

VOLUMETRIC NEXT-GENERATION *IN SITU* SEQUENCER

BACKGROUND OF THE INVENTION

[0001] Biological samples contain complex and heterogenous genetic information spanning the length scales of individual cells and whole tissues. Spatial patterns of nucleic acids within a cell may reveal properties and abnormalities of cellular function; cumulative distributions of RNA expression may define a cell type or function; and systematic variation in the locations of cell types within a tissue may define tissue function. The combination of anatomical connectivity information encoded in nucleic acids and tissue-wide cell type distributions may span many tissue regions and sections. Techniques for *in situ* nucleic acid sequencing must therefore be able to bridge resolutions as small as individual molecules and as large as entire brains. Efficiently collecting and recording this information across orders-of-magnitude differences in lengths requires novel inventions to enhance the robustness, rapidity, automated-, and high throughput-nature of *in situ* sequencing techniques.

SUMMARY OF THE INVENTION

[0002] A sequencing device for automated *in situ* sequencing of volumetric tissue samples is provided. In particular, an automated volumetric *in situ* sequencing device capable of operating at high resolution on multiple samples in parallel is provided. The sequencing device combines automated immersion with automated *in situ* sequencing functions. The sequencing device is especially useful for combinatorial sequencing, which benefits from its high resolution capability. Methods of fabrication and use of the sequencer are also provided.

[0003] In one aspect, a sequencing device is provided, the device comprising: (a) an illumination and detection module comprising a spinning disk confocal component comprising a plurality of laser lines for illumination with flat illumination correction, wherein the plurality of laser lines are used to illuminate a sample with excitation light at one or more wavelengths, a bandpass emission filter, a long-pass image splitter, a first camera that detects fluorescence emissions in a first wavelength range and a second camera that detects fluorescence emissions in a second wavelength range, wherein the first camera and the second camera can detect emissions simultaneously; (b) a microscope module comprising a motorized stage capable of multi-axis positioning along x, y, and z axes, an objective Z drive, an objective turret wheel comprising multiple objectives, wherein each objective provides a different magnification, wherein one or more objectives are immersion objectives, wherein each immersion objective has an objective immersion collar and optics, wherein the optics route light from the objectives to the illumination and detection module; (c) an automated immersion media module comprising i) a container comprising immersion media, ii) fluidic lines

coupled to the container and to the objective immersion collars of the objectives of the microscope module, wherein the fluidic lines carry immersion media to and from the objective immersion collars, wherein the immersion collars capture excess immersion media, and iii) a series of pumps connected to the fluidic lines and to a microcontroller, wherein the microcontroller controls the pumps addition and removal of the immersion media through the fluidic lines, wherein the automated immersion media module provides controlled volumes of the immersion media to the objective immersion collars at the tops of the objectives during imaging; (d) a multi-well plate, wherein the motorized stage can be moved to position a well of the multi-well plate under the objective used for imaging; (e) a fluidic coupling tower, wherein the fluidic coupling tower is on top of the motorized stage and positions the fluidic lines in wells of the multi-well plate; (f) a fluidic management module comprising a symmetrical rotary valve comprising a rotary valve mechanism, a pump, wherein the pump is connected to the fluidic lines, and bubble detectors, wherein the bubble detectors are positioned on either side of the fluidic lines leading to the pump, wherein the fluidic management module allows unidirectional or bidirectional movement of reagents, buffers, and waste through the fluidic lines; (g) a reagent, buffer, and waste module comprising a i) sliding tray, wherein reagent cartridges and buffer cartridges can be positioned in the sliding tray and coupled to the fluidic management module, ii) a waste module comprising a waste container, wherein the waste container is coupled to a fluidic line from the fluid management pump, and iii) a capping mechanism, wherein the capping mechanism closes the waste container when the waste container is removed from the system for waste disposal and opens the waste container when the waste container is placed back into the system; (h) an electrical module comprising: i) a first firmware board controlling media dispensing from the automated immersion media module and ii) a second firmware board controlling the fluid management module and the reagent, buffer, and waste module, wherein the electrical module regulates power to the other modules of the system; and (i) a processor programmed to provide a user interface and operate the modules of the sequencing device.

[0004] In certain embodiments, the plurality of laser lines comprises at least 4 laser lines. In certain embodiments, the plurality of laser lines comprises at least 5 laser lines. In some embodiments, the bandpass emission filter is a penta-bandpass emission filter.

[0005] In certain embodiments, the motorized stage has a piezo z-axis.

[0006] In certain embodiments, the immersion media is water.

[0007] In certain embodiments, the immersion media is filtered and bubble-free.

[0008] In certain embodiments, the sequencing device further comprises an O-ring and a shrink-wrapped coating over each objective.

- [0009] In certain embodiments, the sequencing device further comprises a pressure monitor to monitor pressure in the fluidic lines, wherein increases in pressure in a fluidic line can be used to detect a potential blockage of the fluidic line.
- [0010] In certain embodiments, the sequencing device further comprises a plurality of light-emitting diodes (LEDs), wherein each LED can emit light to provide a status indication for the system.
- [0011] In certain embodiments, the sequencing device further comprises a display component for displaying information and providing a user interface.
- [0012] In certain embodiments, the processor is further programmed to perform steps comprising: (a) locating a selected sample in the multi-well plate; (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning; (c) using the signal to segment an XY bounding box around the sample; (d) imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent; (e) displaying the image produced in step (d); (f) providing an interface for a user to refine a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample; (g) imaging the sample in the selected XY region of interest across the previously sampled Z extents; (h) calculating a volume of the region of interest in the sample and displaying the calculated sample volume of the region of interest to the user; (i) segmenting the image of the sample in the region of interest along the Z extents; (j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and (k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.
- [0013] In certain embodiments, the processor is further programmed to perform steps comprising: providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents that are available and the selected sequencing protocol; providing constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints.

- [0014] In certain embodiments, the processor is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.
- [0015] In certain embodiments, the processor is further programmed to perform steps comprising: performing a rapid confocal sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position; determining the sample top and bottom interface using a segmentation method; and setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.
- [0016] In certain embodiments, the sequencing is in situ sequencing of a target nucleic acid in a tissue sample. In some embodiments, the tissue sample is a thick tissue slice having a thickness of 50-200 μm . In other embodiments, the tissue sample is a thin tissue slice having a thickness of 5-20 μm . In some embodiments, the in situ sequencing is sequential or combinatorial in situ sequencing.
- [0017] In another aspect, a method of using the sequencing device, described herein, is provided, the method comprising: loading samples into the multi-well plate; selecting which samples in the multi-well plate are sequenced; selecting a sequencing protocol; and sequencing nucleic acids in the selected samples using the sequencing device described herein. In certain embodiments, the sequencing is in situ sequencing of a target nucleic acid in a tissue sample. In some embodiments, the tissue sample is a thick tissue slice having a thickness of 50-200 μm . In other embodiments, the tissue sample is a thin tissue slice having a thickness of 5-20 μm . In some embodiments, the in situ sequencing is sequential or combinatorial in situ sequencing.
- [0018] In another aspect, a computer implemented method is provided, the computer performing steps comprising: (a) locating a selected sample in the multi-well plate; (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning; (c) using the signal to segment an XY bounding box around the sample; (d) imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent; (e) displaying the image produced in step (d); (f) providing an interface for a user to select a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample; (g) imaging the sample in the selected XY region of interest across the previously sampled Z extents; (h) calculating a sample volume of the region of interest and displaying the calculated sample volume

of the region of interest to the user; (i) segmenting the image of the sample in the region of interest along Z extents; (j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and (k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.

[0019] In another aspect, a computer implemented method is provided, the computer performing steps comprising: providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents available and the selected sequencing protocol; providing constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints. In some embodiments, the computer is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.

[0020] In another aspect, a computer implemented method is provided, the computer performing steps comprising: performing a rapid confocal sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position; determining the sample top and bottom interface using a segmentation method; and setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.

[0021] In another aspect, a non-transitory computer-readable medium comprising program instructions that, when executed by a processor in a computer, causes the processor to perform any of the computer implemented methods, described herein, is provided.

[0022] In another aspect, an automated immersion media module comprising: (a) a container comprising immersion media; (b) fluidic lines coupled to the container and to the objective immersion collars of the objectives of the microscope module, wherein the fluidic lines carry immersion media to and from an objective immersion collar on an immersion objective, wherein the immersion collar captures excess immersion media; and (c) a series of pumps connected to the fluidic lines and to a microcontroller, wherein the microcontroller controls the pumps addition and removal of the

immersion media through the fluidic lines, wherein the automated immersion media module provides controlled volumes of the immersion media to the objective immersion collars at the tops of the objectives during imaging.

