA probes), as well as to facilitate the imaging of the probed sample (e.g., by an imaging system of the optics module 150). In various embodiments, the various modules of the opto-fluidic instrument 120 may be separate

components in communication with each other, or at least some of them may be integrated together.

[0145] In various embodiments, the sample module 160 may be configured to receive the sample 110 into the opto-fluidic instrument 120. For instance, the sample module 160 may include a sample interface module (SIM) that is configured to receive a sample device (e.g., cassette) onto which the sample 110 can be deposited. That is, the sample 110 may be placed in the opto-fluidic instrument 120 by depositing the sample 110 (e.g., the sectioned tissue) on a sample device that is then inserted into the SIM of the sample module 160. In some instances, the sample module 160 may also include an X-Y stage onto which the SIM is mounted. The X-Y stage may be configured to move the SIM mounted thereon (e.g., and as such the sample device containing the sample 110 inserted therein) in perpendicular directions along the two-dimensional (2D) plane of the opto-fluidic instrument 120.

[0146] The experimental conditions that are conducive for the detection of the molecules in the sample 110 may depend on the target molecule detection technique that is employed by the opto-fluidic instrument 120. For example, in various embodiments, the opto-fluidic instrument 120 can be a system that is configured to detect molecules in the sample 110 via hybridization of probes. In such cases, the experimental conditions can include molecule hybridization conditions that result in the intensity of hybridization of the target molecule (e.g., nucleic acid) to a probe (e.g., oligonucleotide) being significantly higher when the probe sequence is complementary to the target molecule than when there is a single-base mismatch. The hybridization conditions include the preparation of the sample 110 using reagents such as washing/stripping reagents, hybridizing reagents, etc., and such reagents may be provided by the fluidics module 140.

[0147] In various embodiments, the fluidics module 140 may include one or more components that may be used for storing the reagents, as well as for transporting said reagents to and from the sample device containing the sample 110. For example, the fluidics module 140 may include reservoirs configured to store the reagents, as well as a waste container configured for collecting the reagents (e.g., and other waste) after use by the opto-fluidic instrument 120 to analyze and detect the molecules of the sample 110. Further, the fluidics module 140 may also include pumps, tubes, pipettes, etc., that are configured to facilitate the transport of the reagent to the sample device (e.g., and as such the sample 110). For instance, the fluidics module 140 may include pumps (“reagent pumps”) that are configured to pump washing/stripping reagents to the sample device for use in

washing/stripping the sample 110 (e.g., as well as other washing functions such as washing an objective lens of the imaging system of the optics module 150).

[0148] In various embodiments, the ancillary module 170 can be a cooling system of the opto-fluidic instrument 120, and the cooling system may include a network of coolant-carrying tubes that are configured to transport coolants to various modules of the opto-fluidic instrument 120 for regulating the temperatures thereof. In such cases, the fluidics module 140 may include coolant reservoirs for storing the coolants and pumps (e.g., "coolant pumps") for generating a pressure differential, thereby forcing the coolants to flow from the reservoirs to the various modules of the opto-fluidic instrument 120 via the coolant-carrying tubes. In some instances, the fluidics module 140 may include returning coolant reservoirs that may be configured to receive and store returning coolants, i.e., heated coolants flowing back into the returning coolant reservoirs after absorbing heat discharged by the various modules of the opto-fluidic instrument 120. In such cases, the fluidics module 140 may also include cooling fans that are configured to force air (e.g., cool and/or ambient air) into the returning coolant reservoirs to cool the heated coolants stored therein. In some instance, the fluidics module 140 may also include cooling fans that are configured to force air directly into a component of the opto-fluidic instrument 120 so as to cool said component. For example, the fluidics module 140 may include cooling fans that are configured to direct cool or ambient air into the system controller 130 to cool the same.

[0149] As discussed above, the opto-fluidic instrument 120 may include an optics module 150 which include the various optical components of the opto-fluidic instrument 120, such as but not limited to a camera, an illumination module (e.g., light source such as LEDs), an objective lens, and/or the like. The optics module 150 may include a fluorescence imaging system that is configured to image the fluorescence emitted by the probes (e.g., oligonucleotides) in the sample 110 after the probes are excited by light from the illumination module of the optics module 150.

[0150] In some instances, the optics module 150 may also include an optical frame onto which the camera, the illumination module, and/or the X-Y stage of the sample module 160 may be mounted.

[0151] In various embodiments, the system controller 130 may be configured to control the operations of the opto-fluidic instrument 120 (e.g., and the operations of one or more modules thereof). In some instances, the system controller 130 may take various forms, including a processor, a single computer (or computer system), or multiple computers in communication with each other. In various embodiments, the system controller 130 may be

communicatively coupled with data storage, set of input devices, display system, or a combination thereof. In some cases, some or all of these components may be considered to be part of or otherwise integrated with the system controller 130, may be separate components in communication with each other, or may be integrated together. In other examples, the system controller 130 can be, or may be in communication with, a cloud computing platform.

[0152] In various embodiments, the opto-fluidic instrument 120 may analyze the sample 110 and may generate the output 190 that includes indications of the presence of the target molecules in the sample 110. For instance, with respect to the example embodiment discussed above where the opto-fluidic instrument 120 employs a hybridization technique for detecting molecules, the opto-fluidic instrument 120 may cause the sample 110 to undergo successive rounds of fluorescent probe hybridization (using two or more sets of fluorescent probes, where each set of fluorescent probes is excited by a different color channel) and be imaged to detect target molecules in the probed sample 110. In such cases, the output 190 may include optical signatures (*e.g.*, a codeword) specific to each gene, which allow the identification of the target molecules.

[0153] In some instances, an assembly for transilluminating a substrate can include a sample carrier device (e.g.., a microfluidic chip or glass slide), a thermal control module configured to control the temperature of the sample carrier device (*e.g.*, a thermoelectric module), and a light source configured to illuminate the sample carrier device. In some instances, the assembly includes a heat exchanger (*e.g.*, a fluid block having a cooling fluid flowing therethrough). In some instances, an assembly for transilluminating can include sample carrier device (e.g., a sample substrate), an optically transparent substrate, a light source configured to illuminate the optically transparent substrate, a light scattering layer configured to scatter light from the light source, and/or a thermal control module configured to control the temperature of the sample carrier device and/or optically transparent substrate.

[0154] In some embodiments, the sample carrier device (*e.g.*, a cassette) can be configured to receive a sample. In some embodiments, the sample carrier device can include one or more microfluidic channels, *e.g.*, sample chambers or microfluidic channels etched into a planar substrate or chambers within a flow cell or microfluidic device.

[0155] A sample carrier device for the systems disclosed herein can include, but is not limited to, a substrate configured to receive a sample, a microscope slide and/or an adapter configured to mount microscope slides (with or without coverslips) on a microscope stage or automated stage (*e.g.*, an automated translation or rotational stage), a substrate, and/or an adapter configured to mount slides on a microscope stage or automated stage, a substrate

comprising etched sample containment chambers (*e.g.*, chambers open to the environment) and/or an adapter configured to mount such substrates on a microscope stage or automated stage, a flow cell and/or an adapter configured to mount flow cells on a microscope stage or automated stage, or a microfluidic device and/or an adapter configured to mount microfluidic devices on a microscope stage or automated stage. In some embodiments, the sample carrier device further includes a cassette configured to secure a substrate (*e.g.*, a glass slide). In some embodiments, the cassette includes two or more components (*e.g.*, a top half and a bottom half) into which the substrate is secured.

[0156] In some instances, the one or more sample carrier devices can be designed for performing a variety of chemical analysis, biochemical analysis, nucleic acid analysis, cell analysis, or tissue analysis applications. In some instances, for example, the sample carrier device (e.g., flow cells and microfluidic devices) may comprise a sample, e.g., a tissue sample. In some instances, the sample carrier device (e.g., flow cells and microfluidic devices) may comprise a sample, e.g., a tissue sample, placed in contact with, e.g., a substrate (e.g., a surface of the flow cell or microfluidic device).

[0157] The sample carrier devices for the disclosed systems (*e.g.*, microscope slides, substrates comprising one or more etched microfluidic channel, flow cells or microfluidic devices comprising one or more microfluidic channels, *etc.*) can be fabricated from any of a variety of materials known to those of skill in the art including, but not limited to, glass (*e.g.*, borosilicate glass, soda lime glass, *etc.*), fused silica (quartz), silicon, polymer (*e.g.*, polystyrene (PS), macroporous polystyrene (MPPS), polymethylmethacrylate (PMMA), polycarbonate (PC), polypropylene (PP), polyethylene (PE), high density polyethylene (HDPE), cyclic olefin polymers (COP), cyclic olefin copolymers (COC), polyethylene terephthalate (PET), polydimethylsiloxane (PDMS), *etc.*), polyetherimide (PEI) and perfluoroelastomer (FFKM) as more chemically inert alternatives, or any combination thereof. FFKM is also known as Kalrez.

[0158] The one or more materials used to fabricate sample carrier devices for the disclosed systems (*e.g.*, substrates configured to receive a sample, microscope slides, substrates comprising one or more etched microfluidic channels, flow cells or microfluidic devices comprising one or more microfluidic channels or sample chambers, *etc.*) can be optically transparent to facilitate use with spectroscopic or imaging-based detection techniques. In some instances, the entire sample carrier device can be optically transparent. Alternatively, in some instances, only a portion of the sample carrier device (*e.g.*, an optically transparent "window") can be optically transparent.

[0159] The sample carrier devices for the disclosed systems (*e.g.*, substrates configured to receive a sample, microscope slides, substrates comprising one or more etched microfluidic channels, flow cells or microfluidic devices comprising one or more microfluidic channels or sample chambers, *etc.*) can be fabricated using any of a variety of techniques known to those of skill in the art, where the choice of fabrication technique is often dependent on the choice of material used, and vice versa. Examples of suitable sample carrier device fabrication techniques include, but are not limited to, extrusion, drawing, precision computer numerical control (CNC) machining and boring, laser photoablation, photolithography in combination with wet chemical etching, deep reactive ion etching (DRIE), micro-molding, embossing, 3D-printing, thermal bonding, adhesive bonding, anodic bonding, and the like (see, e.g., Gale, et al. (2018), "A Review of Current Methods in Microfluidic Device Fabrication and Future Commercialization Prospects", Inventions 3, 60, 1 – 25, which is hereby incorporated by reference in its entirety).

[0160] FIG. 5A illustrates a cross-sectional view of an optics module 200 in an imaging system. One or more illumination sources 210, *e.g.*, one or more light emitting diodes (LEDs), provides light through one or more optical components and an objective lens 220 to thereby illuminate a sample 250. In various embodiments, the optical components include a collimator 211. In various embodiments, the optical components include a field stop 212. In various embodiments, the optical components include one or more excitation filters 213. In various embodiments, the one or more excitation filters 213 are configured to filter light from the illumination source(s) 210 for a predetermined range of wavelengths (*e.g.*, each filter has one or more blocking band(s) and/or transmission band(s) that may be different or may overlap at least in part) and each excitation filter 213 is aligned with appropriate illumination sources (*e.g.*, blue LEDs, green LEDs, yellow LEDs, red LEDs, ultraviolet LEDs, *etc.*). In various embodiments, the optical components include a condenser 214. In various embodiments, the optical components include a beam splitter 215. An optical axis 251 is illustrated extending through the center of the optical surfaces in the objective lens 220 and its path includes an image plane, a focal plane, and input/output pupils (illustrated in FIG. 5B).

[0161] A sensor array 260 (*e.g.*, CMOS sensor) receives light signals from the sample 250. In various embodiments, the optical components include one or more emission filters 265. In various embodiments, the one or more emission filters 265 are configured to filter light from the sample (*e.g.*, emitted from one or more fluorophores, autofluorescence, *etc.*) for a predetermined range of wavelengths (*e.g.*, each filter has one or more blocking band(s) and/or

transmission band(s) that may be different or may overlap at least in part). In various embodiments, the emission filters 265 align (*e.g.*, via motorized translation) with optics and/or the sensor array. In various embodiments, the sample 250 is probed with fluorescent probes configured to bind to a target (*e.g.*, DNA or RNA) that, when illuminated with a particular wavelength (or range of wavelengths) of light, emit light signals that can be detected by the sensor array 260. In various embodiments, the sample 250 is repeatedly probed with two or more (*e.g.*, two, three, four, five, six, *etc.*) different sets of probes. In various embodiments, each set of probes corresponds to a specific color (*e.g.*, blue, green, yellow, or red) such that, when illuminated by that color, probes bound to a target emit light signals. In some embodiments, the sensor array 260 is aligned with the optical axis 251 of the objective lens 220 (*i.e.*, the optical axis of the camera is coincident with and parallel to the optical axis of the objective lens 220). In various embodiments, the sensor array 260 is positioned perpendicularly to the objective lens 220 (*i.e.*, the optical axis of the camera is perpendicular to and intersects the optical axis of the objective lens 220). In various embodiments, a tube lens 261 is mounted in the optical path to focus light on the sensor array 260 thereby allowing for image formation with infinity-corrected objectives. Descriptions of optical modules and illumination assemblies for use in opto-fluidic instruments can be found in U.S. provisional patent application no. 63/427,282, filed on November 22, 2022, titled "Systems and Methods for Illuminating a Sample" and U.S. provisional patent application no. 63/427,360, file on November 22, 2022, titled "Systems and Methods for Imaging Samples," each of which is incorporated by reference in its entirety.

[0162] In various embodiments, the sample is illuminated with one or more wavelengths configured to induce fluorescence in the sample. In various embodiments, the sample is probed during one or more probing cycles with one or more fluorescent probes configured to bind to one or more target analytes. In various embodiments, the one or more wavelengths are selected to induce fluorescence in a subset of the one or more fluorescent probes. In various embodiments, each probing cycle includes illumination with two or more (*e.g.*, four) colors of light. In various embodiments, the sample is treated with a fluorescent stain configured to illuminate one or more structures within the sample. In various embodiments, the sample is contacted with a nuclear stain. In various embodiments, the sample is contacted with 4′,6-diamidino-2-phenylindole ("DAPI") configured to bind to adenine–thymine-rich regions in DNA. In various embodiments, illumination of the sample causes autofluorescence of the sample. In various embodiments, autofluorescence is the natural emission of light by biological structures when they have absorbed light, and may be used to

distinguish the light originating from artificially added fluorescent markers. In various embodiments, fluorescence of the sample through fluorescent probes, autofluorescence, and/or a fluorescent stain can be used with the methods described herein to determine one or more focus metrics of a tissue sample.

[0163] In various embodiments, the sample is illuminated via edge lighting or transillumination along one or more edges of the sample and/or sample substrate. In various embodiments, the edge lighting provides dark-field illumination of the sample. In various embodiments, edge lighting is provided by one or more light sources positioned to provide light substantially perpendicular to a normal of the substrate surface on which the sample is disposed. In various embodiments, the substrate is a glass slide. In various embodiments, the substrate is configured as a wave guide to thereby guide light emitted from the edge lighting towards the sample. In various embodiments, illumination of the sample via edge lighting can be used with the methods described herein to determine one or more focus metrics of a tissue sample.

[0164] Referring now to **FIG. 6**, a schematic of an example of a computing node is shown. Computing node 10 is only one example of a suitable computing node and is not intended to suggest any limitation as to the scope of use or functionality of embodiments of the invention described herein. Regardless, computing node 10 is capable of being implemented and/or performing any of the functionality set forth hereinabove.

[0165] In computing node 10 there is a computer system/server 12, which is operational with numerous other general purpose or special purpose computing system environments or configurations. Examples of well-known computing systems, environments, and/or configurations that may be suitable for use with computer system/server 12 include, but are not limited to, personal computer systems, server computer systems, thin clients, thick clients, handheld or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputer systems, mainframe computer systems, and distributed cloud computing environments that include any of the above systems or devices, and the like.

[0166] Computer system/server 12 may be described in the general context of computer system-executable instructions, such as program modules, being executed by a computer system. Generally, program modules may include routines, programs, objects, components, logic, data structures, and so on that perform particular tasks or implement particular abstract data types. Computer system/server 12 may be practiced in distributed cloud computing environments where tasks are performed by remote processing devices that are linked

through a communications network. In a distributed cloud computing environment, program modules may be located in both local and remote computer system storage media including memory storage devices.

[0167] As shown in **FIG. 6**, computer system/server 12 in computing node 10 is shown in the form of a general-purpose computing device. The components of computer system/server 12 may include, but are not limited to, one or more processors or processing units 16, a system memory 28, and a bus 18 that couples various system components including system memory 28 to processor 16.

[0168] Bus 18 represents one or more of any of several types of bus structures, including a memory bus or memory controller, a peripheral bus, an accelerated graphics port, and a processor or local bus using any of a variety of bus architectures. By way of example, and not limitation, such architectures include Industry Standard Architecture (ISA) bus, Micro Channel Architecture (MCA) bus, Enhanced ISA (EISA) bus, Video Electronics Standards Association (VESA) local bus, and Peripheral Component Interconnect (PCI) bus.

[0169] Computer system/server 12 typically includes a variety of computer system readable media. Such media may be any available media that is accessible by computer system/server 12, and it includes both volatile and non-volatile media, removable and non-removable media.

[0170] System memory 28 can include computer system readable media in the form of volatile memory, such as random access memory (RAM) 30 and/or cache memory 32. Computer system/server 12 may further include other removable/non-removable, volatile/non-volatile computer system storage media. By way of example only, storage system 34 can be provided for reading from and writing to a non-removable, non-volatile magnetic media (not shown and typically called a “hard drive”). Although not shown, a magnetic disk drive for reading from and writing to a removable, non-volatile magnetic disk (e.g., a “floppy disk”), and an optical disk drive for reading from or writing to a removable, non-volatile optical disk such as a CD-ROM, DVD-ROM or other optical media can be provided. In such instances, each can be connected to bus 18 by one or more data media interfaces. As will be further depicted and described below, memory 28 may include at least one program product having a set (e.g., at least one) of program modules that are configured to carry out the functions of embodiments of the invention.

[0171] Program/utility 40, having a set (at least one) of program modules 42, may be stored in memory 28 by way of example, and not limitation, as well as an operating system, one or more application programs, other program modules, and program data. Each of the operating

system, one or more application programs, other program modules, and program data or some combination thereof, may include an implementation of a networking environment. Program modules 42 generally carry out the functions and/or methodologies of embodiments of the invention as described herein.