[0023] In another aspect, a method of using the automated immersion media module is provided, the method comprising using the automated immersion media module to deliver immersion media to an objective immersion collar attached to an immersion objective of a microscope.

[0024] In another aspect, a fluidic management module is provided, the module comprising a symmetrical rotary valve comprising a rotary valve mechanism, a pump, wherein the pump is connected to the fluidic lines, and bubble detectors, wherein the bubble detectors are positioned on either side of the fluidic lines leading to the pump, wherein the fluidic management module allows bidirectional or unidirectional movement of reagents, buffers, and waste through the fluidic lines.

[0025] In another aspect, a reagent, buffer, and waste module is provided, the module comprising: (a) a sliding tray, wherein reagent cartridges and buffer cartridges can be positioned in the sliding tray and coupled to the fluidic management module; (b) a waste module comprising a waste container, wherein the waste container is coupled to a fluidic line from the fluid management pump; and (c) a capping mechanism, wherein the capping mechanism closes the waste container when the waste container is removed from the system for waste disposal and opens the waste container when the waste container is placed back into the system.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] **FIG. 1** shows the sequencing device including its various modules.

[0027] **FIG. 2** shows a dual camera capable of dual imaging in widefield and confocal modes.

[0028] **FIG. 3** shows 5 laser lines covering 405 nm, 488 nm, 561 nm, 637 nm, and 730 nm and a beam conditioning unit.

[0029] **FIG. 4** shows a diagram of the microscope components.

[0030] **FIG. 5** shows a six-position nosepiece with objectives having 4x, 20x, 40x, and 60x magnification. Water collars are shown on the 40x and 60x objectives.

[0031] **FIG. 6** shows a hand actuated fluidic coupling system for a multi-well plate with 24 wells.

[0032] **FIG. 7** shows the fluidic coupling system with a multi-well plate on top of an XY motorized stage having a nosepiece and piezo Z.

[0033] **FIG. 8** shows the automated fluid delivery system.

[0034] **FIG. 9** shows a diagram of the components of the automated fluid delivery system.

[0035] **FIG. 10** shows an enclosure for light sensitive samples on top of a custom table that provides vibration isolation. The 5 laser lines with a cover and beam conditioning unit and a workstation with

a processor for high data throughput imaging are shown on a shelf beneath the top of the table. A 4K display component connected to the table is also shown.

- [0036] **FIG. 11** shows the board ports for the automated fluid delivery system.
- [0037] **FIG. 12** shows assembly of the modular parts of the sequencing device.
- [0038] **FIG. 13** shows schematics of a stand-alone fluidic module that couples to an existing imaging set up including 1) a microscope and table, 2) a multi-well plate, and 3) a sequencer, shown from various angles.
- [0039] **FIG. 14** shows a schematic of a multi-well plate and cover.
- [0040] **FIG. 15** shows a schematic of a multi-well plate with multiple fluidic lines connected to the multi-well plate and inserted into some selected wells of the multi-well plate.
- [0041] **FIG. 16** shows a schematic of a multi-well plate with multiple fluidic lines connected to the multi-well plate and inserted into all the wells of the multi-well plate.
- [0042] **FIG. 17** shows a multi-well plate with a cover over the wells. For each well of the multi-well plate, the cover comprises a holder for a fluidic line that guides the insertion of the fluidic line into a hole in the cover over the well.
- [0043] **FIG. 18** shows a schematic of the 1) microscope and table, 2) covered multi-well plate, and 3) sequencer from various angles with the multi-well plate connected to the sequencer or removed from the sequencer.
- [0044] **FIG. 19** shows a design of a compact automated fluid delivery system.
- [0045] **FIG. 20** shows a design of a compact automated fluid delivery system.
- [0046] **FIG. 21** shows designs for buffer and reagent trays.
- [0047] **FIG. 22** shows a design of a buffer tray containing a carrier for sealed bottles of buffers and an RFID tag for tracking.
- [0048] **FIG. 23** show alternate designs for a buffer tray. At top, is shown a buffer tray with caps for individual buffers. At bottom, is shown a buffer tray designed to hold sealed bottles of buffers.
- [0049] **FIG. 24** shows a design of a reagent tray containing a carrier for Eppendorf tubes, a seal, and an RFID tag for tracking.
- [0050] **FIG. 25** shows a weighing station for reagent filling verification, a fixture to hold reagent consumables on the scale, and a filling manifold.
- [0051] **FIG. 26** shows a buffer filling station.
- [0052] **FIG. 27** shows objectives with collars on the 40x and 60x objectives.
- [0053] **FIG. 28** shows a fluid diagram for providing fluid to the collars on the 40x and 60x objectives.
- [0054] **FIG. 29** shows a feed collar. The 60x objective has 1 O-ring and the 40x objective has 2 O-rings.

- [0055] **FIG. 30** shows a microscope with connections for an immersion water dispenser.
- [0056] **FIG. 31** shows an immersion water dispenser for use with a Nikon Ti2e microscope with connections for the 40x and 60x objectives.
- [0057] **FIG. 32** shows a schematic of an enclosure for light sensitive samples on top of a table with a shelf underneath for a workstation and a display component attached to the table.
- [0058] **FIG. 33** shows a schematic of an enclosure for light sensitive samples on top of a table with a shelf underneath for a workstation and a display component on the table.
- [0059] **FIG. 34** shows a fluidics diagram for the automated fluid delivery system showing fluid lines connections to the reagent tray, buffer tray, peristaltic pump, motor driven rotary valve, pressure sensors, and bubble detectors.
- [0060] **FIG. 35** shows a fluidics diagram with a series of pumps connected to the fluid lines with connections to the immersion media module and the 40x and 60x microscope objectives.
- [0061] **FIG. 36** shows a fluidics diagram with connections to a syringe pump, motor driven rotary valve, and multi-well plate.
- [0062] **FIG. 37** shows an aspiration dual valve schematic.

DETAILED DESCRIPTION OF THE INVENTION

- [0063] A sequencer for automated *in situ* sequencing of volumetric tissue samples is provided. In particular, an automated volumetric *in situ* sequencing device capable of operating on multiple samples in parallel is provided. Methods of fabrication and use of the sequencer are also provided.
- [0064] Before the sequencer for automated *in situ* sequencing of volumetric tissue samples and methods of fabrication and use of such a sequencer are described, it is to be understood that this invention is not limited to particular devices, methods, or compositions described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.
- [0065] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in that stated range is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included or excluded in the range, and each range where either, neither or both limits are included in the smaller ranges is also encompassed within the invention, 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 invention.

[0066] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, some potential and preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. It is understood that the present disclosure supersedes any disclosure of an incorporated publication to the extent there is a contradiction.

[0067] As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present invention. Any recited method can be carried out in the order of events recited or in any other order which is logically possible.

[0068] It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" includes a plurality of such cells and reference to "the peptide" includes reference to one or more peptides and equivalents thereof, e.g. oligopeptides or polypeptides known to those skilled in the art, and so forth.

[0069] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

Definitions

[0070] The term "about", particularly in reference to a given quantity, is meant to encompass deviations of plus or minus five percent.

[0071] The terms "peptide", "oligopeptide", "polypeptide", and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms also apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymers. Both full-length proteins and fragments thereof are encompassed by

the definition. The terms also include post-expression modifications of the polypeptide, for example, phosphorylation, glycosylation, acetylation, hydroxylation, oxidation, and the like as well as chemically or biochemically modified or derivatized amino acids and polypeptides having modified peptide backbones. The terms also include fusion proteins, including, but not limited to, fusion proteins with a heterologous amino acid sequence, fusions with heterologous and homologous leader sequences, with or without N-terminal methionine residues; immunologically tagged proteins; and the like. The terms include polypeptides including one or more of a fatty acid moiety, a lipid moiety, a sugar moiety, and a carbohydrate moiety.

[0072] As used herein, the term "target nucleic acid" is any polynucleotide nucleic acid molecule (e.g., DNA molecule; RNA molecule, modified nucleic acid, etc.) present in a single cell. In some embodiments, the target nucleic acid is a coding RNA (e.g., mRNA). In some embodiments, the target nucleic acid is a non-coding RNA (e.g., tRNA, rRNA, microRNA (miRNA), mature miRNA, immature miRNA; etc.). In some embodiments, the target nucleic acid is a splice variant of an RNA molecule (e.g., mRNA, pre-mRNA, etc.) in the context of a cell. A suitable target nucleic acid can therefore be an unspliced RNA (e.g., pre-mRNA, mRNA), a partially spliced RNA, or a fully spliced RNA, etc. Target nucleic acids of interest may be variably expressed, i.e. have a differing abundance, within a cell population, wherein the methods of the invention allow profiling and comparison of the expression levels of nucleic acids, including without limitation RNA transcripts, in individual cells. A target nucleic acid can also be a DNA molecule, e.g. a denatured genomic, viral, plasmid, etc. For example, the methods can be used to detect copy number variants, e.g. in a cancer cell population in which a target nucleic acid is present at different abundance in the genome of cells in the population; a virus-infected cells to determine the virus load and kinetics, and the like.

[0073] The terms "oligonucleotide," "polynucleotide," and "nucleic acid molecule", used interchangeably herein, refer to polymeric forms of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. Thus, this term includes, but is not limited to, single-, double-, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer including purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases. The backbone of the polynucleotide can include sugars and phosphate groups (as may typically be found in RNA or DNA), or modified or substituted sugar or phosphate groups. Alternatively, the backbone of the polynucleotide can include a polymer of synthetic subunits such as phosphoramidites, and/or phosphorothioates, and thus can be an oligodeoxynucleoside phosphoramidate or a mixed phosphoramidate-phosphodiester oligomer. Peyrottes et al. (1996) Nucl. Acids Res. 24:1841-1848; Chaturvedi et al. (1996) Nucl. Acids Res. 24:2318-2323. The polynucleotide may include one or more L-nucleosides. A polynucleotide may include modified

nucleotides, such as methylated nucleotides and nucleotide analogs, uracyl, other sugars, and linking groups such as fluororibose and thioate, and nucleotide branches. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be modified to include N3'-P5' (NP) phosphoramidate, morpholino phosphorociamidate (MF), locked nucleic acid (LNA), 2'-O-methoxyethyl (MOE), or 2'-fluoro, arabino-nucleic acid (FANA), which can enhance the resistance of the polynucleotide to nuclease degradation (see, e.g., Faria et al. (2001) *Nature Biotechnol.* 19:40-44; Toulme (2001) *Nature Biotechnol.* 19:17-18). A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Other types of modifications included in this definition are caps, substitution of one or more of the naturally occurring nucleotides with an analog, and introduction of means for attaching the polynucleotide to proteins, metal ions, labeling components, other polynucleotides, or a solid support. Immunomodulatory nucleic acid molecules can be provided in various formulations, e.g., in association with liposomes, microencapsulated, etc., as described in more detail herein. A polynucleotide used in amplification is generally single-stranded for maximum efficiency in amplification, but may alternatively be double-stranded. If double-stranded, the polynucleotide can first be treated to separate its strands before being used to prepare extension products. This denaturation step is typically affected by heat, but may alternatively be carried out using alkali, followed by neutralization.

[0074] By "isolated" is meant, when referring to a protein, polypeptide, or peptide, that the indicated molecule is separate and discrete from the whole organism with which the molecule is found in nature or is present in the substantial absence of other biological macro molecules of the same type. The term "isolated" with respect to a polynucleotide is a nucleic acid molecule devoid, in whole or part, of sequences normally associated with it in nature; or a sequence, as it exists in nature, but having heterologous sequences in association therewith; or a molecule disassociated from the chromosome.

[0075] The terms "individual", "subject", "host", and "patient", are used interchangeably herein and refer to invertebrates and vertebrates including, but not limited to, arthropods (e.g., insects, crustaceans, arachnids), cephalopods (e.g., octopuses, squids), amphibians (e.g., frogs, salamanders, caecilians), fish, reptiles (e.g., turtles, crocodilians, snakes, amphisbaenians, lizards, tuatara), mammals, including human and non-human mammals such as non-human primates, including chimpanzees and other apes and monkey species; laboratory animals such as mice, rats, rabbits, hamsters, guinea pigs, and chinchillas; domestic animals such as dogs and cats; farm animals such as sheep, goats, pigs, horses and cows; and birds such as domestic, wild and game birds, including chickens, turkeys and other gallinaceous birds, ducks, and geese. In some cases, the methods of the invention find use in experimental animals, in veterinary application, and in the

development of animal models for disease, including, but not limited to, rodents including mice, rats, and hamsters; primates, and transgenic animals.

[0076] The term “user” as used herein refers to a person that interacts with a device and/or system disclosed herein for performing one or more steps of the presently disclosed methods. The user may be a subject using the sequencing device described herein.

Sequencing Device

[0077] A sequencer for automated *in situ* sequencing of volumetric tissue samples is provided. In particular, an automated volumetric *in situ* sequencing device capable of operating on multiple samples in parallel is provided. In some embodiments, the sequencing device comprises an illumination and detection module, a microscope module, an automated immersion media module, a multi-well plate, a fluidic coupling tower or stand-alone fluidic module, a fluidic management module, a reagent, buffer, and waste module, an electrical module, and a processor.

[0078] In some embodiments, the illumination and detection module comprises a spinning disk confocal component comprising a plurality of laser lines for illumination with flat illumination correction, wherein the plurality of laser lines are used to illuminate a sample with excitation light at one or more wavelengths, a bandpass emission filter, a long-pass image splitter, a first camera that detects fluorescence emissions in a first wavelength range and a second camera that detects fluorescence emissions in a second wavelength range, wherein the first camera and the second camera can detect emissions simultaneously. In certain embodiments, the plurality of laser lines comprises at least 4 laser lines. In some embodiments, the plurality of laser lines comprises 5 laser lines used with a penta-bandpass emission filter. In certain embodiments, the motorized stage has a piezo z-axis. In other embodiments, the objective z axis drive is used and the motorized stage z is kept constant.

[0079] The microscope module comprises a motorized stage capable of multi-axis positioning along x, y, and z axes, an objective Z drive, an objective turret wheel comprising multiple objectives, wherein each objective provides a different magnification, and optics, wherein the optics route light from the objectives to the illumination and detection module. The objectives may include immersion objectives wherein each immersion objective has an objective immersion collar. In certain embodiments, an immersion objective further comprises an O-ring and a shrink-wrapped coating a multi-well plate, wherein the motorized stage, e.g., to prevent spills that can damage the optics or mechanical parts of the microscope. The objectives may also include dry objectives that have no immersion collar. In some embodiments, the microscope module comprises a confocal microscope. In some embodiments, the microscope module comprises an epifluorescent microscope.

[0080] The automated immersion media module comprises i) a container comprising immersion media, ii) fluidic lines coupled to the container and to the objective immersion collars of the objectives of the microscope module, wherein the fluidic lines carry immersion media to and from the objective immersion collars, wherein the immersion collars capture excess immersion media, and iii) a series of pumps connected to the fluidic lines and to a microcontroller, wherein the microcontroller controls the pumps addition and removal of the immersion media through the fluidic lines, wherein the automated immersion media module provides controlled volumes of the immersion media to the objective immersion collars at the tops of the objectives during imaging. In certain embodiments, the immersion media is water. In certain embodiments, the immersion media is filtered and bubble-free.

[0081] The motorized stage can be moved to position a well of the multi-well plate under the objective used for imaging. In some embodiments, a fluidic coupling tower is on top of the motorized stage and positions the fluidic lines in wells of the multi-well plate to allow addition or removal of the sample from the wells using the fluidic line. In some embodiments, the fluidic coupling interface is not attached to the motorized stage. Instead, a stand-alone fluidic coupling interface module is used, which is placed manually by a user over the sample plate and affixed to the stage, wherein the fluidic coupling interface couples the fluidic lines to the samples for the duration of sequencing.

[0082] The fluidic management module comprises a symmetrical rotary valve comprising a rotary valve mechanism, a pump, wherein the pump is connected to the fluidic lines, and bubble detectors, wherein the bubble detectors are positioned on either side of the fluidic lines leading to the pump, wherein the fluidic management module allows unidirectional or bidirectional movement of reagents, buffers, and waste through the fluidic lines. A series of bubble detectors may be used to ensure that the immersion fluid line is free of bubbles. In addition, bubbles may be avoided by adding a volume of fluid, removing excess fluid, then adding more fluid, and moving the stage to the well edge and back to the sample center to remove any additional bubbles that may form during immersion fluid addition.

[0083] The reagent, buffer, and waste module comprises a i) sliding tray, wherein reagent cartridges and buffer cartridges can be positioned in the sliding tray and coupled to the fluidic management module, ii) a waste module, wherein the waste module comprises a waste container coupled to a fluidic line from the fluid management pump, and iii) a capping mechanism, wherein the capping mechanism closes the waste container when the waste container is removed from the system for waste disposal and opens the waste container when the waste container is placed back into the system.

[0084] The electrical module comprises: i) a first firmware board controlling media dispensing from the automated immersion media module and ii) a second firmware board controlling the fluid

management module and the reagent, buffer, and waste module, wherein the electrical module regulates power to the other modules of the system.

[0085] In certain embodiments, the sequencing device further comprises a pressure monitor to monitor pressure in the fluidic lines, wherein increases in pressure in a fluidic line can be used to detect a potential blockage of the fluidic line.