[0172] Computer system/server 12 may also communicate with one or more external devices 14 such as a keyboard, a pointing device, a display 24, etc.; one or more devices that enable a user to interact with computer system/server 12; and/or any devices (e.g., network card, modem, etc.) that enable computer system/server 12 to communicate with one or more other computing devices. Such communication can occur via Input/Output (I/O) interfaces 22. Still yet, computer system/server 12 can communicate with one or more networks such as a local area network (LAN), a general wide area network (WAN), and/or a public network (e.g., the Internet) via network adapter 20. As depicted, network adapter 20 communicates with the other components of computer system/server 12 via bus 18. It should be understood that although not shown, other hardware and/or software components could be used in conjunction with computer system/server 12. Examples, include, but are not limited to: microcode, device drivers, redundant processing units, external disk drive arrays, RAID systems, tape drives, and data archival storage systems, etc.

[0173] The present invention may be a system, a method, and/or a computer program product. The computer program product may include a computer readable storage medium (or media) having computer readable program instructions thereon for causing a processor to carry out aspects of the present invention.

[0174] The computer readable storage medium can be a tangible device that can retain and store instructions for use by an instruction execution device. The computer readable storage medium may be, for example, but is not limited to, an electronic storage device, a magnetic storage device, an optical storage device, an electromagnetic storage device, a semiconductor storage device, or any suitable combination of the foregoing. A non-exhaustive list of more specific examples of the computer readable storage medium includes the following: a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or Flash memory), a static random access memory (SRAM), a portable compact disc read-only memory (CD-ROM), a digital versatile disk (DVD), a memory stick, a floppy disk, a mechanically encoded device such as punch-cards or raised structures in a groove having instructions recorded thereon, and any suitable combination of the foregoing. A computer readable storage medium, as used herein, is not to be construed as being transitory signals per se, such as radio

waves or other freely propagating electromagnetic waves, electromagnetic waves propagating through a waveguide or other transmission media (e.g., light pulses passing through a fiber-optic cable), or electrical signals transmitted through a wire.

[0175] Computer readable program instructions described herein can be downloaded to respective computing/processing devices from a computer readable storage medium or to an external computer or external storage device via a network, for example, the Internet, a local area network, a wide area network and/or a wireless network. The network may comprise copper transmission cables, optical transmission fibers, wireless transmission, routers, firewalls, switches, gateway computers and/or edge servers. A network adapter card or network interface in each computing/processing device receives computer readable program instructions from the network and forwards the computer readable program instructions for storage in a computer readable storage medium within the respective computing/processing device.

[0176] Computer readable program instructions for carrying out operations of the present invention may be assembler instructions, instruction-set-architecture (ISA) instructions, machine instructions, machine dependent instructions, microcode, firmware instructions, state-setting data, or either source code or object code written in any combination of one or more programming languages, including an object oriented programming language such as Smalltalk, C++ or the like, and conventional procedural programming languages, such as the "C" programming language or similar programming languages. The computer readable program instructions may execute entirely on the user's computer, partly on the user's computer, as a stand-alone software package, partly on the user's computer and partly on a remote computer or entirely on the remote computer or server. In the latter scenario, the remote computer may be connected to the user's computer through any type of network, including a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider). In some embodiments, electronic circuitry including, for example, programmable logic circuitry, field-programmable gate arrays (FPGA), or programmable logic arrays (PLA) may execute the computer readable program instructions by utilizing state information of the computer readable program instructions to personalize the electronic circuitry, in order to perform aspects of the present invention.

[0177] Aspects of the present invention are described herein with reference to flowchart illustrations and/or block diagrams of methods, apparatus (systems), and computer program products according to embodiments of the invention. It will be understood that each block of

the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by computer readable program instructions.

[0178] These computer readable program instructions may be provided to a processor of a general purpose computer, special purpose computer, or other programmable data processing apparatus to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks. These computer readable program instructions may also be stored in a computer readable storage medium that can direct a computer, a programmable data processing apparatus, and/or other devices to function in a particular manner, such that the computer readable storage medium having instructions stored therein comprises an article of manufacture including instructions which implement aspects of the function/act specified in the flowchart and/or block diagram block or blocks.

[0179] The computer readable program instructions may also be loaded onto a computer, other programmable data processing apparatus, or other device to cause a series of operational steps to be performed on the computer, other programmable apparatus or other device to produce a computer implemented process, such that the instructions which execute on the computer, other programmable apparatus, or other device implement the functions/acts specified in the flowchart and/or block diagram block or blocks.

[0180] Referring to **FIG. 7**, an exemplary method 700 of tissue bounds detection is illustrated. In some aspects, provided herein are methods comprising: 702, receiving a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; 704, determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and 706, directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

[0181] In some embodiments, said method further comprising receiving a plurality of images from the probing cycle, wherein the plurality of images comprises the plurality of optical signals; and determining, based on the plurality of optical signals, the first set of 3D positional information. In some such embodiments, the plurality of images comprises a

plurality of z-stacks of the sample. The distance between focal planes in the plurality of z-stacks may be about 0.01 μm to about 2 μm. In some embodiments, the distance between focal planes in the plurality of z-stacks may be about 0.25 μm to about 1 μm. In some preferred embodiments, the distance is about 0.75 μm.

[0182] In some such embodiments, determining the first set of 3D positional information comprises blob detection of the plurality of optical signals. Determining the first set of 3D positional information may comprise image registration and/or alignment of the plurality of focal planes in each z-stack of the plurality of z-stacks.

[0183] In certain embodiments of the invention, the disclosed methods further comprise receiving a second set of 3D positional information of at least a portion of the plurality of biological molecules, wherein the second set of 3D positional information is obtained from a probing cycle of the at least one subsequent probing cycle.

[0184] In some embodiments, the biological sample defines a sample volume, and wherein the first imaging volume is greater than the sample volume. In some such embodiments, the second imaging volume is equal to the sample volume, is less than the sample volume, or is greater than the sample volume.

[0185] In various embodiments, the method provided herein further comprises determining, based on the second set of 3D positional information, a third imaging volume that is greater than or less than the second imaging volume. Some embodiments comprise directing the imaging instrument to image the third imaging volume in at least one subsequent probing cycle of the sample. In some such embodiments, determining the second imaging volume comprises truncating the first set of 3D positional information. In some embodiments, about 0.01% to about 1% of the first set of 3D positional information is truncated at each end. In other embodiments, about 0.2% of the first set of 3D positional information is truncated at each end.

[0186] In some embodiments, the first imaging volume has a height of about 25 μm to about 50 μm. In preferred embodiments, the first imaging volume has a height of about 30 μm.

[0187] In various embodiments, the biological sample has a thickness of about 5 μm to about 20 μm. In some preferred embodiments, the biological sample has a thickness of about 10 μm.

[0188] In embodiments of the invention disclosed herein, the biological molecules comprise nucleic acid molecules. Such nucleic acid molecules may comprise DNA, RNA, or

combinations thereof. In some such embodiments, the biological molecules comprise at least one protein.

[0189] In some aspects of the invention as provided herein, is a computer program product comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a processor to cause the processor to perform a method comprising receiving a first set of 3D positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

[0190] In some embodiments, the computer program product further comprises receiving a plurality of images from the probing cycle, wherein the plurality of images comprises the plurality of optical signals; determining, based on the plurality of optical signals, the first set of 3D positional information. In some such embodiments, the plurality of images comprises a plurality of z-stacks of the sample. A distance between focal planes in the plurality of z-stacks may be about 0.25 μm to about 1 μm. In preferred embodiments, the distance is about 0.75 μm.

[0191] In certain embodiments of the computer program product disclosed herein, determining the first set of 3D positional information comprises blob detection of the plurality of optical signals. In some such embodiments, determining the first set of 3D positional information comprises image registration and/or alignment of the plurality of focal planes in each z-stack of the plurality of z-stacks. The computer program product of the invention may further comprise receiving a second set of 3D positional information of at least a portion of the plurality of biological molecules, wherein the second set of 3D positional information is obtained from a probing cycle of the at least one subsequent probing cycle. In some embodiments, the biological sample defines a sample volume, and wherein the first imaging volume is greater than the sample volume. In some such embodiments, the second imaging volume is equal to the sample volume, is less than the sample volume, or is greater than the sample volume.

In various embodiments, the computer program product may further comprise determining, based on the second set of 3D positional information, a third imaging volume that is greater

than or less than the second imaging volume. In certain embodiments, directing the imaging instrument to image the third imaging volume in at least one subsequent probing cycle of the sample.

[0192] In some embodiments of the computer program product disclosed herein, determining the second imaging volume comprises truncating the first set of 3D positional information. In some such embodiments, about 0.01% to about 1% of the first set of 3D positional information is truncated at each end. In preferred embodiments, about 0.2% of the first set of 3D positional information is truncated at each end.

[0193] In some embodiments, the first imaging volume has a height of about 25 μm to about 50 μm. Preferably, the first imaging volume has a height of about 30 μm.

[0194] In some embodiments, the biological sample has a thickness of about 5 μm to about 20 μm. In preferred embodiments, the biological sample has a thickness of about 10 μm.

[0195] In further embodiments of the computer program product disclosed herein, the biological molecules comprise nucleic acid molecules. Said nucleic acid molecules may comprise DNA, RNA, or combinations thereof. In some embodiments, the biological molecules comprise at least one protein.

[0196] Further aspects of the invention include a system comprising a database; and a computing node comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a processor to cause the processor to perform a method comprising receiving, from the database, a first set of 3D positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

[0197] In some embodiments, the system further comprises receiving a plurality of images from the probing cycle, wherein the plurality of images comprises the plurality of optical signals; determining, based on the plurality of optical signals, the first set of 3D positional information. The plurality of images may comprise a plurality of z-stacks of the sample. In some embodiments, a distance between focal planes in the plurality of z-stacks is about 0.25 μm to about 1 μm. In preferred embodiments, the distance is about 0.75 μm. In various

embodiments, determining the first set of 3D positional information comprises blob detection of the plurality of optical signals. Said determining of the first set of 3D positional information may comprise image registration and/or alignment of the plurality of focal planes in each z-stack of the plurality of z-stacks.

[0198] Embodiments of the system disclosed herein further comprise receiving a second set of 3D positional information of at least a portion of the plurality of biological molecules, wherein the second set of 3D positional information is obtained from a probing cycle of the at least one subsequent probing cycle.

[0199] In some embodiments, the biological sample defines a sample volume, and the first imaging volume is greater than the sample volume. In some such embodiments, the second imaging volume is equal to the sample volume, is less than the sample volume, or is greater than the sample volume.

[0200] In further embodiments, the system further comprises determining, based on the second set of 3D positional information, a third imaging volume that is greater than or less than the second imaging volume. In some such embodiments, directing the imaging instrument to image the third imaging volume in at least one subsequent probing cycle of the sample.

[0201] In certain embodiments, determining the second imaging volume comprises truncating the first set of 3D positional information. In some such embodiments, about 0.01% to about 1% of the first set of 3D positional information is truncated at each end. In preferred embodiments, about 0.2% of the first set of 3D positional information is truncated at each end. Said first imaging volume may have a height of about 25 μm to about 50 μm. In preferred embodiments, first imaging volume has a height of about 30 μm.

[0202] In further embodiments, the biological sample has a thickness of about 5 μm to about 20 μm. In preferred embodiments, the biological sample has a thickness of about 10 μm.

[0203] In some embodiments of the system provided herein, the biological molecules comprise nucleic acid molecules. Said nucleic acid molecules may comprise DNA, RNA, or a combination thereof. In certain embodiments, the biological molecules comprise at least one protein.

[0204] The system provided herein may further comprise an optical assembly configured to image the sample in a plurality of focal planes. Said optical assembly may comprise an objective and a z-axis motion control apparatus configured to move a focal plane of the objective along a z-axis. In some embodiments, the objective has a numerical aperture of

about 0.8 to about 1.50. Preferably, the objective has a numerical aperture of about 1.0. Said objective may comprise a water immersion, a water dipping, or an oil immersion objective.

[0205] Referring to **FIG. 8**, an exemplary method 800 of tissue bounds detection is illustrated. In certain aspects of the invention, provided herein are methods comprising: 802, imaging a first imaging volume to determine a first set of 3D positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules; 804, determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and 806, imaging the second imaging volume in at least one subsequent probing cycle of the sample.

[0206] In some embodiments, imaging the second imaging volume comprises imaging a plurality of sub-volumes. In some such embodiments, each sub-volume of the plurality of sub-volumes comprises a different range in at least one dimension than the other sub-volumes of the plurality of volumes.

[0207] The method provided herein may further comprise data corresponding to the first imaging volume being written to a memory while simultaneously being read from the memory during the probing cycle. In some embodiments, at least a portion of the determination of the first set of 3D positional information occurs simultaneously with the data corresponding to the first imaging volume being written to the memory. In some such embodiments, the memory is a cache memory.

[0208] The flowchart and block diagrams in the Figures illustrate the architecture, functionality, and operation of possible implementations of systems, methods, and computer program products according to various embodiments of the present invention. In this regard, each block in the flowchart or block diagrams may represent a module, segment, or portion of instructions, which comprises one or more executable instructions for implementing the specified logical function(s). In some alternative implementations, the functions noted in the block may occur out of the order noted in the figures. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved. It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of blocks in the block diagrams and/or flowchart illustration, can be implemented by special purpose hardware-based systems that perform the specified functions or acts or carry out combinations of special purpose hardware and computer instructions.

[0209] The descriptions of the various embodiments of the present invention have been presented for purposes of illustration, but are not intended to be exhaustive or limited to the embodiments disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art without departing from the scope and spirit of the described embodiments. The terminology used herein was chosen to best explain the principles of the embodiments, the practical application or technical improvement over technologies found in the marketplace, or to enable others of ordinary skill in the art to understand the embodiments disclosed herein.

Examples

[0210] Example 1: Blob-Finding

[0211] Volumetric imaging enables discrimination between individual targets, such as point sources (e.g., rolling-circle amplification (RCA) products (RCPs) containing concatenated repeats of fluorescently labeled probes), and density clouds (e.g., proteins that are observed as an amorphous concentration of probe signal). In various embodiments, volumetric imaging is performed with an epifluorescent, step-scanning widefield optical system having high magnification (*e.g.*, 20x, 40x, *etc.*). For example, the numerical aperture of the objective is at least 0.8 and has a field of view measuring about 1 to about 1.5mm diagonally. This allows the design of sample preparation and codebooks to capture the most targets (e.g., point sources, observed as blobs) per single experiment. (**FIG. 9.**) **FIG. 9** illustrates exemplary 2-dimensional XZ slices of an imaged volume during a single cycle across a plurality of excitation illumination color channels (*e.g.*, C1, C2, C3, C4) where each slice shows a plurality of blobs (representing detected light signals or punctae of targets) detected in each color channel and distributed throughout the imaged volume. In various embodiments, when the sample is illuminated in each excitation illumination color channel (*e.g.*, C1, C2, C3, C4), a plurality of light signals is detected from within the imageable volume of the sample (if targets within the sample have fluorescent tags with an absorption spectrum corresponding to the specific channel of excitation illumination). In various embodiments, the images (*e.g.*, z-slices) obtained for each z-stack contain a plurality of light signals (including both in-focus light and out-of-focus light). In various embodiments, the z-stacks of images are provided to a blob detection module configured to call blobs based on the plurality of light signals (a full z-stack of images may be required for the blob detection module to determine blobs in the volume defined by the z-stack). In various embodiments, the blobs detected from each image

are represented as a matrix of blob features (center, radius, brightness, color channel, 3D location information, *etc.*). As shown in **FIG. 9**, an imaging instrument (such as an optofluidic instrument described above) obtains volumetric images (*e.g.*, a plurality of z-stacks representing an imageable volume) of a sample during a plurality of cycles. In each cycle, the sample is illuminated sequentially with a plurality of excitation illumination color channels (*e.g.*, C1, C2, C3, and C4) to thereby generate a composite volumetric image (COMP). In various embodiments, the plurality of cycles includes about 15 cycles. In various embodiments, the plurality of cycles includes up to 30 cycles (for example, for analyzing 500 to 1000 target molecules). In various embodiments, the plurality of cycles includes up to 100 cycles (for example, for analyzing 5000 target molecules). In various embodiments, the plurality of cycles includes over 100 cycles (for a full transcriptome analysis of over 20,000 genes). For example, the plurality of cycles may include about 100 to about 500 cycles. In various embodiments, the sample is a biological tissue sample. In various embodiments, the sample is a hydrogel that, for example, has a plurality of target molecules (*e.g.*, RNA transcripts or a sequence library) distributed within and affixed to the hydrogel. In various embodiments, each cycle in the plurality of cycles includes probing with a plurality of probes. In various embodiments, the plurality of probes includes a plurality of fluorescently tagged oligonucleotides. In various embodiments, the plurality of probes includes fluorescently tagged nucleotides (*e.g.*, dNTPs). In various embodiments, the plurality of probes includes fluorescent tags that have absorption spectra corresponding to one or more excitation color channels. For example, the plurality of probes may include a first set of probes having an absorption spectrum corresponding to a red excitation color channel (*e.g.*, C1), a second set of probes having an absorption spectrum corresponding to a yellow excitation color channel (*e.g.*, C2), a third set of probes having an absorption spectrum corresponding to a green excitation color channel (*e.g.*, C3), a fourth set of probes having an absorption spectrum corresponding to a blue excitation color channel (*e.g.*, C4). In various embodiments, during each cycle, the sample is imaged in the plurality of excitation color channels. For example, the sample may be imaged in at least one of near UV, blue, green, yellow, and red excitation color channels. In various embodiments, during at least one of the cycles, the sample is contacted with one or more stains (*e.g.*, DAPI) that are imaged in a unique color channel. For example, DAPI is imaged in the near UV color channel.

[0212] In various embodiments, during high-throughput image acquisition, a sample thickness (Z-dimension) is overestimated at the start of imaging (*e.g.*, at the first cycle). For example, for an approximately 10 μm thick tissue section, a 30μm-thick volume is imaged to

support variations in tissue section thickness and to provide margins that mitigate risk of tissue drifting outside of the imaged volume due to thermomechanical deformation, non-orthogonality, and various accumulated position control errors. (**FIG. 10.**) In various embodiments, image acquisition and/or image processing are drivers of total instrument run duration and the extra imaged Z-dimension increases the number of total images to be acquired, thereby increasing instrument run duration (and decreasing throughput). For example, at maximum regions of interest in two cassettes (which may be user-selected or automatically determined), a single cycle involves 1 hour of chemistry (*e.g.* fluidic operations) followed by 3 hours of imaging. Measurement of the signals of interest at the approximate volumetric bounds that it resides within a sample would allow for imaging a reduced, optimal volume, even after applying drift margins.