[0086] In certain embodiments, the sequencing device further comprises a plurality of light-emitting diodes (LEDs), wherein each LED can emit light to provide a status indication for the system.

[0087] In certain embodiments, the sequencing device comprises a processor programmed to provide a user interface and operate the modules of the sequencing device. In some embodiments, the sequencing device further comprises a display component for displaying information and providing a user interface.

[0088] In certain embodiments, the processor is further programmed to perform steps comprising: (a) locating a selected sample in the multi-well plate; (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning; (c) using the signal to segment an XY bounding box around the sample; (d) imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent; (e) displaying the image produced in step (d); (f) providing an interface for a user to refine a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample; (g) imaging the sample in the selected XY region of interest across the previously sampled Z extents; (h) calculating a volume of the region of interest in the sample and displaying the calculated sample volume of the region of interest to the user; (i) segmenting the image of the sample in the region of interest along the Z extents; (j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and (k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.

[0089] In certain embodiments, the processor is further programmed to perform steps comprising: providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents that are available and the selected sequencing protocol; providing

constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints. In certain embodiments, the processor is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.

[0090] In certain embodiments, the processor is further programmed to perform steps comprising: performing a rapid confocal or epifluorescent sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position; determining the sample top and bottom interface using a segmentation method; and setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.

[0091] In certain embodiments, the sequencing is in situ sequencing of a target nucleic acid in a tissue sample. In some embodiments, the tissue sample is a thick tissue slice having a thickness of 50-200 μm . In other embodiments, the tissue sample is a thin tissue slice having a thickness of 5-20 μm . In some embodiments, the in situ sequencing is sequential or combinatorial in situ sequencing.

Modular Uses of Sequencer Components

[0092] In some embodiments, the fluidics components may act as a standalone sequencing module for use with any compatible imaging system. The fluid lines to the sample may be coupled to a sample plate and microscope stage magnetically and/or mechanically, such that the coupling is easily attachable to and detachable from an engaged position, and such that the fluidic components are coupled to sample wells. In one instance of this coupling, fluid lines addressed to each sample well are bundled together and routed to each sample well via a detachable plate lid (see schematic), which couples via a mechanical guide and magnetic fixture to a microscope stage. In another embodiment of the coupling, a modular coupling tower is provided to be affixed to the microscope stage. When used as a standalone sequencing module, the fluidic components facilitate the use of reagent and buffer kits and the automation of fluid exchange from multiple sample wells across multiple cycles of fluid addition and removal. When used as a device for in situ sequencing, the fluidic components may be coupled to an existing microscopy set up that is compatible with the sample format, for example, an inverted microscope. When used with thin section samples (5-20 μm), the microscope may be an epifluorescent microscope with 3, 4, or 5 illumination or detection channels. When used with thin or thick section samples, the microscope may be an epifluorescent microscope,

a confocal microscope (spinning disk or point scanning), a structured illumination microscope, or a light sheet or oblique-plane light sheet microscope.

[0093] An immersion water dispensing module (IWD) can be used as a submodule of an integrated fluidic system for a sequencer. Alternatively, it can be used as a standalone kit for automated immersion of microscope immersion objectives. In one instance it is used in tandem with the reagent/buffer/consumables fluidic module to enable parallel and automated sequencing of samples on a separate and existing microscope set up. In this instance the immersion fluid reservoir and the immersion fluid waste are external to the fluidic device so that the user may manually fill the immersion fluid reservoir and empty the waste reservoir. The immersion water dispensing module connects to microscope objectives via the immersion collar, which is designed to flow immersion liquid across the imaging glass of the objective such that no bubbles are introduced and the volume of liquid and flow rate are precise and consistent, while providing a tight seal against the objective body such that excess liquid can be removed. The exact dimensions of the immersion collar are adjusted to match specific objective lenses to ensure a proper fit but the function of the other immersion water dispensing module subcomponents are agnostic to the make and manufacture of the imaging system.

[0094] The software controlling the sequencer provides an abstraction layer over the control of the illumination, detection, microscope, and stage components and thus may be used modularly with a variety of imaging set ups, provided that an appropriate configuration file or other hardware plugins are provided. Thus, a particular imaging and microscopy set up is not privileged in the operation of the sequencer and the software, objective immersion module, sample, reagent and buffer fluidics, and consumables may be used modularly and reconfigured into one or more combinations of components.

[0095] In some embodiments, the reagent and buffer fluid components draw fluids from reusable reservoirs. In another embodiment of the sequencer, the reagent and buffer fluid components draw fluids from consumable reservoirs. In one instance consumable reservoirs are sealed after they are filled, and the seal is punctured by the sipper needles of the reagent/buffer fluidic module. In one instance the seals are supported mechanically in the assembly of the consumable to ensure consistent puncturing of the seal and to avoid excess forces on the sipper needles or forces not aligned to the parallel axes of the sipper needles. Consumable reservoirs are normally replaced at the beginning of each use of the sequencer and their use is tracked programmatically through detection of the identity of the consumable. In one instance the detection of the consumable performed through the use of a RFID integrated into the consumable and an RFID reader in the sequencer fluidic module. In another embodiment the detection of the consumable is performed

through the use of a barcode on the consumable and a barcode scanner integrated into the sequencer fluidic module.

[0096] In another embodiment, the fluid module draws from buffers and reagents used in in situ sequencing cycles. In one instance of the fluidic module, some or all of the buffers or reagents are chilled by a refrigeration component. In some embodiments of the fluidic module, some or all of the buffers or reagents are temperature sensitive, for example, enzymes such as ligase or molecules such as ATP involved in SCAL, SEDAL, or SEDAL2 sequencing chemistries. In another embodiment, the buffer and reagent fluidic modules draw liquids used in other sequencing or cyclical labeling chemistries, for instance, oligos used to hybridize to sequences in a sample, or fluorescently labeled oligos used to detect hybridization events in a sample. In another embodiment, the buffer and reagent fluidic modules draw liquids used for the labeling of samples with dyes. In another embodiment, the buffer and reagent fluidic modules draw liquids used for CLICK chemistry reactions with the sample. In another embodiment, the buffer and reagent fluidic modules draw liquids to quench fluorescent signal in the sample. In another embodiment, the buffer and reagent fluidic modules draw fluids containing enzymatic components that add or remove signals from a sample.

Computer Implemented Methods

[0097] The present disclosure provides systems and computer implemented methods which find use in using the sequencing device described herein. In certain embodiments, the sequencing device comprises a processor programmed to provide a user interface and operate the modules of the sequencing device. In some embodiments, the sequencing device further comprises a display component for displaying information and providing a user interface. The system may also comprise one or more graphic boards for processing and outputting graphical information of a tissue image to the display component.

[0098] In some embodiments, a computer implemented method is used to provide an interface between a user and the sequencer firmware and hardware, for example, to perform sequencing run set up, select sequencing run options, and select and define a sample region of interest (ROI). The computer implemented methods may be used to control the different modules of the sequencing device and parallelization of sequencing across samples, and provide logging, error monitoring, data acquisition, management and transfer, and run progress monitoring.

[0099] In one embodiment, a computer implemented method is provided, the computer performing steps comprising: (a) locating a selected sample in the multi-well plate; (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning; (c) using the signal to segment an XY bounding box around the sample; (d)

imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent; (e) displaying the image produced in step (d); (f) providing an interface for a user to select a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample; (g) imaging the sample in the selected XY region of interest across the previously sampled Z extents; (h) calculating a sample volume of the region of interest and displaying the calculated sample volume of the region of interest to the user; (i) segmenting the image of the sample in the region of interest along Z extents; (j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and (k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.

[00100] In another embodiment, a computer implemented method is provided, the computer performing steps comprising: providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents available and the selected sequencing protocol; providing constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints. In some embodiments, the computer is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.

[00101] In another embodiment, a computer implemented method is provided, the computer performing steps comprising: performing a rapid confocal sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position; determining the sample top and bottom interface using a segmentation method; and setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.

[00102] The method can be implemented in digital electronic circuitry, or in computer software, firmware, or hardware. The disclosed and other embodiments can be implemented as one or more computer program products, i.e., one or more modules of computer program instructions encoded on a computer readable medium for execution by, or to control the operation of, a data processing apparatus. The computer readable medium can be a machine-readable storage device, a machine-readable storage substrate, a memory device, a composition of matter effecting a machine-readable propagated signal, or any combination thereof.

[00103] A computer program (also known as a program, software, software application, script, or code) can be written in any form of programming language, including compiled or interpreted languages, and it can be deployed in any form, including as a stand-alone program or as a module, component, subroutine, or other unit suitable for use in a computing environment. A computer program does not necessarily correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, sub programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

[00104] In a further aspect, the system for performing the computer implemented method, as described, may include a processor, a storage component (i.e., memory), a display component, and other components typically present in general purpose computers. The storage component stores information accessible by the processor, including instructions that may be executed by the processor and data that may be retrieved, manipulated or stored by the processor.

[00105] The storage component includes instructions. For example, the storage component may include instructions for providing a user interface for the sequencing device, operating the sequencing device, and processing in situ sequencing imaging data, as described herein. The computer processor is coupled to the storage component and configured to execute the instructions stored in the storage component in order to receive in situ sequencing imaging data and analyze the data according to one or more algorithms, as described herein. The display component displays information and provides a user interface.

[00106] The storage component may be of any type capable of storing information accessible by the processor, such as a hard-drive, memory card, ROM, RAM, DVD, CD-ROM, USB Flash drive, write-capable, and read-only memories. The processor may be any well-known processor, such as processors from Intel Corporation. Alternatively, the processor may be a dedicated controller such as an ASIC.

[00107] In certain embodiments, the in situ sequencing imaging data are uploaded and stored in a cloud data storage system. In some embodiments, the cloud data storage system is a public cloud storage system. In other embodiments, the cloud data storage system is a private cloud storage system. Cloud data storage may be used to store raw images, intermediate processed files, and final data products. Processing may begin with the upload of a dataset into cloud storage by a data acquisition system. Configuration parameters such as the encoding scheme, codebook, image acquisition parameters, and sample metadata can be input by a user using a data management web interface or generated automatically from a configuration file uploaded into cloud storage along with the sequencing data. Each set of configuration parameters is stored in a cloud database. In some cases, multiple processing runs using different configuration parameters can be applied to a single dataset to optimize processing parameters.

[00108] The instructions may be any set of instructions to be executed directly (such as machine code) or indirectly (such as scripts) by the processor. In that regard, the terms "instructions," "steps" and "programs" may be used interchangeably herein. The instructions may be stored in object code form for direct processing by the processor, or in any other computer language including scripts or collections of independent source code modules that are interpreted on demand or compiled in advance.

[00109] Data may be retrieved, stored or modified by the processor in accordance with the instructions. For instance, although the system is not limited by any particular data structure, the data may be stored in computer registers, in a relational database as a table having a plurality of different fields and records, XML documents, or flat files. The data may also be formatted in any computer-readable format such as, but not limited to, binary values, ASCII or Unicode. Moreover, the data may comprise any information sufficient to identify the relevant information, such as numbers, descriptive text, proprietary codes, pointers, references to data stored in other memories (including other network locations) or information which is used by a function to calculate the relevant data.

[00110] In certain embodiments, the processor and storage component may comprise multiple processors and storage components that may or may not be stored within the same physical housing. For example, some of the instructions and data may be stored on removable CD-ROM and others within a read-only computer chip. Some or all of the instructions and data may be stored in a location physically remote from, yet still accessible by, the processor. Similarly, the processor may comprise a collection of processors which may or may not operate in parallel.

[00111] In some embodiments, the method can be performed using a cloud computing system. In some embodiments, the image data files and programming for processing the imaging data can be

exported to a cloud computer, which runs the program, and returns an output to the user. The method may include optional compression of the imaging data before transfer to reduce data size and increase transfer speed. During the data acquisition process, acquired images are coupled with metadata files detailing the optical specifications, stage positions, and sequencing information; the optional compression of imaging data as a separate process from the imaging acquisition; and the optional offloading of data from the acquisition to a remote cloud storage medium, a networked attached storage system, or a separate large file system.

[00112] Components of systems for carrying out the presently disclosed methods are further described in the examples below.

In Situ Gene Sequencing

[00113] The sequencing device disclosed herein may be used for in situ gene sequencing of a target nucleic acid in a cell in an intact tissue. In situ sequencing may be performed by a method comprising: (a) contacting a fixed and permeabilized intact tissue with at least a pair of oligonucleotide primers under conditions to allow for specific hybridization, wherein the pair of primers comprise a first oligonucleotide and a second oligonucleotide; wherein each of the first oligonucleotide and the second oligonucleotide comprises a first complementarity region, a second complementarity region sequence, and a third complementarity region; wherein the second oligonucleotide further comprises a barcode sequence; wherein the first complementarity region of the first oligonucleotide is complementary to a first portion of the target nucleic acid, wherein the second complementarity region of the first oligonucleotide is complementary to the first complementarity region of the second oligonucleotide, wherein the third complementarity region of the first oligonucleotide is complementary to the third complementarity region of the second oligonucleotide, wherein the second complementarity region of the second oligonucleotide is complementary to a second portion of the target nucleic acid, wherein the first portion of the target nucleic acid is adjacent to the second portion of the target nucleic acid; (b) adding ligase to ligate the second oligonucleotide and generate a closed nucleic acid circle; (c) performing rolling circle amplification in the presence of a nucleic acid molecule, wherein the performing comprises using the second oligonucleotide as a template and the first oligonucleotide as a primer for a polymerase to form one or more amplicons; (d) embedding the one or more amplicons in the presence of hydrogel subunits to form one or more hydrogel-embedded amplicons; (e) contacting the one or more hydrogel-embedded amplicons having the barcode sequence with a set of sequencing primers under conditions to allow for ligation, wherein the set of sequencing primers comprises a third oligonucleotide configured to decode bases and a fourth oligonucleotide configured to convert decoded bases into a signal, wherein the ligation only occurs

when both the third oligonucleotide and the fourth oligonucleotide are complementary to adjacent sequences of the same amplicon; (f) reiterating step (e); and (g) imaging the one or more hydrogel-embedded amplicons using the sequencing device described herein to determine a gene sequence of the target nucleic acid in situ in the cell in the intact tissue.

[00114] In some embodiments, in situ sequencing is performed using Sequencing with Error-correction by Dynamic Annealing and Ligation (SEDAL) to determine a sequence of a target nucleic acid. The SEDAL method comprises contacting one or more hydrogel-embedded amplicons having the barcode sequence with a pair of primers under conditions to allow for ligation, wherein the pair of primers include a third oligonucleotide and a fourth oligonucleotide, wherein the ligation only occurs when both the third oligonucleotide and the fourth oligonucleotide ligate to the same amplicon. In some embodiments, SEDAL is used with STARmap. In such embodiments, the method herein includes operating at room temperature for best preservation of tissue morphology with low background noise and error reduction. In such other embodiments, the contacting the one or more hydrogel-embedded amplicons includes eliminating error accumulation as sequencing proceeds.

[00115] In some embodiments, the contacting the one or more hydrogel-embedded amplicons occurs two times or more, including, but not limited to, e.g., three times or more, four times or more, five times or more, six times or more, or seven times or more. In certain embodiments, the contacting the one or more hydrogel-embedded amplicons occurs four times or more for thin tissue specimens. In other embodiments, the contacting the one or more hydrogel-embedded amplicons occurs six times or more for thick tissue specimens. In some embodiments, one or more amplicons can be contacted by a pair of primers for 24 or more hours, 24 or less hours, 18 or less hours, 12 or less hours, 8 or less hours, 6 or less hours, 4 or less hours, 2 or less hours, 60 or less minutes, 45 or less minutes, 30 or less minutes, 25 or less minutes, 20 or less minutes, 15 or less minutes, 10 or less minutes, 5 or less minutes, or 2 or less minutes.

[00116] Specimens prepared using the subject methods may be analyzed by any of a number of different types of microscopy, for example, optical microscopy (e.g. bright field, oblique illumination, dark field, phase contrast, differential interference contrast, interference reflection, epifluorescence, confocal, etc., microscopy), laser microscopy, electron microscopy, and scanning probe microscopy. In some aspects, a non-transitory computer readable medium transforms raw images acquired through microscopy of multiple rounds of in situ sequencing first into decoded gene identities and spatial locations and then analyzes the per-cell composition of gene expression.

SEDAL Oligonucleotide Primers

[00117] In some embodiments, the methods disclosed include a third oligonucleotide and a fourth oligonucleotide. In certain aspects, the third oligonucleotide is configured to decode bases and the fourth oligonucleotide is configured to convert decoded bases into a signal. In some aspects, the signal is a fluorescent signal. In exemplary aspects, the contacting the one or more hydrogel-embedded amplicons having the barcode sequence with a pair of primers under conditions to allow for ligation involves each of the third oligonucleotide and the fourth oligonucleotide ligating to form a stable product for imaging only when a perfect match occurs. In certain aspects, the mismatch sensitivity of a ligase enzyme is used to determine the underlying sequence of the target nucleic acid molecule.