[0213] One approach is to fit a curve using a focus metric (*e.g.*, Discrete Cosine Transform, Tenenbaum Gradient (Tenengrad), Normalized Variance, Vollath's F4) and assign bounds using a heuristic (*e.g.*, 2nd derivative argmax). For instance, a dark field imaging channel and/or nuclear stain channel may be used, since both provide global tissue signals that persist throughout the imaging and chemistry cycles. (**FIG. 11.**)

[0214] In various embodiments, the resultant curve elbow finding heuristic may be arbitrary, i.e., not rooted in the physical properties of the tissue sample itself, and focus curve shapes are variable and complex across tissue types and individual samples. (**FIG. 12.**). Various methods of using inherent contrast of a sample to determine tissue bounds is described in U.S. patent application no. 18/394,370, which is incorporated by reference herein in its entirety.

[0215] In various embodiments, a tissue sample is sectioned to a nominal thickness (*e.g.*, 10 μm) as determined by a sectioning instrument (*e.g.*, a microtome). In various embodiments, the actual thickness cannot be known precisely without applying a metrology method. In various embodiments, the tissue section undergoes a non-uniform/-isometric deformation in varying conditions (e.g., a section has shape A when paraffin embedded, shape B when deparaffinized, shape C when frozen, shape D when in aqueous solution, shape E when in a solution with high alcohol concentration, etc.).

[0216] For the measured physical properties to be reliable and useful, tissue thickness was determined after all sample handling steps and under the sequencing conditions. Volumetrically localized blobs (detected areas of observed signal, such as punctae, whether discrete, clustered, or overlapping, or diffuse clouds of probe signal, such as fluorescence of labelled proteins) were used to identify the portion of the tissue thickness that contains

decoding-relevant information. In various embodiments, this strategy of using volumetrically localized blobs to determine tissue bounds uses no heuristics or guesswork in estimating the information content of images and instead uses empirically observed blob signals. Also, there is no new image processing as the existing primary sample image analysis is leveraged. Notably, a full volumetric scan must be acquired to compute blob distribution. (**FIG. 13.**)

[0217] In various embodiments, z-slices in a z-stack are obtained at predetermined heights, *e.g.*, relative to a surface (*e.g.*, a sample surface, a substrate surface, *etc.*). In various embodiments, the predetermined heights for the z-slices are the same across all FOVs of the sample. In various embodiments, a tissue thickness is determined as a subset of z-slices within a z-stack (*e.g.*, the smallest set of z-slices that can be imaged to obtain the target signal or blobs across all FOVs). In various embodiments, a tissue thickness is determined based on a predetermined delta between slices. For example, if the delta between slices is 0.75 μm and the detected signal is located between slices, a ceiling function may be applied to the upper thickness z-height to round the upper thickness z-height up to the next z-slice and/or a floor function may be applied to a lower thickness z-height to round the lower thickness z-height down to the next z-slice to ensure that the thickness containing the full target signal (*e.g.*, blobs) is obtained. In various embodiments, to improve robustness of the tissue thickness determination against outlier signal, one or more z-slice is subtracted from the upper z-slice (associated with the determined upper sample bounds) and one or more z-slice is subtracted from the lower z-slice (associated with the determined lower sample bounds). For example, if 40 z-slices (slices 1 through 40) are obtained with 0.75 μm spacing (representing a 30 μm initial imageable volume), and blobs are detected from slice 15 to slice 25 across all FOVs, a reduced imageable volume may be determined after the first cycle to be slices 10 through 30. This represents a 15 μm reduced imageable volume (a 50% reduction in imageable volume that will result in significant throughput gains because half of the z-slices are required to be imaged to obtain the same results). In another example, the imageable volume may be further reduced to slices 15 to 25, corresponding to only the slices where blobs were detected across all FOVs. Using this example, for robustness against outliers, the imageable volume may be further reduced by one slice on each of the upper and lower bounds such that only slices 16 to 24 are imaged across all FOVs.

[0218] *In silico* analysis of historical run-data assessed the percent of all decoded point sources (e.g., amplified signals of RCPs observed as various punctae as described herein, representing the signal of interest for blob-finding analysis) with reference to imaged Z-slices. (**FIG. 14A-14D**.) As shown in **FIG. 14A**, an imaged volume includes a plurality of z-

stacks from a plurality of FOVs of a sample. The plurality of z-stacks is provided to a blob finding module. In various embodiments, at the blob finding module, flat field correction is performed on the images. In various embodiments, at the blob finding module, point spread function (PSF) fitting is performed on the images. In various embodiments, object localization is performed on each image to call blobs from the z-stack volumetric image data. As shown in **FIG. 14B**, the distribution of called blobs is tilted at an angle, possibly due to bending or tilt of the substrate on which the sample is positioned. As shown in **FIG. 14B**, the distribution of blobs is detected between an upper z-slice 1410 (approximately z-slice number 11) and a lower z-slice 1412 (approximately z-slice number 27). In various embodiments, a reduced imageable volume can be determined based on the upper z-slice 1410 and the lower z-slice 1412. In various embodiments, a reduced volume can be determined by applying a rounding function (*e.g.*, a ceiling) to the upper z-slice and/or applying a rounding function (*e.g.*, a floor) to the lower z-slice. For example, a reduced volume is determined to be from a lower z-slice of 10 to an upper z-slice of 30. In another example, the reduced volume is determined to be from a lower z-slice of 11 to an upper z-slice of 27. Referring back to **FIG. 14A**, in various embodiments, the detected blobs are provided to a thickness profiling module including statistical bounding and profile synthesis. **FIG. 14C** illustrates the process of statistical bounding where violin plots of detected blobs across all color channels and all FOVs containing sample are generated. **FIG. 14D** illustrates the process of profile synthesis where a thickness profile is generated across all FOVs of the sample based on the detected blobs. As shown in **FIG. 14D**, a section of mouse brain tissue has an imageable volume that is reduced to about 20 z-slices (from an initial scan of 40 z-slices).

[0219] Experimental data suggested a significant reduction in the thickness (*i.e.*, the number of z-slices) imaged was possible while retaining >99% of decoded RCPs. (**FIG. 15**.) As shown in **FIG. 15**, over 99% of decodable RCPs are covered for most tissue samples after reducing the imageable volume to a size of about 10 z-slices (corresponding to an imaged thickness of about 7.5 μm). In various embodiments, as shown in **FIG. 15**, experiments showed that over 99.5% of decodable RCPs were covered after reducing the imageable volume to a size of about 15 to about 20 z-slices (corresponding to an imaged thickness of about 11.25 μm to about 15 μm). In various embodiments, more stringent drift compensation methods or additional margins can be implemented to be tolerant of non-orthogonality and drift. (**FIG. 16.**) As shown in **FIG. 16**, tilt of a sample substrate (*e.g.*, a glass slide) can cause significant variations in the z-heights that must be imaged to obtain full coverage of decodable RCPs, resulting in an increased number of z-slices needed (assuming that a fixed

upper and lower bound is used across all FOVs). In various embodiments, when substrate tilt is detected to be above a predetermined threshold, the initial thickness for z-stack imaging is increased. For example, the initial z-stacks may include 50 z-slices having a 0.75 μm spacing where, for a substantially flat sample substrate, the initial z-stacks would have been imaged using 40 z-slices having the same spacing. In various embodiments, when substrate tilt is detected to be above a predetermined threshold, the upper and lower bounds may be determined on a FOV-by-FOV basis and, thus, the imaging instrument obtains a different subset of z-slices for each z-stack. For example, in a first FOV, z-slices 11 to 27 are imaged while, in a second FOV, z-slices 20 to 35 are imaged due to tilt of the sample substrate. In various embodiments, the magnitude of additional margins was determined using analysis of the evolution of slide tilt and drift across a subpopulation of instruments.

[0220] Referring now to **FIG. 17**, an exemplary instrument data flow is illustrated. An exemplary thickness profile sequencing suitable for execution in conjunction with this instrument data flow is shown in **FIG. 18**. In various embodiments, the implementation of blob-finding analysis within this sample analysis pipelines required the full volumetric scan data to be complete (e.g. Cycle 1) and then operated on in bulk. In various embodiments, the volumetric sample imaging instrument has a plurality of imaging cycles, such as 15 imaging cycles in this example. In various embodiments, the first cycle includes a DAPI imaging cycle (using near ultraviolet illumination). In various embodiments, any of the additional cycles after the first cycle may include a DAPI imaging cycle. In various embodiments, any cycle may include two illumination channels (*e.g.*, NUV and one of blue, green, yellow, or red). Preferably, the first cycle includes two illumination channels (*e.g.*, DAPI and green). In various embodiments, any cycle includes three illumination channels. In various embodiments, any cycle includes four illumination channels. In various embodiments, any cycle includes five illumination channels (*e.g.*, NUV, blue, green, yellow, red). In various embodiments, additional cycles after the first cycle include four illumination channels (*e.g.*, blue, green, yellow, red). In various embodiments, Cycle 2 is imaged at a 40-slice volume representing approximately a 30um, similar to Cycle 1 (i.e., thickness overestimation) as the computation is not yet complete and is, therefore, wasted time. The thickness profile was produced by the third imaging cycle (Cycle 3) due to the pipeline computation model. Despite the known inefficiency, per cycle imaging duration was reduced by 50 minutes for a 11.25μm thick tissue section across 4 cm^2 total. This resulted in approximately 11 hours saved per run. (**FIG. 19**.)

[0221] In various embodiments, the first cycle is imaged using an imaging order of ZCYX. In various embodiments, the first cycle is imaged using an imaging order of ZXYC. In various embodiments, additional cycles after the first cycle are imaged using an imaging order of ZXYC. In various embodiments, additional cycles after the first cycle are imaged using an imaging order of ZCYX. In various embodiments, an imaging order of ZCYX (or ZCXY) images a z-stack of a single FOV in all color channels before moving in X or Y to the next FOV to image the next FOV in all color channels, and so on and so forth until all FOVs are imaged. In various embodiments, an imaging order of ZXYC (or ZYXC) images all z-stacks of all FOVs in a single illumination channel and then switches to the next illumination channel to reimage all z-stacks of all FOVs, and so on and so forth until all illumination channels are imaged.

Example 2: Blob-Finding System Designs

[0222] Referring now to **FIG. 20**, an exemplary optimized profiling sequence is illustrated, also suitable for execution in conjunction with the instrument data flow of **FIG. 17**. Blob detection is performed on in-coming data (i.e. "streaming", "real time", "on the fly"). By carefully managing the read/write access pattern to the filesystem and the disk banks, the data is guaranteed to stay in the filesystem cache in RAM, thus never hitting the spinning disk write head. This will allow real-time read and write operations, such that blob detection could be performed while acquiring the scan. Time is not wasted while waiting for the scan to complete prior to starting blob detection on acquired images, thus, results will be available faster, with minimal (*e.g.*, no) postponement of subsequent cycles.

CLAIMS

What is claimed is:

1. A method comprising:

receiving a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules;

determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and

directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

2. The method of claim 1, further comprising:

receiving a plurality of images from the probing cycle, wherein the plurality of images comprises the plurality of optical signals;

determining, based on the plurality of optical signals, the first set of 3D positional information.

3. The method of claim 2, wherein the plurality of images comprises a plurality of z-stacks of the sample.
4. The method of claim 3, wherein a distance between focal planes in the plurality of z-stacks is about 0.25 μm to about 1 μm.
5. The method of claim 4, wherein the distance is about 0.75 μm.
6. The method of any one of claims 2 to 5, wherein determining the first set of 3D positional information comprises blob detection of the plurality of optical signals.
7. The method of any one of claims 2 to 6, wherein determining the first set of 3D positional information comprises image registration and/or alignment of the plurality of focal planes in each z-stack of the plurality of z-stacks.
8. The method of any one of claims 1 to 7, further comprising receiving a second set of 3D positional information of at least a portion of the plurality of biological molecules, wherein the second set of 3D positional information is obtained from a probing cycle of the at least one subsequent probing cycle.

9. The method of any one of claims 1 to 8, wherein the biological sample defines a sample volume, and wherein the first imaging volume is greater than the sample volume.
10. The method of claim 9, wherein the second imaging volume is equal to the sample volume.
11. The method of claim 9, wherein the second imaging volume is less than the sample volume.
12. The method of claim 9, wherein the second imaging volume is greater than the sample volume.
13. The method of any one of claims 1 to 12, further comprising:

determining, based on the second set of 3D positional information, a third imaging volume that is greater than or less than the second imaging volume.
14. The method of claim 13, directing the imaging instrument to image the third imaging volume in at least one subsequent probing cycle of the sample.
15. The method of any one of claims 1 to 14, wherein determining the second imaging volume comprises truncating the first set of 3D positional information.
16. The method of claim 15, wherein about 0.01% to about 1% of the first set of 3D positional information is truncated at each end.
17. The method of claim 16, wherein about 0.2% of the first set of 3D positional information is truncated at each end.
18. The method of any one of claims 1 to 17, wherein the first imaging volume has a height of about 25 μm to about 50 μm.
19. The method of claim 18, wherein the first imaging volume has a height of about 30 μm.
20. The method of any one of claims 1 to 19, wherein the biological sample has a thickness of about 5 μm to about 20 μm.
21. The method of claim 20, Wherein the biological sample has a thickness of about 10 μm.
22. The method of any one of claims 1 to 21, wherein the biological molecules comprise nucleic acid molecules.
23. The method of claim 22, wherein the nucleic acid molecules comprise DNA.
24. The method of claim 22 or claim 23, wherein the nucleic acid molecules comprise RNA.

25. The method of any one of claims 22 to 24, wherein the biological molecules comprise at least one protein.
26. A computer program product comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a processor to cause the processor to perform a method comprising:

receiving a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules;

determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and

directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

27. The computer program product of claim 26, further comprising:

receiving a plurality of images from the probing cycle, wherein the plurality of images comprises the plurality of optical signals;

determining, based on the plurality of optical signals, the first set of 3D positional information.

28. The computer program product of claim 27, wherein the plurality of images comprises a plurality of z-stacks of the sample.
29. The computer program product of claim 28, wherein a distance between focal planes in the plurality of z-stacks is about 0.25 μm to about 1 μm.
30. The computer program product of claim 29, wherein the distance is about 0.75 μm.
31. The computer program product of any one of claims 27 to 30, wherein determining the first set of 3D positional information comprises blob detection of the plurality of optical signals.
32. The computer program product of any one of claims 27 to 31, wherein determining the first set of 3D positional information comprises image registration and/or alignment of the plurality of focal planes in each z-stack of the plurality of z-stacks.
33. The computer program product of any one of claims 26 to 32, further comprising receiving a second set of 3D positional information of at least a portion of the

plurality of biological molecules, wherein the second set of 3D positional information is obtained from a probing cycle of the at least one subsequent probing cycle.

34. The computer program product of any one of claims 26 to 33, wherein the biological sample defines a sample volume, and wherein the first imaging volume is greater than the sample volume.
35. The computer program product of claim 34, wherein the second imaging volume is equal to the sample volume.
36. The computer program product of claim 34, wherein the second imaging volume is less than the sample volume.
37. The computer program product of claim 34, wherein the second imaging volume is greater than the sample volume.
38. The computer program product of any one of claims 26 to 37, further comprising:

determining, based on the second set of 3D positional information, a third imaging volume that is greater than or less than the second imaging volume.

39. The computer program product of claim 38, directing the imaging instrument to image the third imaging volume in at least one subsequent probing cycle of the sample.
40. The computer program product of any one of claims 26 to 39, wherein determining the second imaging volume comprises truncating the first set of 3D positional information.
41. The computer program product of claim 40, wherein about 0.01% to about 1% of the first set of 3D positional information is truncated at each end.
42. The computer program product of claim 41, wherein about 0.2% of the first set of 3D positional information is truncated at each end.
43. The computer program product of any one of claims 26 to 42, wherein the first imaging volume has a height of about 25 μm to about 50 μm.
44. The computer program product of claim 43, wherein the first imaging volume has a height of about 30 μm.
45. The computer program product of any one of claims 26 to 44, wherein the biological sample has a thickness of about 5 μm to about 20 μm.
46. The computer program product of claim 45, Wherein the biological sample has a thickness of about 10 μm.

47. The computer program product of any one of claims 26 to 46, wherein the biological molecules comprise nucleic acid molecules.
48. The computer program product of claim 47, wherein the nucleic acid molecules comprise DNA.
49. The computer program product of claim 47 or claim 48, wherein the nucleic acid molecules comprise RNA.
50. The computer program product of any one of claims 47 to 49, wherein the biological molecules comprise at least one protein.
51. A system comprising:

a database; and

a computing node comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a processor to cause the processor to perform a method comprising:

receiving, from the database, a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules;

determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and

directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample.

52. The system of claim 51, further comprising:

receiving a plurality of images from the probing cycle, wherein the plurality of images comprises the plurality of optical signals;

determining, based on the plurality of optical signals, the first set of 3D positional information.

53. The system of claim 52, wherein the plurality of images comprises a plurality of z-stacks of the sample.
54. The system of claim 53, wherein a distance between focal planes in the plurality of z-stacks is about 0.25 μm to about 1 μm.
55. The system of claim 54, wherein the distance is about 0.75 μm.

56. The system of any one of claims 52 to 55, wherein determining the first set of 3D positional information comprises blob detection of the plurality of optical signals.
57. The system of any one of claims 52 to 56, wherein determining the first set of 3D positional information comprises image registration and/or alignment of the plurality of focal planes in each z-stack of the plurality of z-stacks.
58. The system of any one of claims 51 to 57, further comprising receiving a second set of 3D positional information of at least a portion of the plurality of biological molecules, wherein the second set of 3D positional information is obtained from a probing cycle of the at least one subsequent probing cycle.
59. The system of any one of claims 51 to 58, wherein the biological sample defines a sample volume, and wherein the first imaging volume is greater than the sample volume.
60. The system of claim 59, wherein the second imaging volume is equal to the sample volume.
61. The system of claim 59, wherein the second imaging volume is less than the sample volume.
62. The system of claim 59, wherein the second imaging volume is greater than the sample volume.
63. The system of any one of claims 51 to 62, further comprising:

determining, based on the second set of 3D positional information, a third imaging volume that is greater than or less than the second imaging volume.
64. The system of claim 63, directing the imaging instrument to image the third imaging volume in at least one subsequent probing cycle of the sample.
65. The system of any one of claims 51 to 64, wherein determining the second imaging volume comprises truncating the first set of 3D positional information.
66. The system of claim 65, wherein about 0.01% to about 1% of the first set of 3D positional information is truncated at each end.
67. The system of claim 66, wherein about 0.2% of the first set of 3D positional information is truncated at each end.
68. The system of any one of claims 51 to 67, wherein the first imaging volume has a height of about 25 μm to about 50 μm.
69. The system of claim 68, wherein the first imaging volume has a height of about 30 μm.