[00118] The term “perfectly matched”, when used in reference to a duplex means that the polynucleotide and/or oligonucleotide strands making up the duplex form a double stranded structure with one another such that every nucleotide in each strand undergoes Watson-Crick base pairing with a nucleotide in the other strand. The term “duplex” includes, but is not limited to, the pairing of nucleoside analogs, such as deoxyinosine, nucleosides with 2-aminopurine bases, peptide nucleic acids (PNAs), and the like, that may be employed. A “mismatch” in a duplex between two oligonucleotides means that a pair of nucleotides in the duplex fails to undergo Watson-Crick bonding.

[00119] In some embodiments, the method includes a plurality of third oligonucleotides, including, but not limited to, 5 or more third oligonucleotides, e.g., 8 or more, 10 or more, 12 or more, 15 or more, 18 or more, 20 or more, 25 or more, 30 or more, 35 or more that hybridize to target nucleotide sequences. In some embodiments, a method of the present disclosure includes a plurality of third oligonucleotides, including, but not limited to, 15 or more third oligonucleotides, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different first oligonucleotides that hybridize to 15 or more, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different target nucleotide sequences. In some embodiments, the methods include a plurality of fourth oligonucleotides, including, but not limited to, 5 or more fourth oligonucleotides, e.g., 8 or more, 10 or more, 12 or more, 15 or more, 18 or more, 20 or more, 25 or more, 30 or more, 35 or more. In some embodiments, a method of the present disclosure includes a plurality of fourth oligonucleotides including, but not limited to, 15 or more fourth oligonucleotides, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different first oligonucleotides that hybridize to 15 or more, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different target nucleotide sequences. A plurality of oligonucleotide pairs can be used in a reaction, where one or more pairs specifically bind to each

target nucleic acid. For example, two primer pairs can be used for one target nucleic acid in order to improve sensitivity and reduce variability. It is also of interest to detect a plurality of different target nucleic acids in a cell, e.g. detecting up to 2, up to 3, up to 4, up to 5, up to 6, up to 7, up to 8, up to 9, up to 10, up to 12, up to 15, up to 18, up to 20, up to 25, up to 30, up to 40 or more distinct target nucleic acids.

[00120] In certain embodiments, SEDAL involves a ligase with activity hindered by base mismatches, a third oligonucleotide, and a fourth oligonucleotide. The term “hindered” in this context refers to activity of a ligase that is reduced by approximately 20% or more, such as by 25% or more, such as by 50% or more, such as by 75% or more, such as by 90% or more, such as by 95% or more, such as by 99% or more, such as by 100%. In some embodiments, the third oligonucleotide has a length of 5-15 nucleotides, including, but not limited to, 5-13 nucleotides, 5-10 nucleotides, or 5-8 nucleotides. In some embodiments, the T_m of the third oligonucleotide is at room temperature (22-25°C). In some embodiments, the third oligonucleotide is degenerate, or partially thereof. In some embodiments, the fourth oligonucleotide has a length of 5-15 nucleotides, including, but not limited to, 5-13 nucleotides, 5-10 nucleotides, or 5-8 nucleotides. In some embodiments, the T_m of the fourth oligonucleotide is at room temperature (22-25°C). After each cycle of SEDAL corresponding to a base readout, the fourth oligonucleotides may be stripped, which eliminates error accumulation as sequencing proceeds. In such embodiments, the fourth oligonucleotides are stripped by formamide.

[00121] In some embodiments, SEDAL involves the washing of the third oligonucleotide and the fourth oligonucleotide to remove unbound oligonucleotides, thereafter revealing a fluorescent product for imaging. In certain exemplary embodiments, a detectable label can be used to detect one or more nucleotides and/or oligonucleotides described herein. In certain embodiments, a detectable label can be used to detect the one or more amplicons. Examples of detectable markers include various radioactive moieties, enzymes, prosthetic groups, fluorescent markers, luminescent markers, bioluminescent markers, metal particles, protein-protein binding pairs, protein-antibody binding pairs and the like. Examples of fluorescent proteins include, but are not limited to, yellow fluorescent protein (YFP), green fluorescence protein (GFP), cyan fluorescence protein (CFP), umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride, phycoerythrin and the like. Examples of bioluminescent markers include, but are not limited to, luciferase (e.g., bacterial, firefly, click beetle and the like), luciferin, aequorin and the like. Examples of enzyme systems having visually detectable signals include, but are not limited to, galactosidases, glucorimidases, phosphatases, peroxidases, cholinesterases and the like. Identifiable markers also include radioactive compounds such as ^{125}I , ^{35}S , ^{14}C , or ^3H . Identifiable markers are commercially available from a variety of sources.

[00122] Fluorescent labels and their attachment to nucleotides and/or oligonucleotides are described in many reviews, including Haugland, Handbook of Fluorescent Probes and Research Chemicals, Ninth Edition (Molecular Probes, Inc., Eugene, 2002); Keller and Manak, DNA Probes, 2nd Edition (Stockton Press, New York, 1993); Eckstein, editor, Oligonucleotides and Analogues: A Practical Approach (IRL Press, Oxford, 1991); and Wetmur, Critical Reviews in Biochemistry and Molecular Biology, 26:227-259 (1991). Particular methodologies applicable to the invention are disclosed in the following sample of references: U.S. Pat. Nos. 4,757,141, 5,151,507 and 5,091,519. In one aspect, one or more fluorescent dyes are used as labels for labeled target sequences, e.g., as disclosed by U.S. Pat. No. 5,188,934 (4,7-dichlorofluorescein dyes); U.S. Pat. No. 5,366,860 (spectrally resolvable rhodamine dyes); U.S. Pat. No. 5,847,162 (4,7-dichlororhodamine dyes); U.S. Pat. No. 4,318,846 (ether-substituted fluorescein dyes); U.S. Pat. No. 5,800,996 (energy transfer dyes); Lee et al.; U.S. Pat. No. 5,066,580 (xanthine dyes); U.S. Pat. No. 5,688,648 (energy transfer dyes); and the like. Labelling can also be carried out with quantum dots, as disclosed in the following patents and patent publications: U.S. Pat. Nos. 6,322,901, 6,576,291, 6,423,551, 6,251,303, 6,319,426, 6,426,513, 6,444,143, 5,990,479, 6,207,392, 2002/0045045 and 2003/0017264. As used herein, the term "fluorescent label" includes a signaling moiety that conveys information through the fluorescent absorption and/or emission properties of one or more molecules. Such fluorescent properties include fluorescence intensity, fluorescence lifetime, emission spectrum characteristics, energy transfer, and the like.

[00123] Commercially available fluorescent nucleotide analogues readily incorporated into nucleotide and/or oligonucleotide sequences include, but are not limited to, Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy5-dUTP (Amersham Biosciences, Piscataway, N.J.), fluorescein-12-dUTP, tetramethylrhodamine-6-dUTP, TEXAS RED™-5-dUTP, CASCADE BLUE™-7-dUTP, BODIPY TMFL-14-dUTP, BODIPY TMR-14-dUTP, BODIPY TMTR-14-dUTP, RHODAMINE GREEN™-5-dUTP, OREGON GREENR™ 488-5-dUTP, TEXAS RED™-12-dUTP, BODIPY™ 630/650-14-dUTP, BODIPY™ 650/665-14-dUTP, ALEXA FLUOR™ 488-5-dUTP, ALEXA FLUOR™ 532-5-dUTP, ALEXA FLUOR™ 568-5-dUTP, ALEXA FLUOR™ 594-5-dUTP, ALEXA FLUOR™ 546-14-dUTP, fluorescein-12-UTP, tetramethylrhodamine-6-UTP, TEXAS RED™-5-UTP, mCherry, CASCADE BLUE™-7-UTP, BODIPY™ FL-14-UTP, BODIPY TMR-14-UTP, BODIPY™ TR-14-UTP, RHODAMINE GREEN™-5-UTP, ALEXA FLUOR™ 488-5-UTP, ALEXA FLUOR™ 546-14-UTP (Molecular Probes, Inc. Eugene, Oreg.) and the like. Protocols are known in the art for custom synthesis of nucleotides having other fluorophores (See, Henegariu et al. (2000) Nature Biotechnol. 18:345).