70. The system of any one of claims 51 to 69, wherein the biological sample has a thickness of about 5 μm to about 20 μm.
71. The system of claim 70, wherein the biological sample has a thickness of about 10 μm.
72. The system of any one of claims 51 to 71, wherein the biological molecules comprise nucleic acid molecules.
73. The system of claim 72, wherein the nucleic acid molecules comprise DNA.
74. The system of claim 72 or claim 73, wherein the nucleic acid molecules comprise RNA.
75. The system of any one of claims 72 to 74, wherein the biological molecules comprise at least one protein.
76. The system of any one of claims 51 to 75, further comprising an optical assembly configured to image the sample in a plurality of focal planes.
77. The system of claim 76, wherein the optical assembly comprises an objective and a z-axis motion control apparatus configured to move a focal plane of the objective along a z-axis.
78. The system of claim 77, wherein the objective has a numerical aperture of about 0.8 to about 1.50.
79. The system of claim 78, wherein the objective has a numerical aperture of about 1.0.
80. The system of any one of claims 77 to 79, wherein the objective comprises a water immersion, a water dipping, or an oil immersion objective.
81. A method comprising:

imaging a first imaging volume to determine a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules;

determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume; and

imaging the second imaging volume in at least one subsequent probing cycle of the sample.

82. The method of claim 81, wherein the imaging the second imaging volume comprises imaging a plurality of sub-volumes.

83. The method of claim 82, wherein each sub-volume of the plurality of sub-volumes comprises a different range in at least one dimension than the other sub-volumes of the plurality of volumes.
84. The method of claim 81, further comprising data corresponding to the first imaging volume being written to a memory while simultaneously being read from the memory during the probing cycle.
85. The method of claim 84, wherein at least a portion of the determination of the first set of 3D positional information occurs simultaneously with the data corresponding to the first imaging volume being written to the memory.
86. The method of claim 84, wherein the memory is a cache memory.

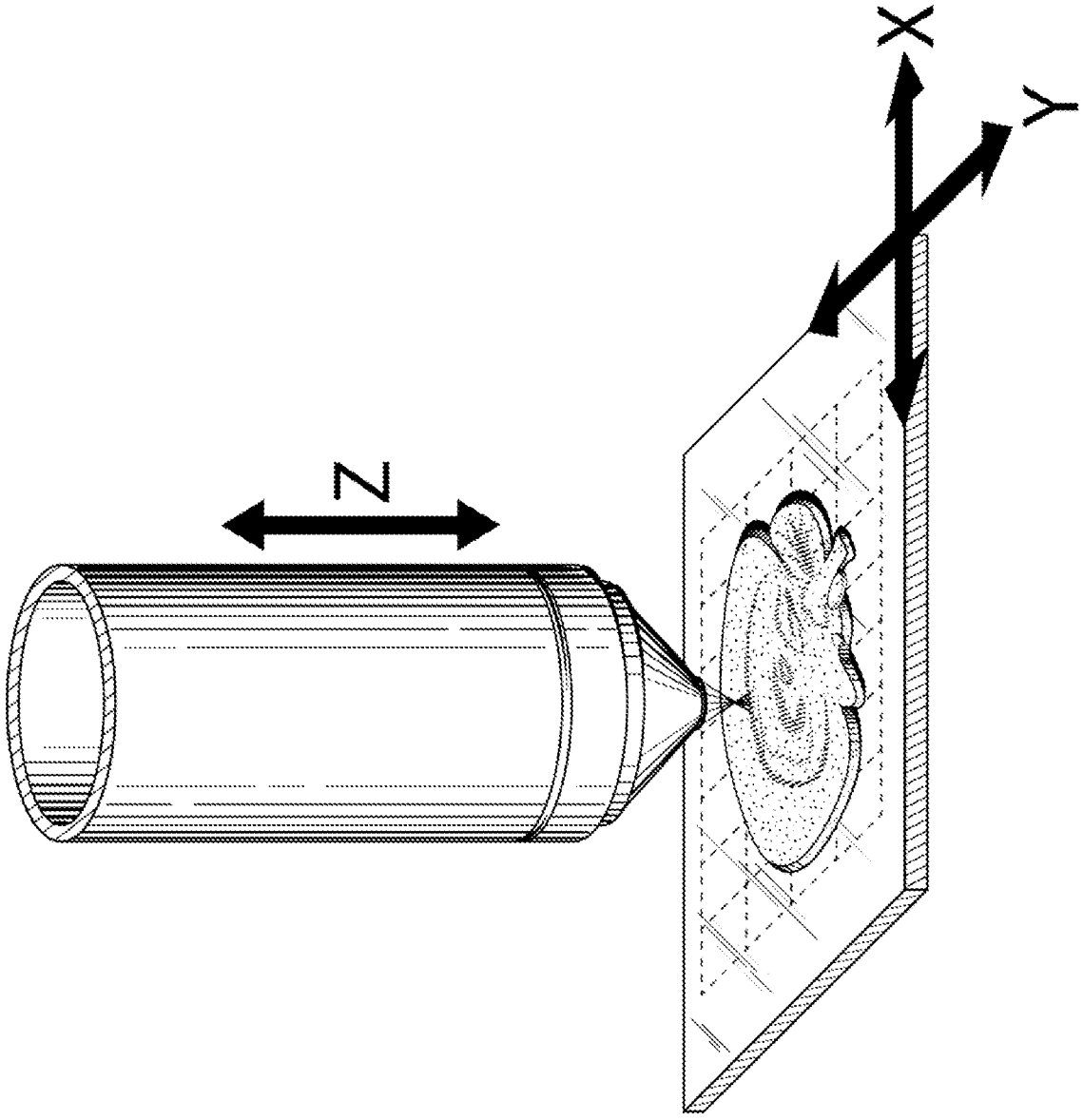


FIG. 1

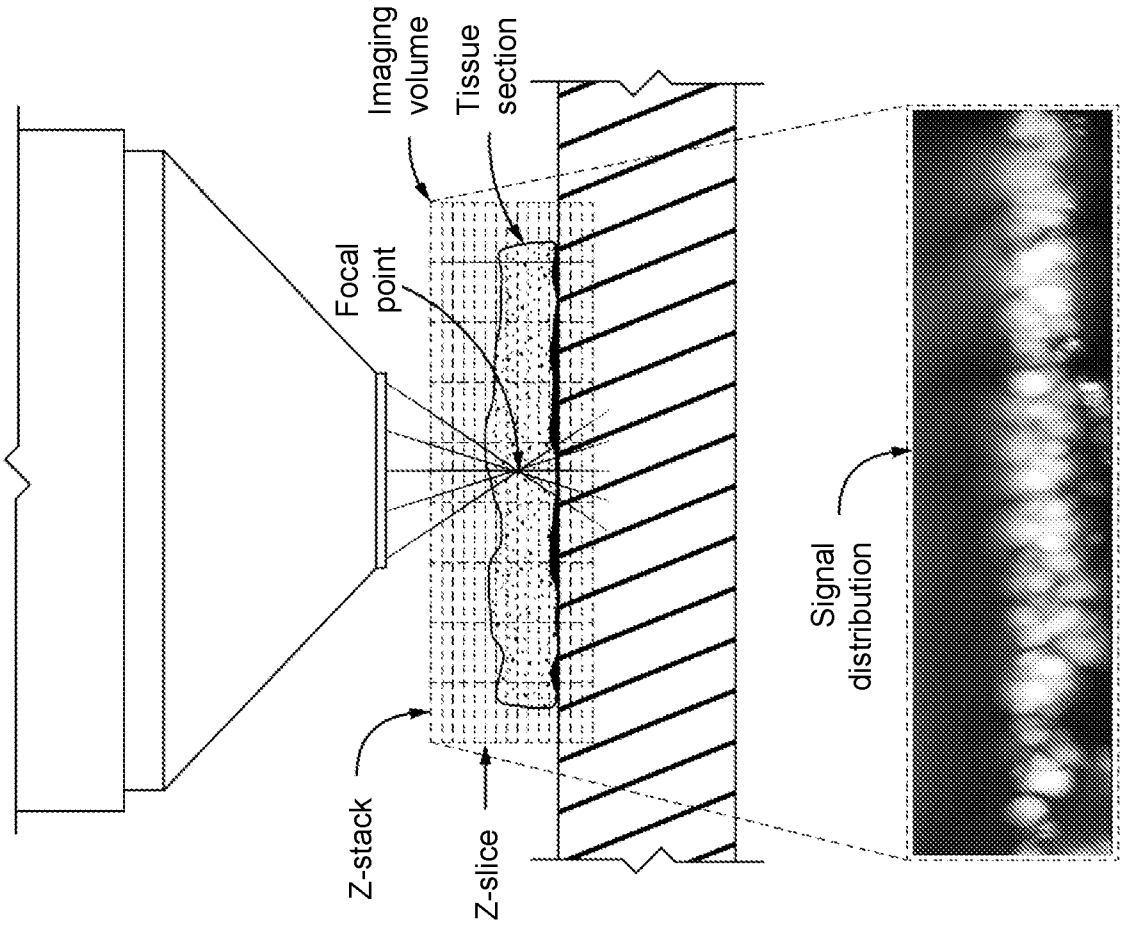

Imaging volume
Tissue section
Focal point
Signal distribution
Z-stack
Z-slice

FIG. 2

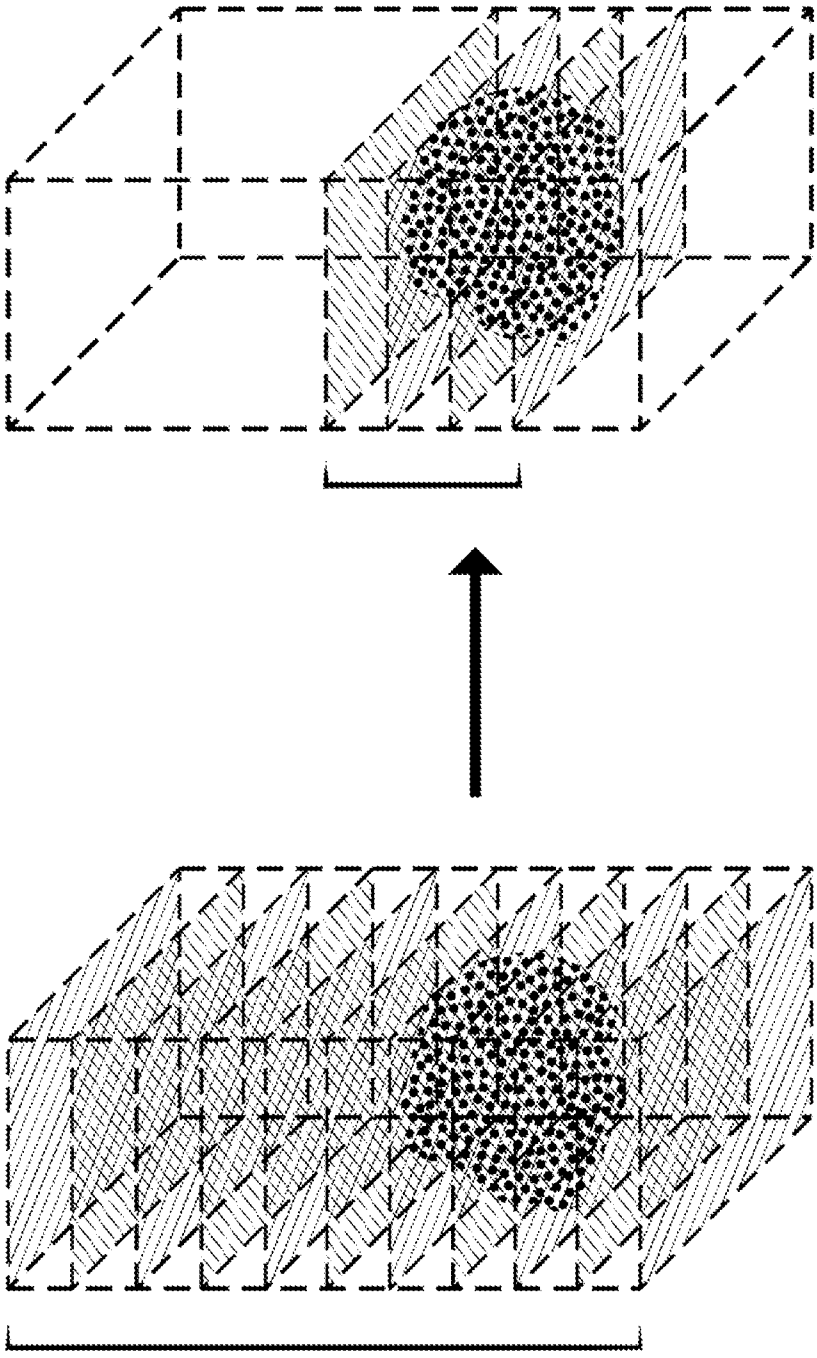


FIG. 3

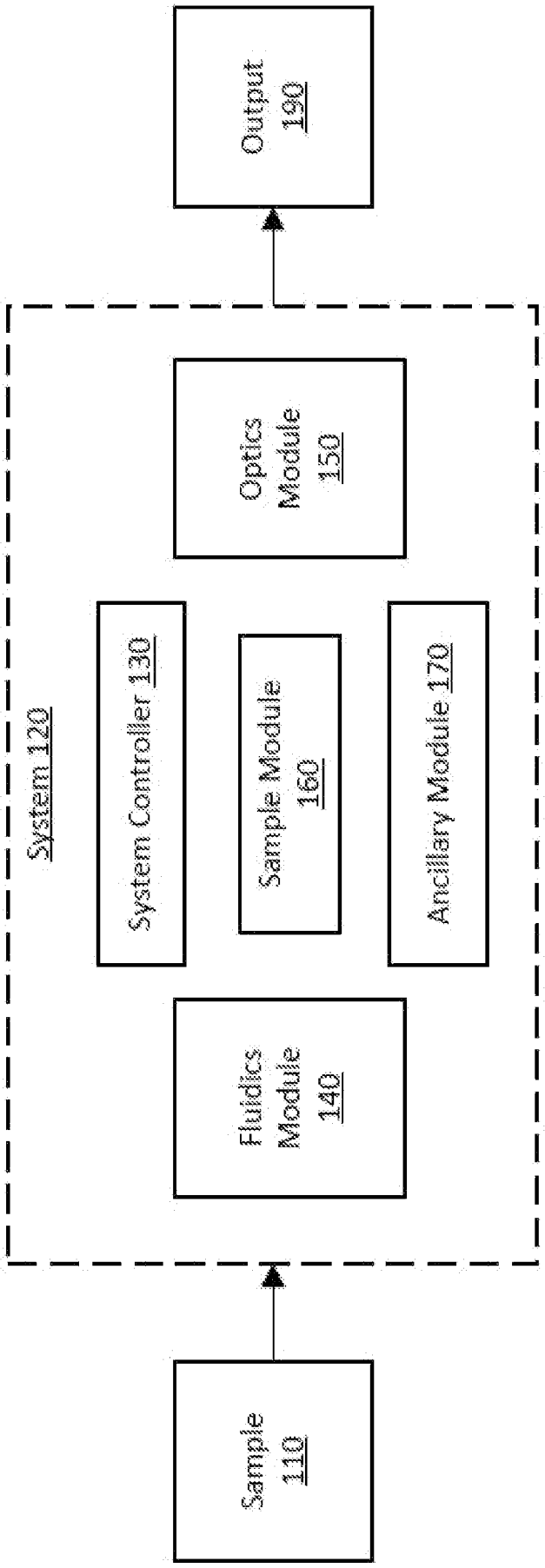

Output
190
System 120
Optics
Module
150
System Controller 130
Sample Module
160
Ancillary Module 170
Fluidics
Module
140
Sample
110