- [00124] Other fluorophores available for post-synthetic attachment include, but are not limited to, ALEXA FLUOR™ 350, ALEXA FLUOR™ 532, ALEXA FLUOR™ 546, ALEXA FLUOR™ 568, ALEXA FLUOR™ 594, ALEXA FLUOR™ 647, BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethyl rhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, Oreg.), Cy2, Cy3.5, Cy5.5, Cy7 (Amersham Biosciences, Piscataway, N.J.) and the like. FRET tandem fluorophores may also be used, including, but not limited to, PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-Texas Red, APC-Cy7, PE-Alexa dyes (610, 647, 680), APC-Alexa dyes and the like.
- [00125] Metallic silver or gold particles may be used to enhance signal from fluorescently labeled nucleotide and/or oligonucleotide sequences (Lakowicz et al. (2003) Bio Techniques 34:62).
- [00126] Biotin, or a derivative thereof, may also be used as a label on a nucleotide and/or an oligonucleotide sequence, and subsequently bound by a detectably labeled avidin/streptavidin derivative (e.g. phycoerythrin-conjugated streptavidin), or a detectably labeled anti-biotin antibody. Digoxigenin may be incorporated as a label and subsequently bound by a detectably labeled anti-digoxigenin antibody (e.g. fluoresceinated anti-digoxigenin). An aminoallyl-dUTP residue may be incorporated into an oligonucleotide sequence and subsequently coupled to an N-hydroxy succinimide (NHS) derivatized fluorescent dye. In general, any member of a conjugate pair may be incorporated into a detection oligonucleotide provided that a detectably labeled conjugate partner can be bound to permit detection. As used herein, the term antibody refers to an antibody molecule of any class, or any sub-fragment thereof, such as an Fab.
- [00127] Other suitable labels for an oligonucleotide sequence may include fluorescein (FAM), digoxigenin, dinitrophenol (DNP), dansyl, biotin, bromodeoxyuridine (BrdU), hexahistidine (6×His), phosphor-amino acids (e.g. P-tyr, P-ser, P-thr) and the like. In one embodiment the following hapten/antibody pairs are used for detection, in which each of the antibodies is derivatized with a detectable label: biotin/α-biotin, digoxigenin/α-digoxigenin, dinitrophenol (DNP)/α-DNP, 5-Carboxyfluorescein (FAM)/α-FAM.
- [00128] In certain exemplary embodiments, a nucleotide and/or an oligonucleotide sequence can be indirectly labeled, especially with a hapten that is then bound by a capture agent, e.g., as disclosed in U.S. Pat. Nos. 5,344,757, 5,702,888, 5,354,657, 5,198,537 and 4,849,336, PCT publication WO 91/17160 and the like. Many different hapten-capture agent pairs are available for use. Exemplary haptens include, but are not limited to, biotin, des-biotin and other derivatives, dinitrophenol, dansyl,

fluorescein, CY5, digoxigenin and the like. For biotin, a capture agent may be avidin, streptavidin, or antibodies. Antibodies may be used as capture agents for the other haptens (many dye-antibody pairs being commercially available, e.g., Molecular Probes, Eugene, Oreg.).

[00129] In some embodiments, in situ sequencing is performed using sequencing by competitive annealing and ligation (SCAL) to determine a sequence of a target nucleic acid, the method comprising performing one or more sequencing cycles, each cycle comprising: (a) contacting the target nucleic acid with a read oligonucleotide and a set of fluorescently labeled decoding probes, wherein the read oligonucleotide comprises a first complementarity region that is complementary to a reading sequence on the target nucleic acid, and wherein each decoding probe comprises a second complementarity region that is complementary to a probe binding site on the target nucleic acid; (b) ligating the read oligonucleotide to one of the decoding probes of the set of fluorescently labeled decoding probes to generate a fluorescent ligation product, wherein the ligation only occurs when the read oligonucleotide and the decoding probe bind to adjacent sequences on the target nucleic acid and both the read oligonucleotide and the decoding probe have sequences that are exactly complementary to the sequence of the target nucleic acid; (c) removing unligated probes; (d) imaging the fluorescent ligation product to detect the fluorescent label of the decoding probe that ligated to the read oligonucleotide, wherein the fluorescent label identifies a nucleotide of the sequence of the target nucleic acid; and (e) removing the fluorescent ligation product from the target nucleic acid by binding a competitor oligonucleotide to the target nucleic acid, wherein the competitor oligonucleotide comprises a third complementarity region comprising a sequence that is complementary to the reading sequence on the target nucleic acid, wherein the fluorescent ligation product dissociates from the target nucleic acid.

[00130] In exemplary aspects, the ligation involves each of the read oligonucleotide and a fluorescently labeled decoding probe ligating to form a stable product for imaging only when a perfect match occurs. In certain aspects, the mismatch sensitivity of a ligase enzyme is used to determine the underlying sequence of the target nucleic acid molecule. Inclusion of a polyethylene glycol (PEG) polymer in the sequencing ligation mixture substantially accelerates signal addition onto target nucleic acids. Exemplary PEG polymers have molecular weights ranging from 300 g/mol to 10,000,000 g/mol. In some embodiments, a PEG 6000 polymer is present during ligation of the read oligonucleotide and a fluorescently labeled decoding probe.

[00131] In certain embodiments, the set of fluorescently labeled decoding probes comprises: a first probe encoding a guanine, wherein the first probe comprises a first fluorescent label, a second probe encoding an adenine, wherein the second probe comprises a second fluorescent label, a third probe

encoding a cytosine, wherein the third probe comprises a third fluorescent label, and a fourth probe encoding a thymine, wherein the fourth probe comprises a fourth fluorescent label.

[00132] In certain embodiments, each fluorescently labeled decoding probe encodes 1 to 3 bases adjacent to a ligation junction where the read oligonucleotide is ligated to the fluorescently labeled decoding probe, wherein fluorescently labeled decoding probes encoding different sequences of bases comprise different fluorescent labels.

[00133] In certain embodiments, the sequences of the fluorescently labeled decoding probes for a current cycle of sequencing are optimized to minimize cross-hybridization with the fluorescently labeled decoding probes for other sequencing cycles.

[00134] In certain embodiments, the read oligonucleotide ranges in length from 8 to 11 nucleotides, including any length within this range such as 8, 9, 10, or 11 nucleotides in length. In some embodiments, the read oligonucleotide has a melting temperature ranging from 17 °C to 20 °C, including any melting temperature within this range such as 17 °C, 18 °C, 19 °C, or 20 °C.

[00135] In certain embodiments, the competitor oligonucleotide further comprises a fourth complementarity region comprising a sequence that is complementary to at least a portion of the probe binding site. In some embodiments, the fourth complementarity region of the competitor oligonucleotide comprises a sequence that is fully complementary to the entire probe binding site on the target nucleic acid. In certain embodiments, the competitor oligonucleotide further comprises a fifth complementarity region comprising a sequence that is complementary to a competitor-specific complementary site adjacent to the reading sequence on the target nucleic acid. In some embodiments, the competitor-specific complementary site ranges in length from 2 nucleotides to 16 nucleotides, including any length within this range such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 nucleotides. In certain embodiments, for sequencing cycles following an initial sequencing cycle, the competitor oligonucleotide used in a previous cycle of sequencing is present during one or more subsequent cycles of sequencing.

[00136] In certain embodiments, the read oligonucleotide further comprises a competitor-specific complementary sequence. In some embodiments, the competitor-specific complementary sequence of the read oligonucleotide ranges in length from 2 nucleotides to 16 nucleotides, including any length within this range such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 nucleotides. In certain embodiments, the sequence of the read oligonucleotide for a current cycle of sequencing is optimized to minimize cross-hybridization with read oligonucleotides for other sequencing cycles.

[00137] In certain embodiments, multiple read oligonucleotides, sets of fluorescently labeled decoding probes, and competitor oligonucleotides having specificity for different target nucleic acids are used to sequence a plurality of different target nucleic acids simultaneously or sequentially.

[00138] In certain embodiments, the competitor oligonucleotides remove ligation products from a previous round of sequencing from different target nucleic acids than target nucleic acids currently undergoing steps (a) or (b) of a sequencing cycle. In certain embodiments, the competitor oligonucleotides remove ligation products from a previous round of sequencing from the same target nucleic acids currently undergoing steps (a) or (b) of a sequencing cycle. In certain embodiments, the competitor oligonucleotide is a round-specific competitor oligonucleotide comprising a fourth complementarity region comprising a sequence that is complementary to the reading sequence for the next cycle of sequencing.

[00139] The sequencing reads may be in a 5' to 3' forward direction or a 3' to 5' reverse direction. For sequencing reads in the forward direction, each fluorescently labeled decoding probe has a fluorophore modification at the 5' end and each read oligonucleotide has a phosphate at the 5' end. For sequencing reads in the reverse direction, each fluorescently labeled decoding probe has a phosphate at the 5' end and a fluorophore modification at the 3' end.

[00140] In certain embodiments, the sequencing is performed with sequential encoding. In some embodiments, each read oligonucleotide comprises a unique sequential orthogonal readout sequence and a unique adjacent competitor-specific complementary sequence for each cycle of sequencing. In some embodiments, the unique sequential orthogonal readout sequence ranges in length from 8 nucleotides to 11 nucleotides, including any length within this range such as 8, 9, 10, or 11 nucleotides. In some embodiments, the unique adjacent competitor-specific complementary sequence ranges in length from 2 nucleotides to 16 nucleotides, including any length within this range such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 nucleotides. In some embodiments, each competitor oligonucleotide comprises a sequence that is complementary to the unique sequential orthogonal readout sequence and the unique adjacent competitor-specific complementary sequence of the read oligonucleotide and at least a portion of the sequence of the fluorescently labeled decoding probe for each cycle of sequencing. In some embodiments, the sequence of the competitor oligonucleotide has partial complementarity or full complementarity to the sequence of the fluorescently labeled decoding probe.