FIG. 4

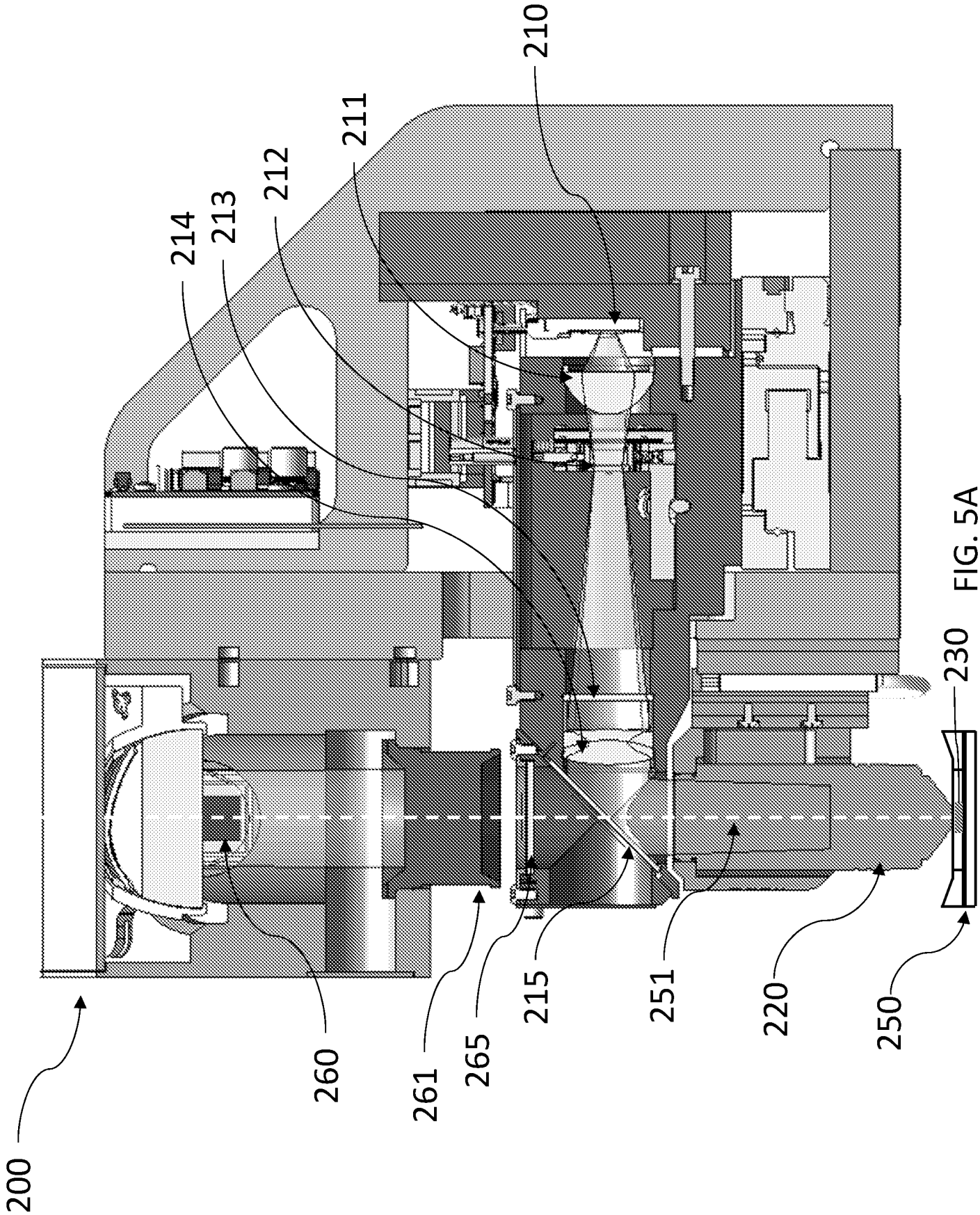


FIG. 5A

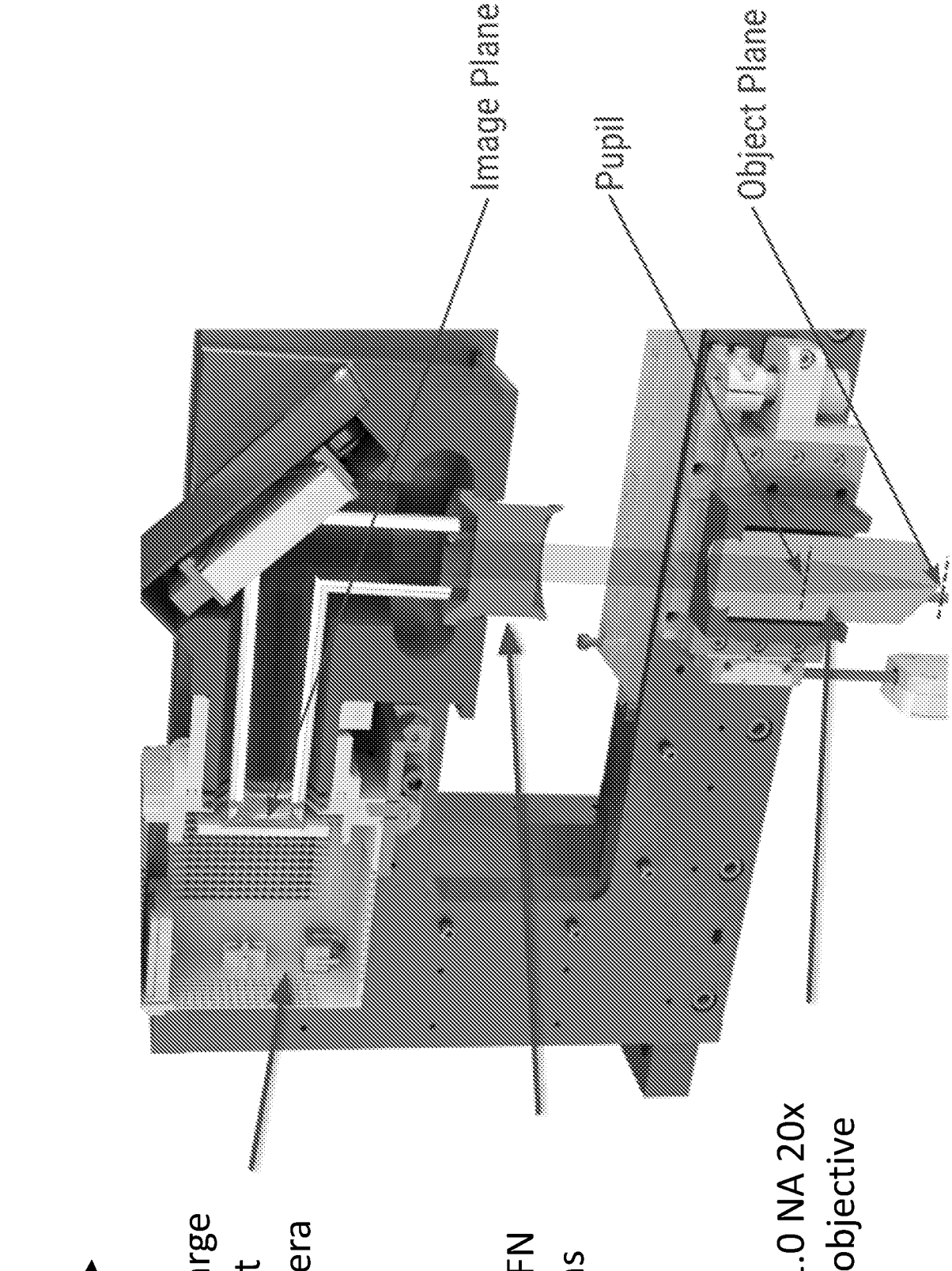

200
Small pixel, large sensor, fast readout camera
26.5mm FN tube lens
1.0 NA 20x objective
Image Plane
Pupil
Object Plane

FIG. 5B

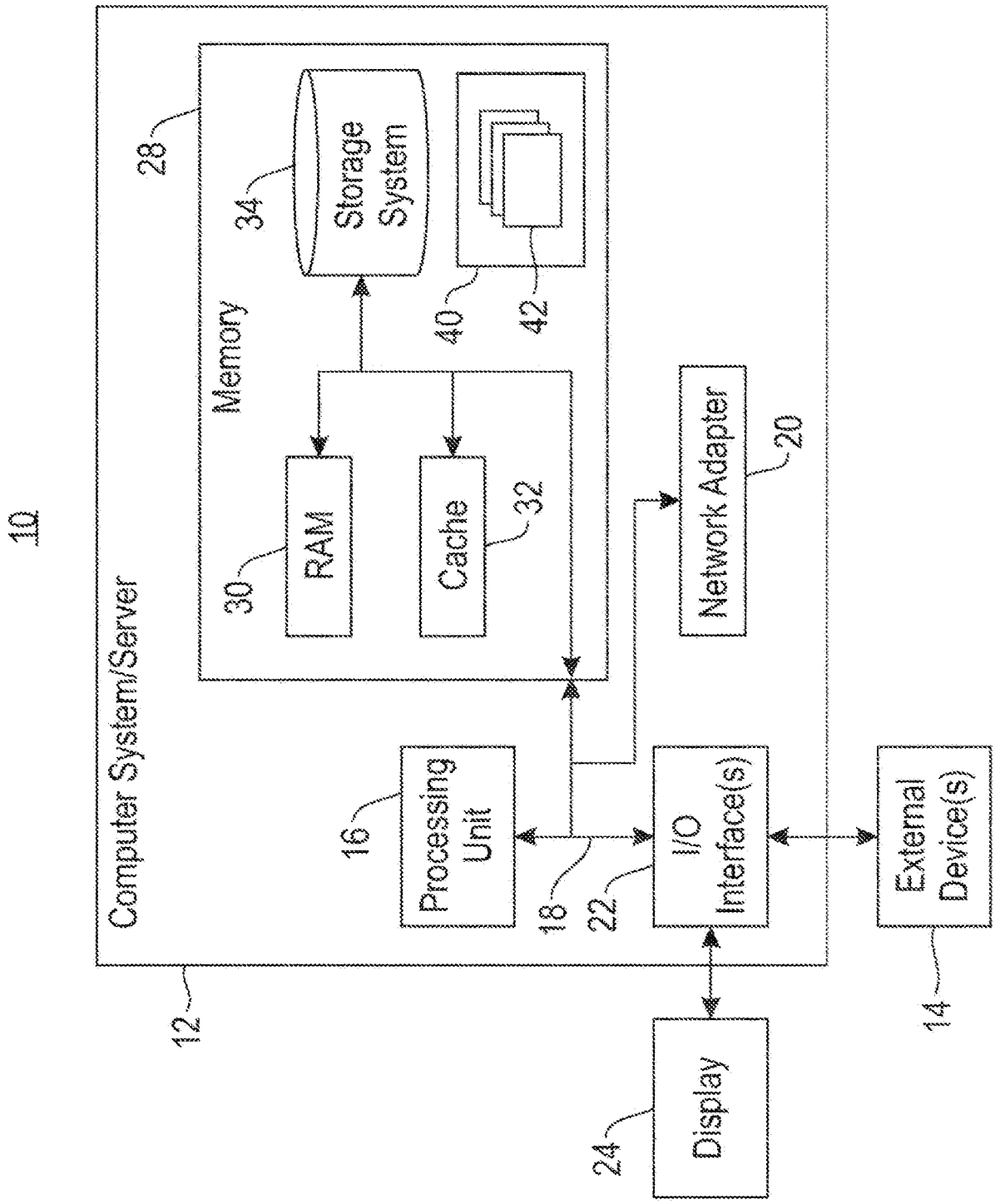

10
Computer System/Server
12
28
Memory
34
Storage System
40
42
30
RAM
Cache
32
Network Adapter
20
16
Processing Unit
18
22
I/O Interface(s)
External Device(s)
14
24
Display

FIG. 6

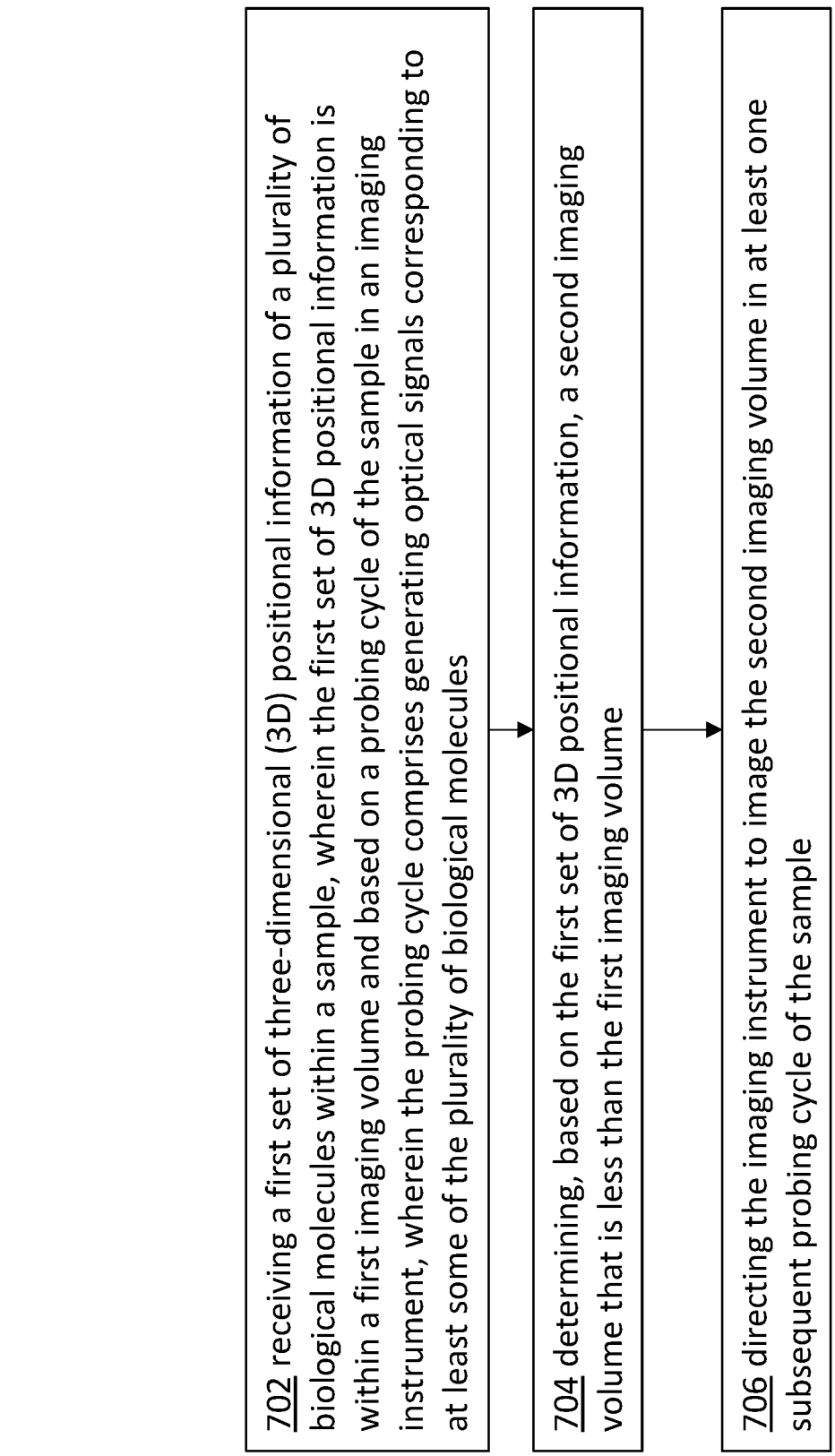

700
702 receiving a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is within a first imaging volume and based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules
704 determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume
706 directing the imaging instrument to image the second imaging volume in at least one subsequent probing cycle of the sample

FIG. 7

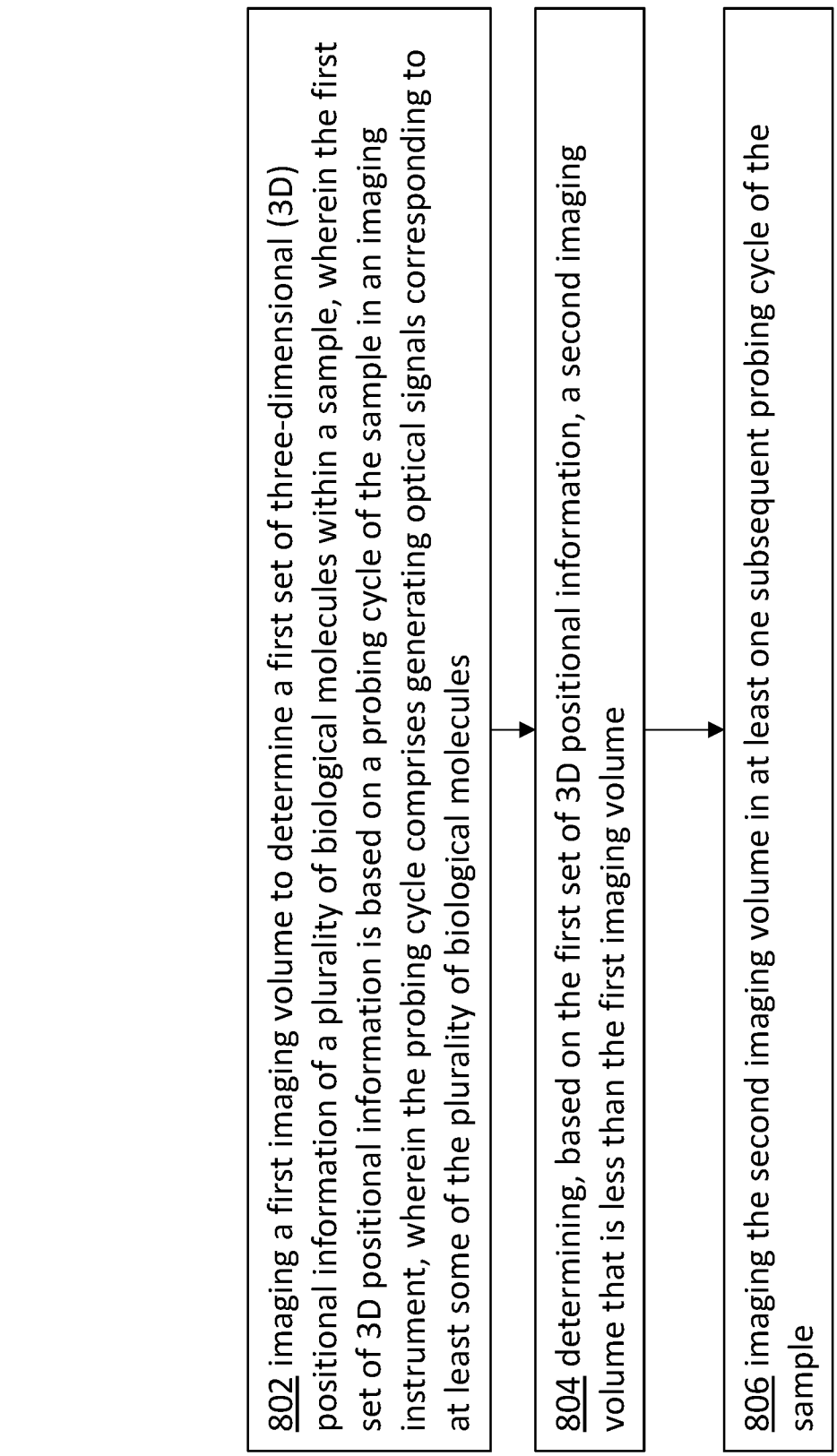

800
802 imaging a first imaging volume to determine a first set of three-dimensional (3D) positional information of a plurality of biological molecules within a sample, wherein the first set of 3D positional information is based on a probing cycle of the sample in an imaging instrument, wherein the probing cycle comprises generating optical signals corresponding to at least some of the plurality of biological molecules
804 determining, based on the first set of 3D positional information, a second imaging volume that is less than the first imaging volume
806 imaging the second imaging volume in at least one subsequent probing cycle of the sample

FIG. 8

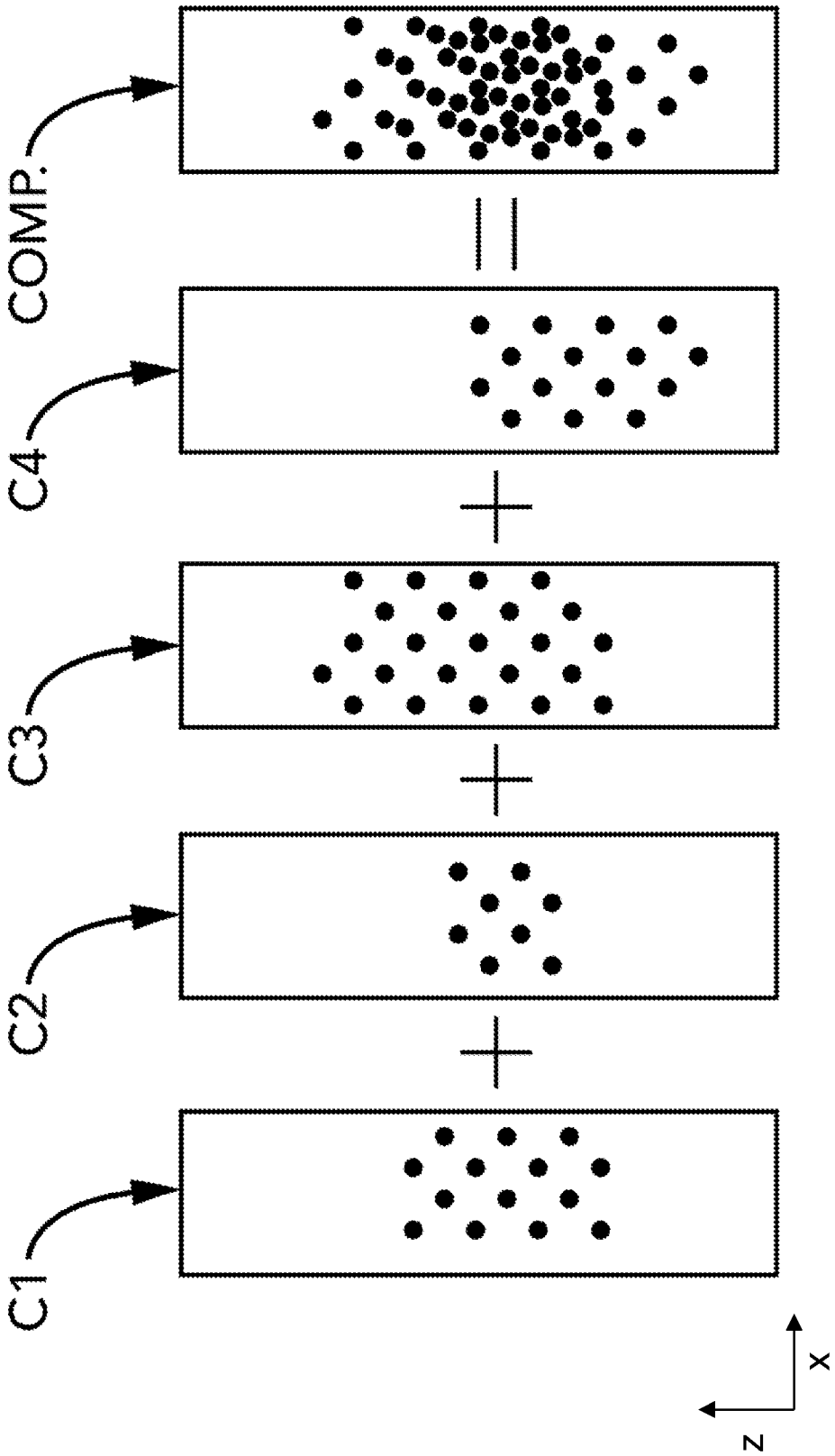


FIG. 9

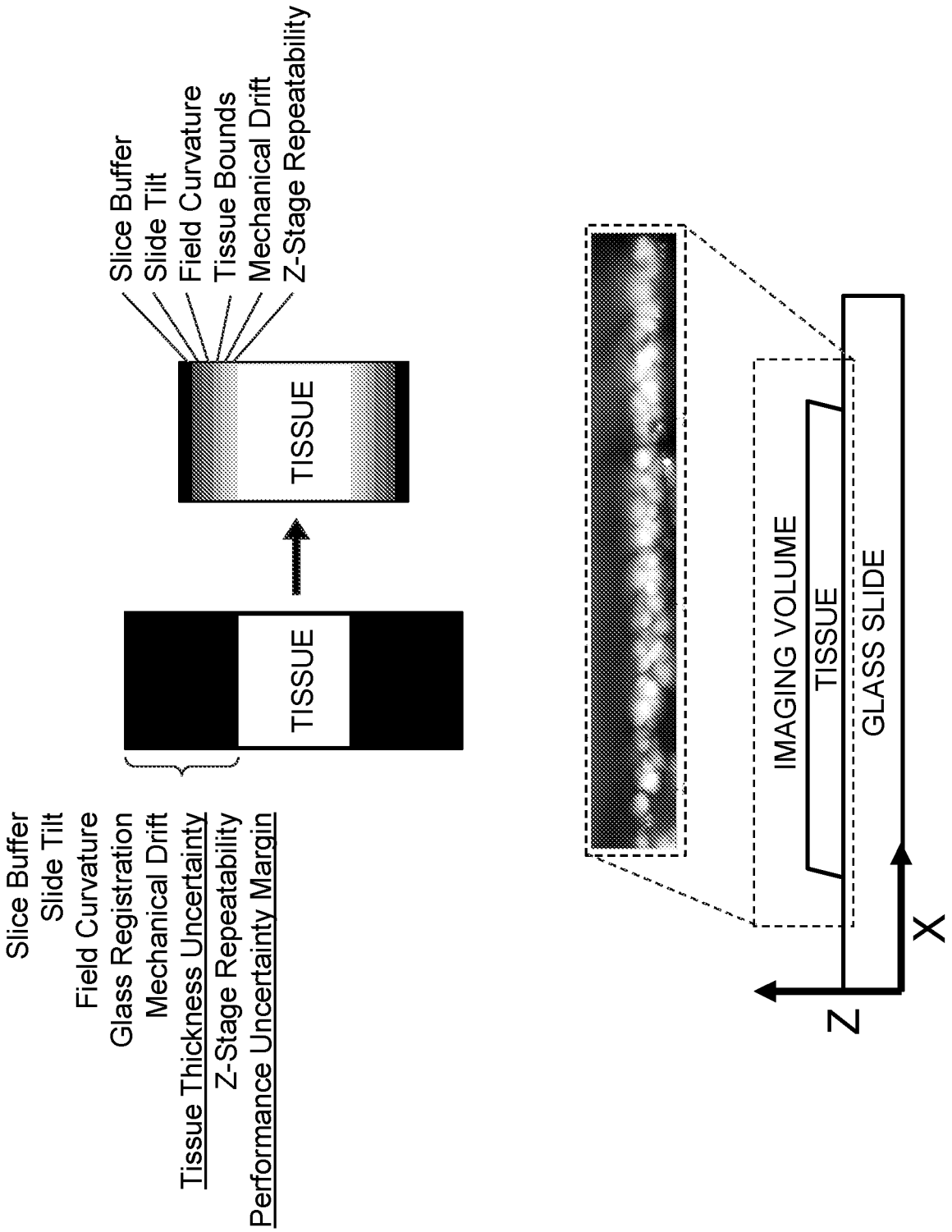

Slice Buffer
Slide Tilt
Field Curvature
Glass Registration
Mechanical Drift
Tissue Thickness Uncertainty
Z-Stage Repeatability
Performance Uncertainty Margin
TISSUE
Slice Buffer
Slide Tilt
Field Curvature
Tissue Bounds
Mechanical Drift
Z-Stage Repeatability
TISSUE
IMAGING VOLUME
TISSUE
GLASS SLIDE
Z
X

FIG. 10

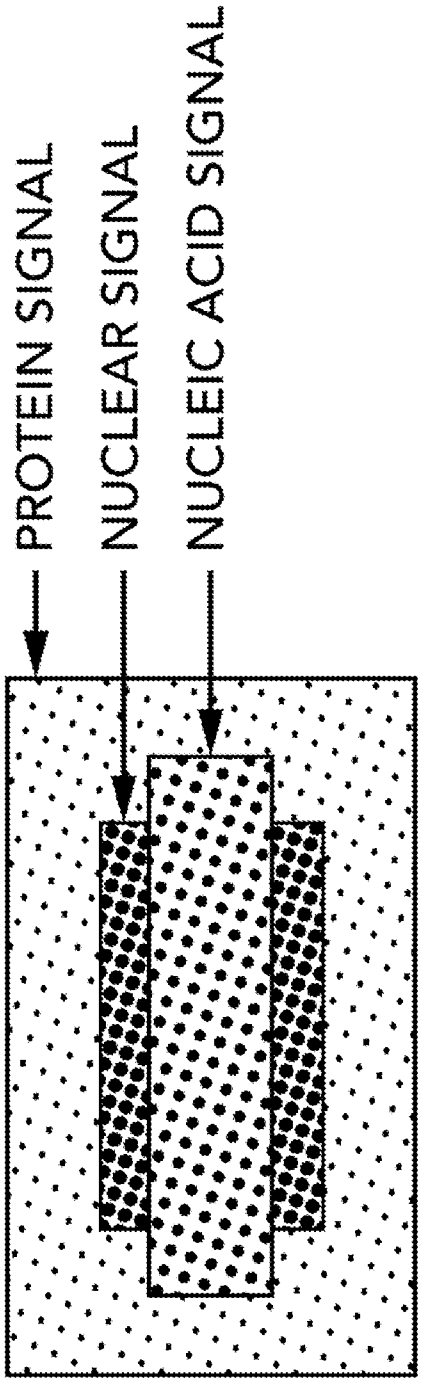


FIG. 11

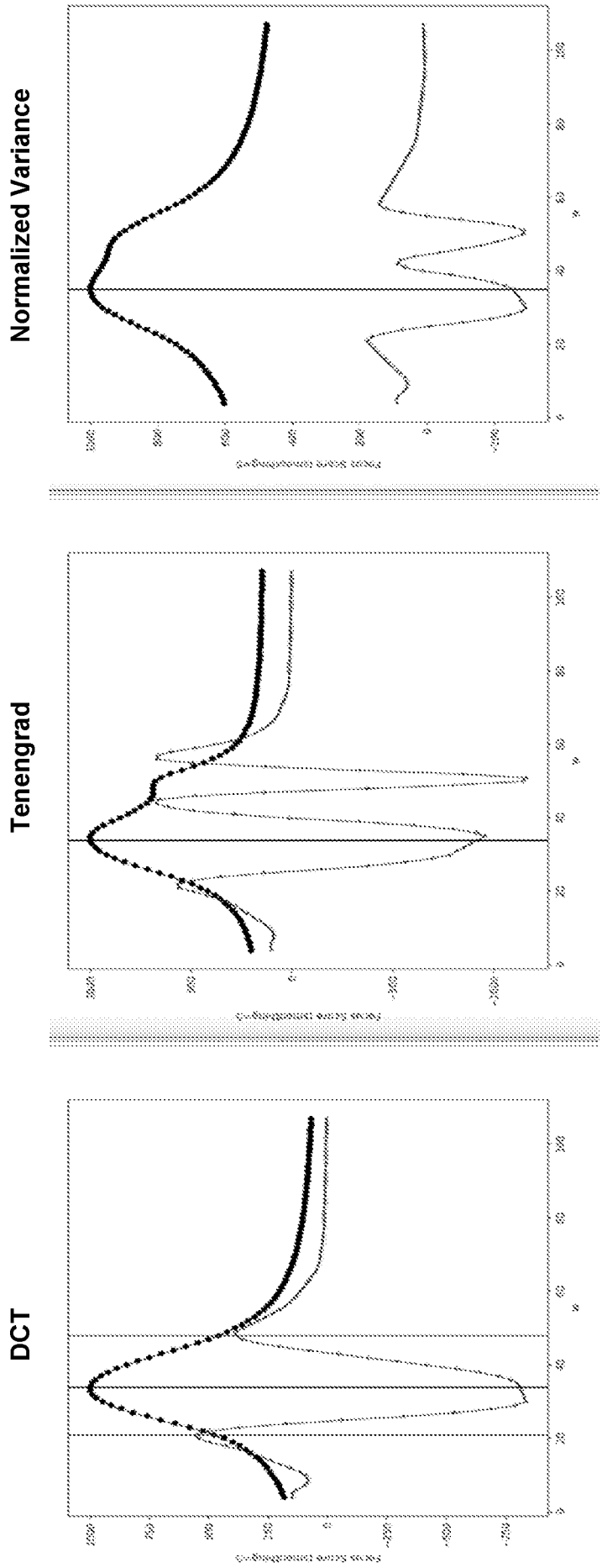


FIG. 12

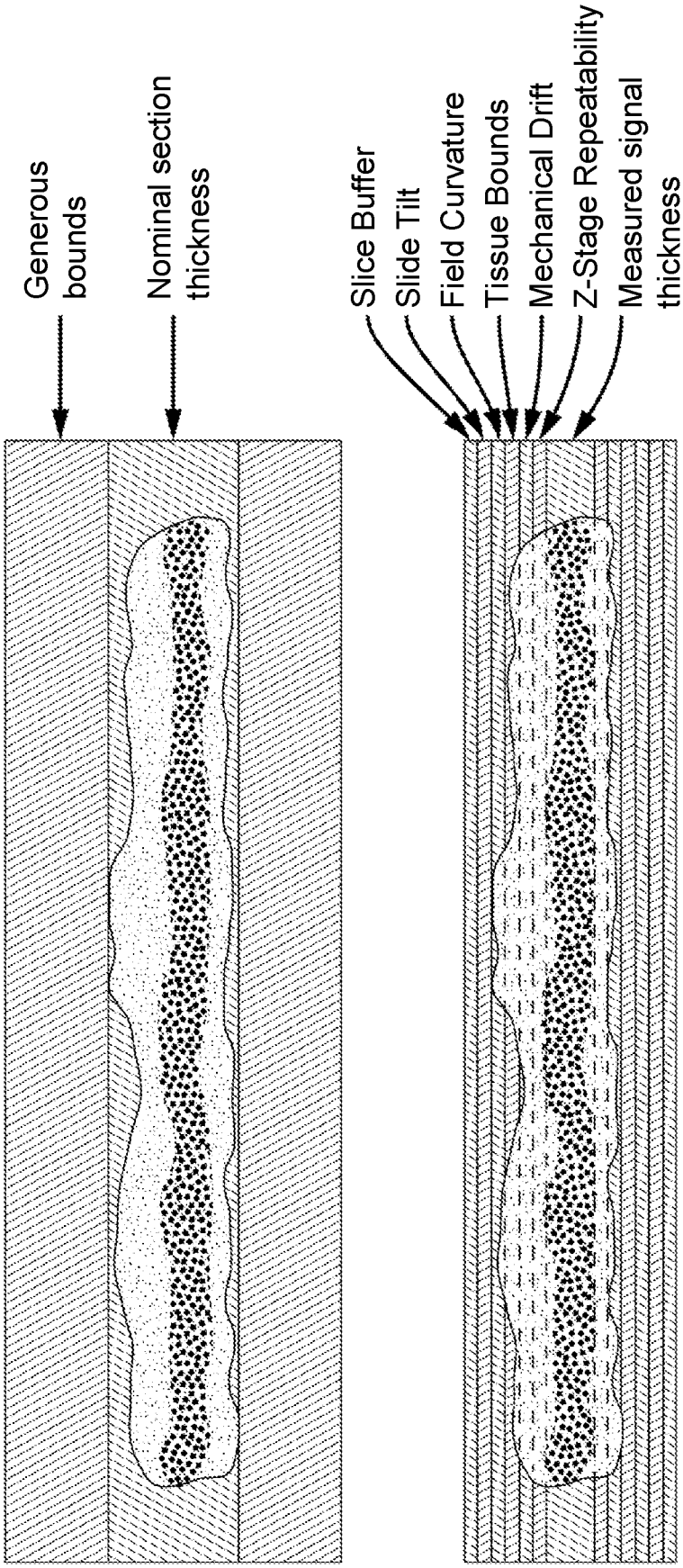

Generous bounds
Nominal section thickness
Slice Buffer
Slide Tilt
Field Curvature
Tissue Bounds
Mechanical Drift
Z-Stage Repeatability
Measured signal thickness

FIG. 13

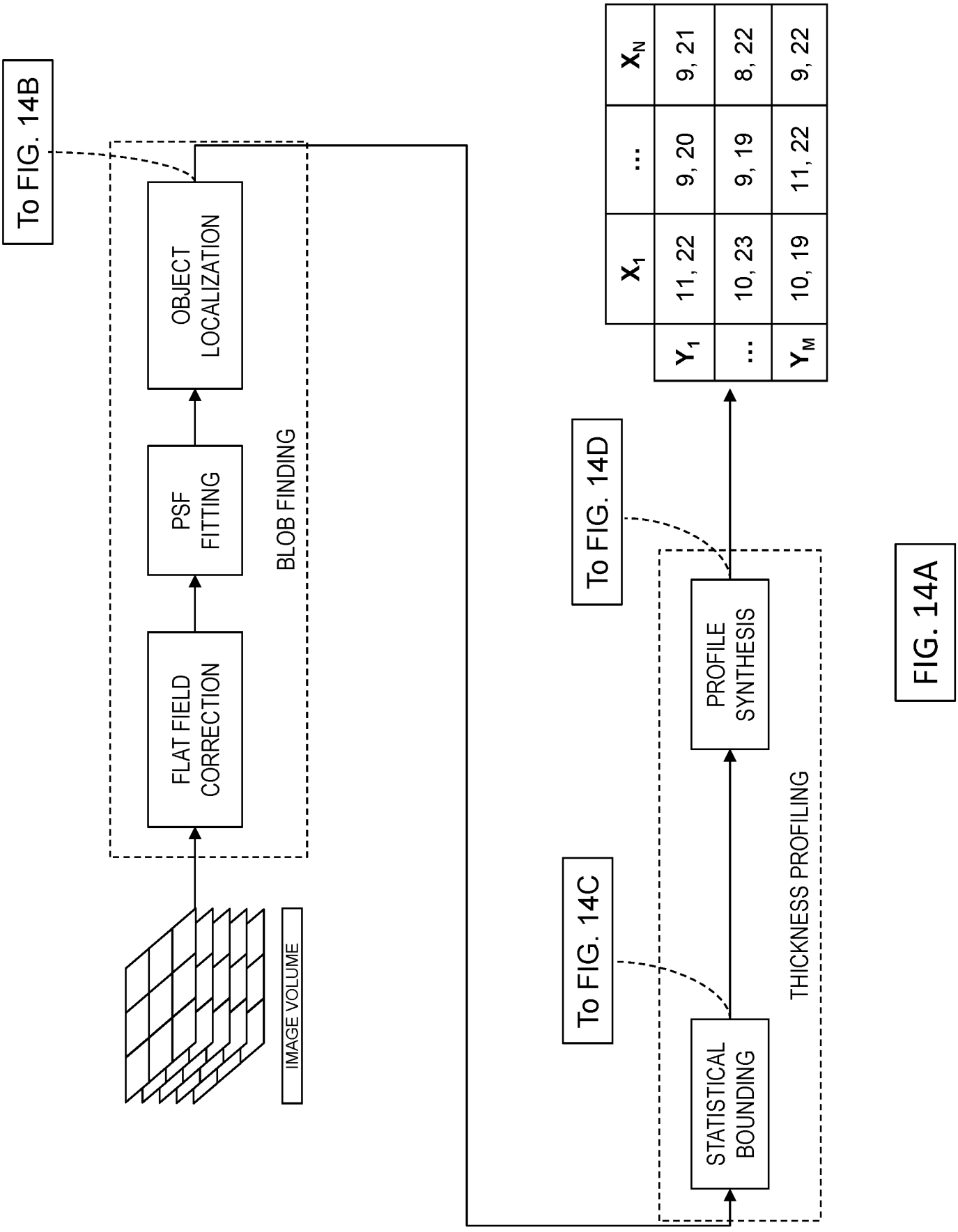

To FIG. 14B
OBJECT LOCALIZATION
PSF FITTING
FLAT FIELD CORRECTION
BLOB FINDING
IMAGE VOLUME
To FIG. 14C
To FIG. 14D
STATISTICAL BOUNDING
PROFILE SYNTHESIS
THICKNESS PROFILING
X_1 ... X_N
Y_1 11, 22 9, 20 9, 21
... 10, 23 9, 19 8, 22
Y_M 10, 19 11, 22 9, 22

FIG. 14A

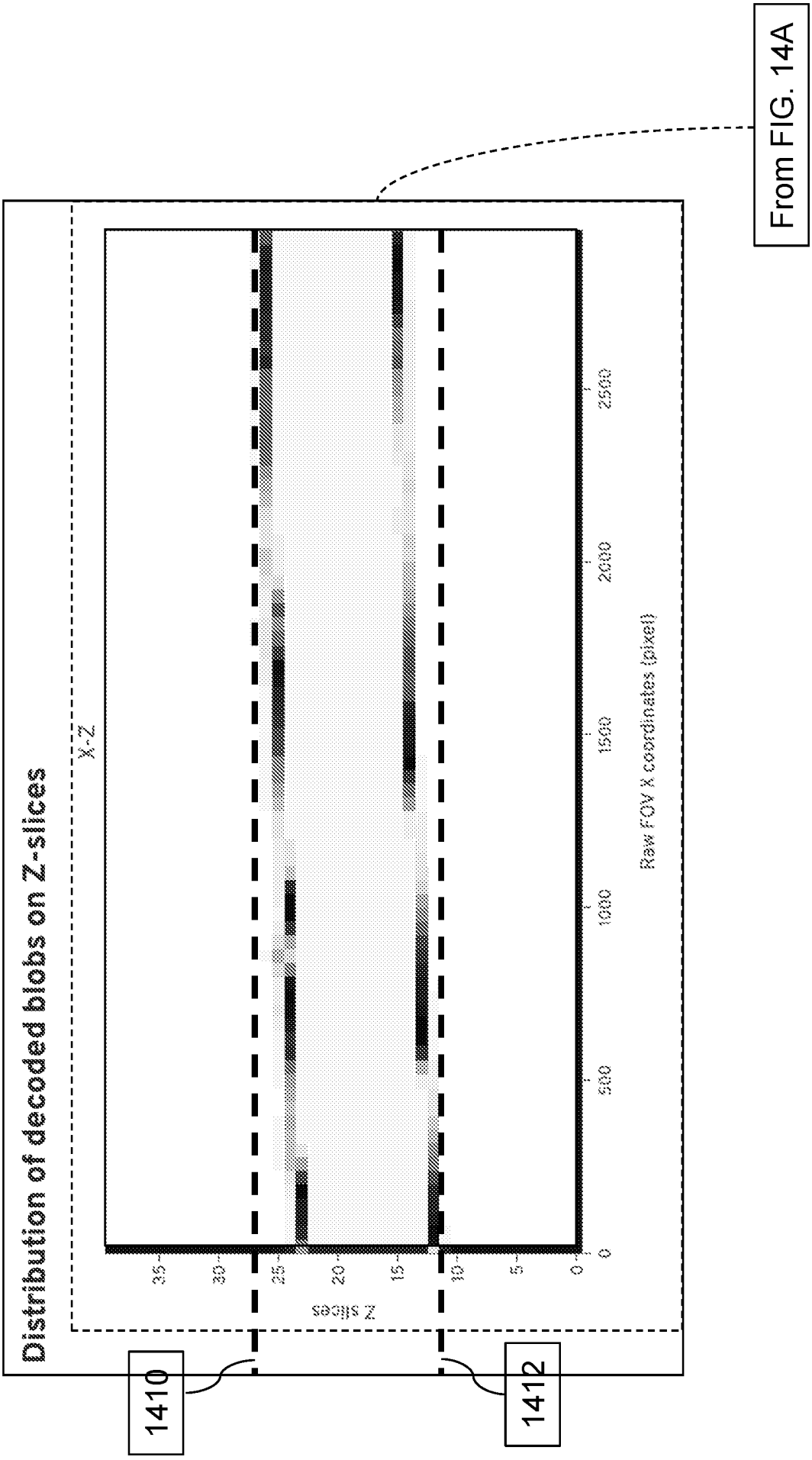
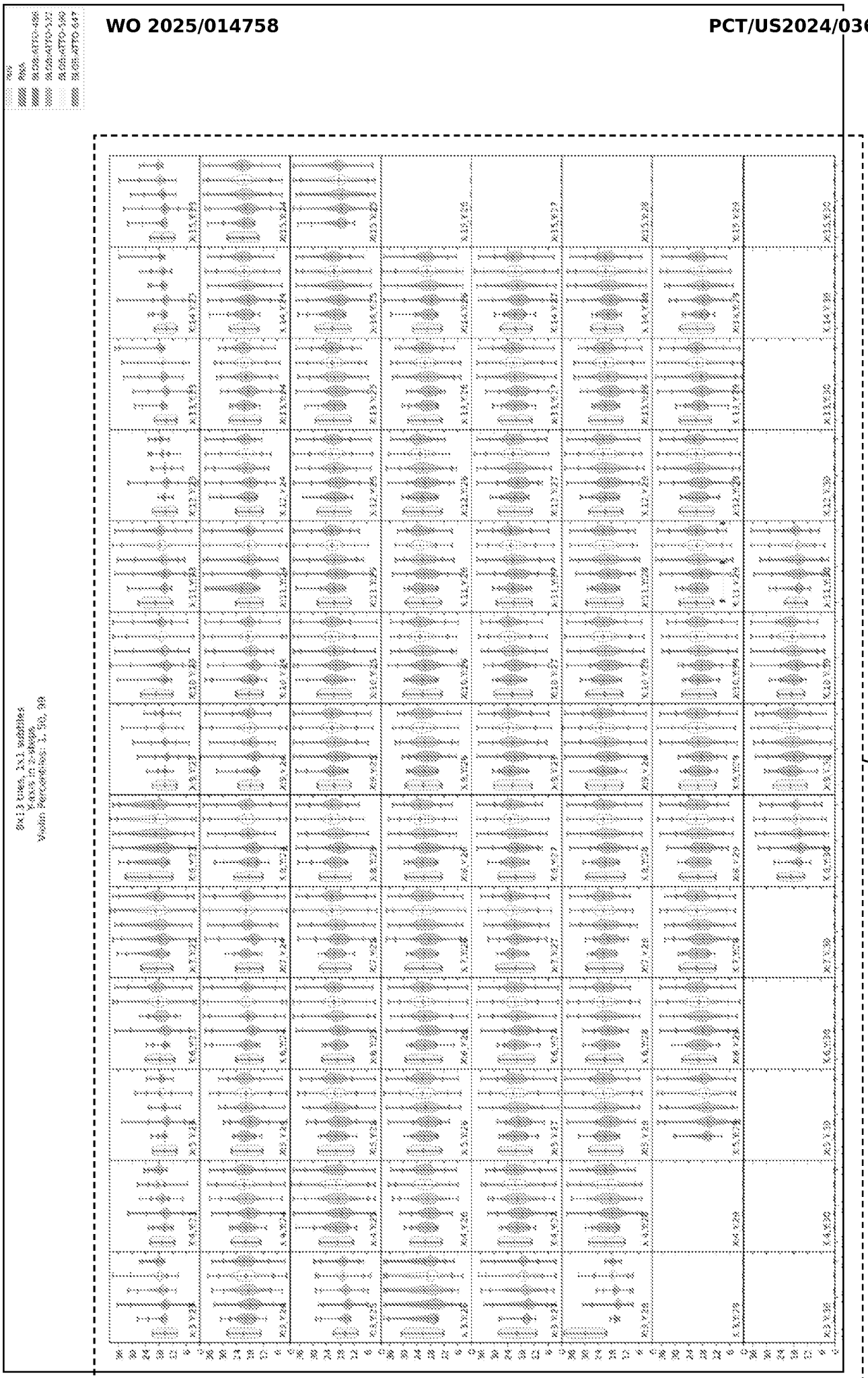
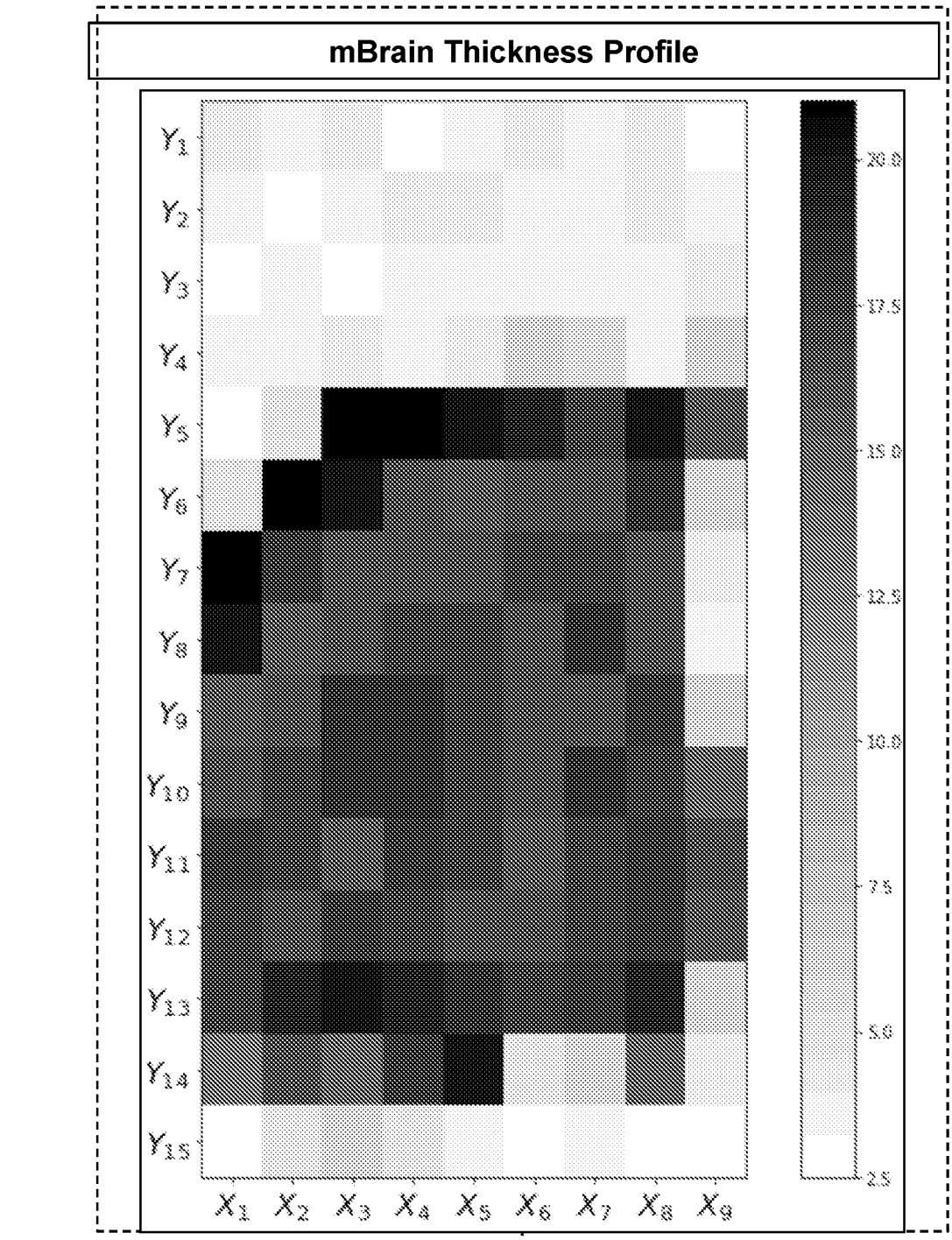

From FIG. 14A
Distribution of decoded blobs on Z-slices
X-Z
Z slices
Raw FOV X coordinates (pixel)
0
5
10
15
20
25
30
35
0
500
1000
1500
2000
2500
1410
1412

FIG. 14B



From FIG. 14A

FIG. 14C



From FIG. 14A

FIG. 14D

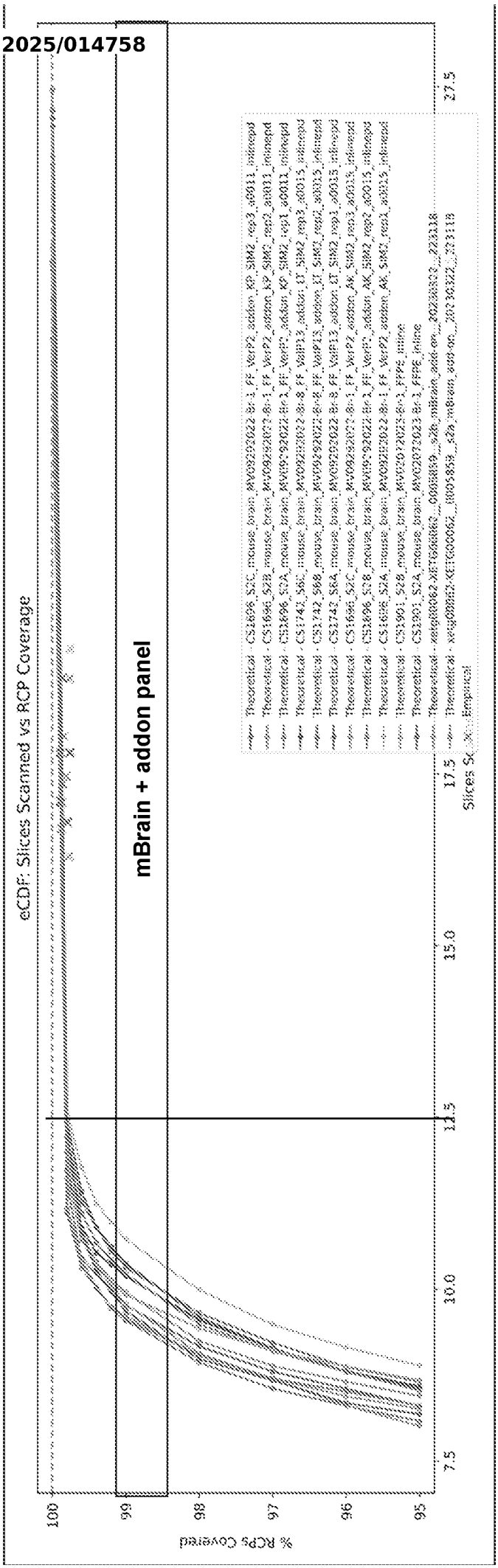

eCDF: Slices Scanned vs RCP Coverage
mBrain + addon panel
% RCPs Covered
100
99
98
97
96
95
7.5
10.0
12.5
15.0
17.5
Slices Sc
Theoretical
Empirical

FIG. 15

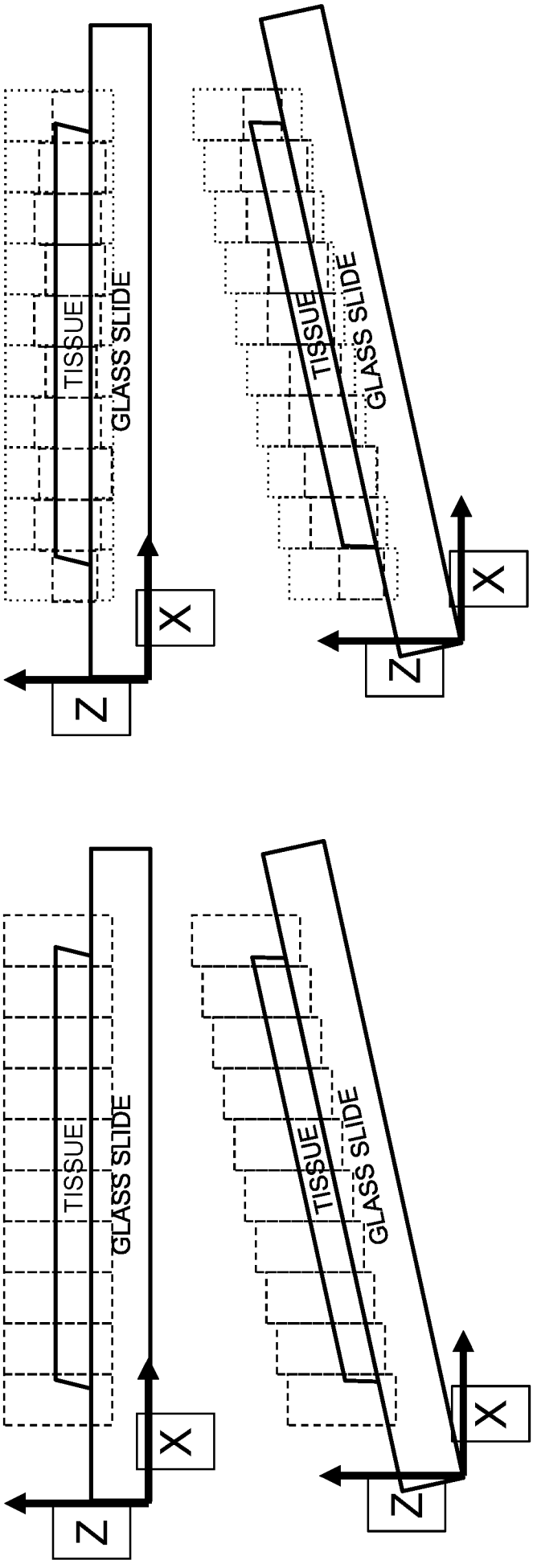

TISSUE
GLASS SLIDE
X
Z
TISSUE
GLASS SLIDE
X
Z
TISSUE
GLASS SLIDE
X
Z
TISSUE
GLASS SLIDE
X
Z

FIG. 16

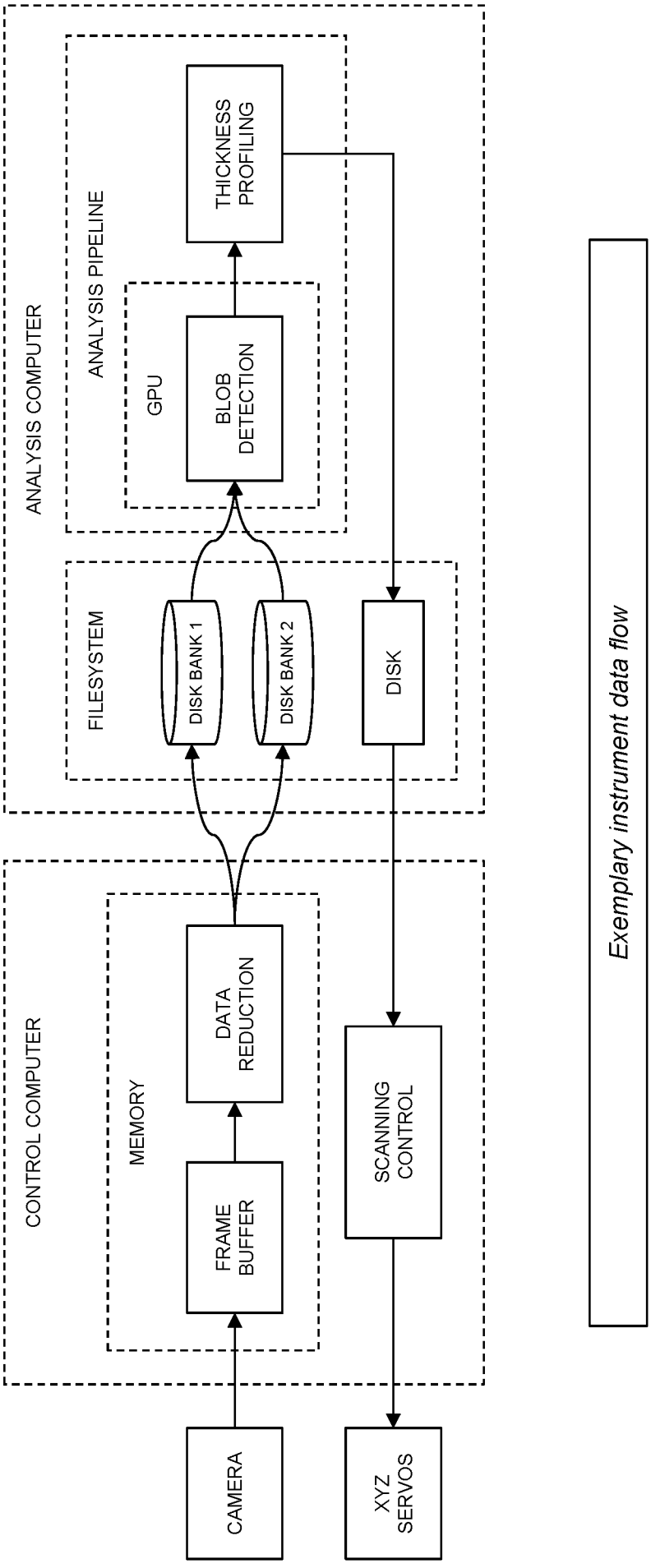

ANALYSIS COMPUTER
ANALYSIS PIPELINE
GPU
BLOB DETECTION
THICKNESS PROFILING
FILESYSTEM
DISK BANK 1
DISK BANK 2
DISK
CONTROL COMPUTER
MEMORY
DATA REDUCTION
FRAME BUFFER
SCANNING CONTROL
CAMERA
XYZ SERVOS
Exemplary instrument data flow

FIG. 17

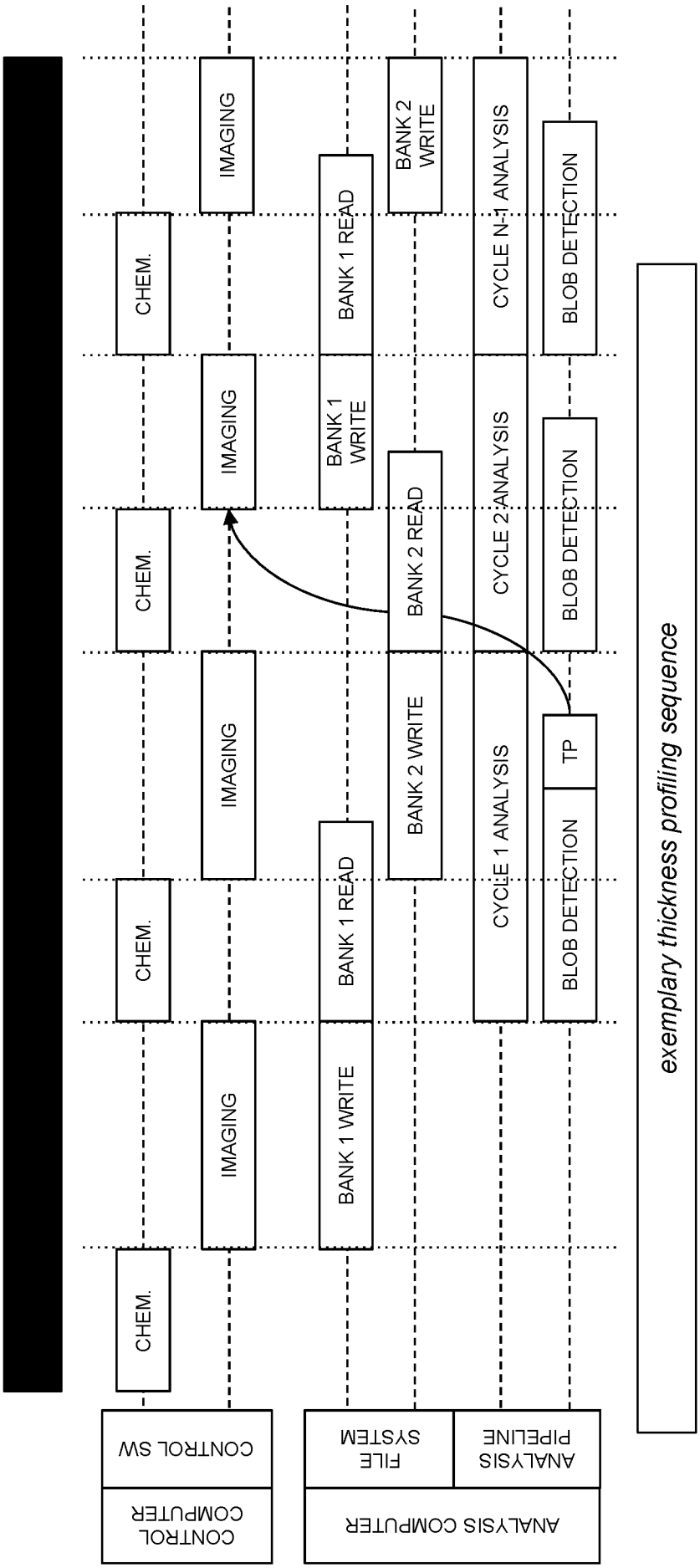

CONTROL COMPUTER
CONTROL SW
ANALYSIS COMPUTER
FILE SYSTEM
ANALYSIS PIPELINE
CHEM.
IMAGING
BANK 1 WRITE
BANK 1 READ
BANK 2 WRITE
BANK 2 READ
CYCLE 1 ANALYSIS
CYCLE 2 ANALYSIS
CYCLE N-1 ANALYSIS
BLOB DETECTION
TP
exemplary thickness profiling sequence

FIG. 18

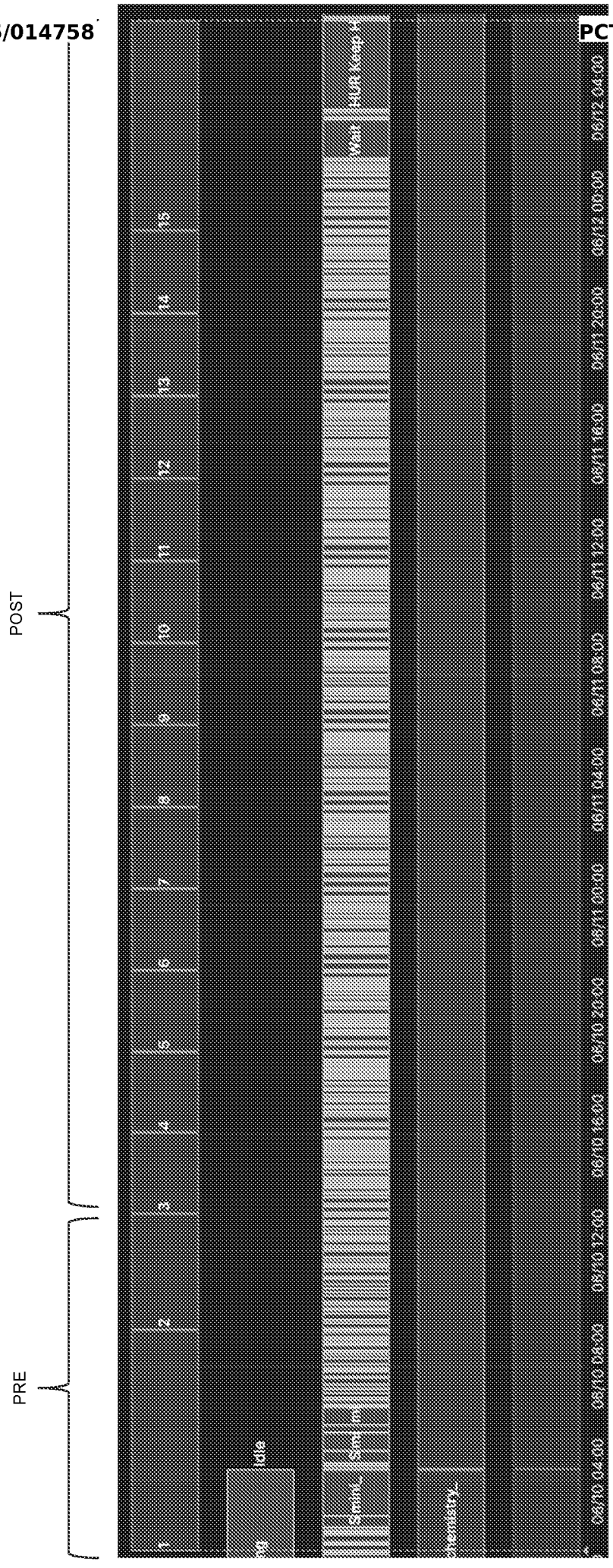

POST
PRE
Idle
Sleep
Wait
Run Keep 1
Memory
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
08/10 04:00
08/10 08:00
08/10 12:00
08/10 16:00
08/10 20:00
08/11 00:00
08/11 04:00
08/11 08:00
08/11 12:00
08/11 16:00
08/11 20:00
08/12 00:00
08/12 04:00

FIG. 19

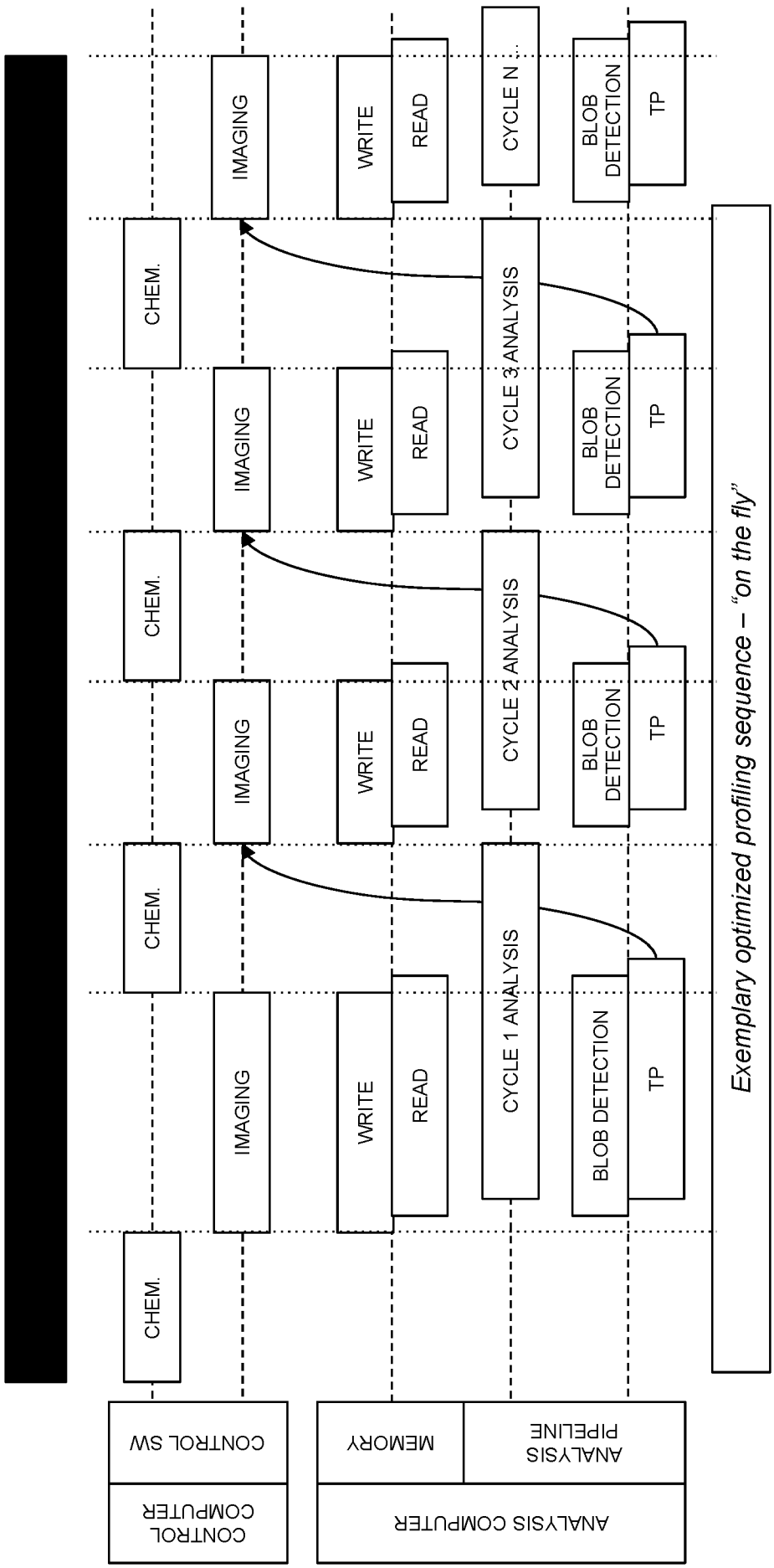

CONTROL COMPUTER
CONTROL SW
ANALYSIS COMPUTER
MEMORY
ANALYSIS PIPELINE
CHEM.
IMAGING
WRITE
READ
CYCLE 1 ANALYSIS
BLOB DETECTION
TP
CHEM.
IMAGING
WRITE
READ
CYCLE 2 ANALYSIS
BLOB DETECTION
TP
CHEM.
IMAGING
WRITE
READ
CYCLE 3 ANALYSIS
BLOB DETECTION
TP
CHEM.
IMAGING
WRITE
READ
CYCLE N ...
BLOB DETECTION
TP
Exemplary optimized profiling sequence – "on the fly"

FIG. 20

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/036755

A. CLASSIFICATION OF SUBJECT MATTER

INV. G02B21/36
ADD. G01N15/14

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G02B G01N

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2023/069991 A1 (BOSIO ANDREAS [DE] ET AL) 9 March 2023 (2023-03-09)	1-6, 8-31, 33-56, 58-78, 81-83
A	figure 3 paragraphs [0001], [0012], [0018], [0024], [0029], [0044], [0049], [0133], [0123], [0137]	7,32,57, 79,80, 84-86

-/--

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 November 2024	15/11/2024

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

1

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

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/036755

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Sergio Marco Salas: "Optimizing Xenium In Situ data utility by quality assessment and best practice analysis workflows", bioRxiv, 14 February 2023 (2023-02-14), XP093157105, DOI: 10.1101/2023.02.13.528102 Retrieved from the Internet: URL:https://www.biorxiv.org/content/10.1101/2023.02.13.528102v1.full.pdf [retrieved on 2024-11-04]	1-78, 81-83
A	abstract; figure 1 page 2, left-hand column, paragraph 2 - right-hand column, paragraph 1	79,80, 84-86

X	WO 2023/108139 A2 (10X GENOMICS INC [US]) 15 June 2023 (2023-06-15) paragraphs [0037], [0109], [0151] - [0153], [0224], [0271] - [0273], [0301], [0324], [0337], [0349], [0363]; claim 96	1-86

A	RONGQIN KE ET AL: "In situ sequencing for RNA analysis in preserved tissue and cells", NATURE METHODS, vol. 10, no. 9, 14 July 2013 (2013-07-14), pages 857-860, XP055163946, ISSN: 1548-7091, DOI: 10.1038/nmeth.2563 abstract; figures 1-3 page 857, right-hand column, paragraph 2 - page 860, right-hand column	1-86

A	LAPIERRE-LANDRY MARYSE ET AL: "Nuclei Detection for 3D Microscopy With a Fully Convolutional Regression Network", IEEE ACCESS, vol. 9, 2 January 2021 (2021-01-02), pages 60396-60408, XP093220680, USA ISSN: 2169-3536, DOI: 10.1109/ACCESS.2021.3073894 [retrieved on 2024-11-04] abstract; figure 1 page 60398, right-hand column, paragraph 2 - page 60401, right-hand column, paragraph 2	1-86
	-----	
1 A	US 2022/043251 A1 (MOORE RODGER MICHAEL [US] ET AL) 10 February 2022 (2022-02-10) figures 1,2 paragraph [0062] - paragraph [0071]	1-86

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2024/036755

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2023069991 A1	09-03-2023	EP 4111194 A1	04-01-2023
		US 2023069991 A1	09-03-2023
		WO 2021170710 A1	02-09-2021
WO 2023108139 A2	15-06-2023	EP 4445377 A2	16-10-2024
		US 2024331348 A1	03-10-2024
		WO 2023108139 A2	15-06-2023
US 2022043251 A1	10-02-2022	AU 2019403134 A1	17-06-2021
		BR 112021011719 A2	31-08-2021
		CA 3122853 A1	25-06-2020
		CN 113939728 A	14-01-2022
		EP 3899500 A1	27-10-2021
		JP 2022513494 A	08-02-2022
		KR 20210113236 A	15-09-2021
		US 2022043251 A1	10-02-2022
		US 2022260826 A1	18-08-2022
		US 2024248294 A1	25-07-2024
		WO 2020131864 A1	25-06-2020