[00141] In certain embodiments, sequencing is performed with combinatorial encoding. In some embodiments, multiple read oligonucleotides are used for sequencing, wherein each read oligonucleotide comprises a first complementarity region comprising a combinatorial readout sequence that is complementary to a reading sequence at a separate combinatorial read position on the target nucleic acid, wherein the reading sequence at each separate position on the target nucleic acid is adjacent to a probe binding site. In some embodiments, each read oligonucleotide further comprises a competitor-specific complementary sequence adjacent to the reading sequence. In

some embodiments, the competitor-specific complementary sequence is not complementary to the fluorescently labeled decoding probe. In some embodiments, the reading sequence ranges in length from 8 nucleotides to 11 nucleotides, including any length within this range such as 8, 9, 10, or 11 nucleotides. In some embodiments, the competitor-specific complementary sequence ranges in length from 2 nucleotides to 16 nucleotides, including any length within this range such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 nucleotides. In some embodiments, for each separate combinatorial read position, the competitor oligonucleotide comprises a sequence that is complementary to the combinatorial readout sequence and the competitor-specific complementary sequence of the read oligonucleotide and at least a portion of the sequence of the fluorescently labeled decoding probe for each cycle of sequencing. In some embodiments, the combinatorial encoding uses a hamming code.

[00142] The sequencing methods described herein can be used for situ gene sequencing of a target nucleic acid in a cell in an intact tissue. In some embodiments, the method of in situ gene sequencing of a target nucleic acid in a cell in an intact tissue comprises: (a) contacting a fixed and permeabilized intact tissue with at least a pair of oligonucleotide primers under conditions to allow for specific hybridization, wherein the pair of primers comprise a first oligonucleotide and a second oligonucleotide; wherein each of the first oligonucleotide and the second oligonucleotide comprises a first complementarity region, a second complementarity region sequence, and a third complementarity region; wherein the second oligonucleotide further comprises a barcode sequence; wherein the first complementarity region of the first oligonucleotide is complementary to a first portion of the target nucleic acid, wherein the second complementarity region of the first oligonucleotide is complementary to the first complementarity region of the second oligonucleotide, wherein the third complementarity region of the first oligonucleotide is complementary to the third complementarity region of the second oligonucleotide, wherein the second complementarity region of the second oligonucleotide is complementary to a second portion of the target nucleic acid, wherein the first portion of the target nucleic is adjacent to the second portion of the target nucleic acid; (b) adding ligase to ligate the second oligonucleotide and generate a closed nucleic acid circle; (c) performing rolling circle amplification in the presence of a nucleic acid molecule, wherein the performing comprises using the second oligonucleotide as a template and the first oligonucleotide as a primer for a polymerase to form one or more amplicons; (d) embedding the one or more amplicons in the presence of hydrogel subunits to form one or more hydrogel-embedded amplicons; (e) sequencing the one or more amplicons according to a method described herein. In certain embodiments, the sequencing is performed with sequential encoding. In other embodiments, the sequencing is performed with combinatorial encoding.

[00143] In some embodiments, the contacting the one or more hydrogel-embedded amplicons occurs two times or more, including, but not limited to, e.g., three times or more, four times or more, five times or more, six times or more, or seven times or more, eight times or more, nine times or more, ten times or more, eleven times or more, or twelve times or more. In certain embodiments, the contacting the one or more hydrogel-embedded amplicons occurs four times or more for thin tissue specimens. In other embodiments, the contacting the one or more hydrogel-embedded amplicons occurs six times or more for thick tissue specimens. In some embodiments, one or more amplicons can be contacted by a pair of primers for 24 or more hours, 24 or less hours, 18 or less hours, 12 or less hours, 8 or less hours, 6 or less hours, 4 or less hours, 2 or less hours, 60 or less minutes, 45 or less minutes, 30 or less minutes, 25 or less minutes, 20 or less minutes, 15 or less minutes, 10 or less minutes, 5 or less minutes, or 2 or less minutes. In some embodiments, 12 or more cycles of sequencing are performed, including 13 or more cycles, 14 or more cycles, 15 or more cycles, 16 or more cycles, 17 or more cycles, or 18 or more cycles of sequencing. In some embodiments, the methods are performed at room temperature for preservation of tissue morphology with low background noise and error reduction. In some embodiments, the contacting the one or more hydrogel-embedded amplicons includes eliminating error accumulation as sequencing proceeds.

[00144] Specimens prepared using the subject methods may be analyzed by any of a number of different types of microscopy, for example, optical microscopy (e.g. bright field, oblique illumination, dark field, phase contrast, differential interference contrast, interference reflection, epifluorescence, confocal, etc., microscopy), laser microscopy, electron microscopy, and scanning probe microscopy. In some aspects, a non-transitory computer readable medium transforms raw images acquired through microscopy of multiple rounds of in situ sequencing first into decoded gene identities and spatial locations and then analyzes the per-cell composition of gene expression.

[00145] The term “perfectly matched”, when used in reference to a duplex means that the polynucleotide and/or oligonucleotide strands making up the duplex form a double stranded structure with one another such that every nucleotide in each strand undergoes Watson-Crick base pairing with a nucleotide in the other strand. The term “duplex” includes, but is not limited to, the pairing of nucleoside analogs, such as deoxyinosine, nucleosides with 2-aminopurine bases, peptide nucleic acids (PNAs), and the like, that may be employed. A “mismatch” in a duplex between two oligonucleotides means that a pair of nucleotides in the duplex fails to undergo Watson-Crick bonding.

[00146] In some embodiments, the method includes a plurality of read oligonucleotides, including, but not limited to, 5 or more read oligonucleotides, e.g., 8 or more, 10 or more, 12 or more, 15 or more, 18 or more, 20 or more, 25 or more, 30 or more, 35 or more that hybridize to target nucleotide

sequences. In some embodiments, a method of the present disclosure includes a plurality of read oligonucleotides, including, but not limited to, 15 or more read oligonucleotides, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different read oligonucleotides that hybridize to 15 or more, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different target nucleotide sequences.

[00147] In some embodiments, the methods include a plurality of fluorescently labeled decoding probes, including, but not limited to, 4 or more fluorescently labeled decoding probes, e.g., 8 or more, 10 or more, 12 or more, 16 or more, 18 or more, 20 or more, 25 or more, 30 or more, 35 or more. In some embodiments, a method of the present disclosure includes a plurality of fluorescently labeled decoding probes including, but not limited to, 15 or more fluorescently labeled decoding probes, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different fluorescently labeled decoding probes that hybridize to 15 or more, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different target nucleotide sequences.

[00148] A plurality of pairs of oligonucleotide primers can be used in a reaction, where one or more pairs specifically bind to each target nucleic acid. For example, two primer pairs can be used for one target nucleic acid in order to improve sensitivity and reduce variability. It is also of interest to detect a plurality of different target nucleic acids in a cell, e.g. detecting up to 2, up to 3, up to 4, up to 5, up to 6, up to 7, up to 8, up to 9, up to 10, up to 12, up to 15, up to 18, up to 20, up to 25, up to 30, up to 40 or more distinct target nucleic acids.

[00149] In certain embodiments, sequencing is performed with a ligase with activity hindered by base mismatches, a read oligonucleotide, and a fluorescently labeled decoding probe. The term "hindered" in this context refers to activity of a ligase that is reduced by approximately 20% or more, such as by 25% or more, such as by 50% or more, such as by 75% or more, such as by 90% or more, such as by 95% or more, such as by 99% or more, such as by 100%. In some embodiments, the third oligonucleotide has a length of 5-15 nucleotides, including, but not limited to, 5-13 nucleotides, 5-10 nucleotides, or 5-8 nucleotides. In some embodiments, the T_m of the third oligonucleotide is at room temperature (22-25°C). In some embodiments, the read oligonucleotide is degenerate, or partially thereof. In some embodiments, the fluorescently labeled decoding probe oligonucleotide has a length of 5-15 nucleotides, including, but not limited to, 5-13 nucleotides, 5-10 nucleotides, or 5-8 nucleotides. In some embodiments, the T_m of the fourth oligonucleotide is at room temperature (22°-25°C). After each cycle of sequencing corresponding to a base readout, the fluorescent ligation product is removed from the target nucleic acid by binding a competitor oligonucleotide to the target nucleic acid, wherein the competitor oligonucleotide comprises a third complementarity region comprising a sequence that is complementary to the reading sequence on

the target nucleic acid, wherein the fluorescent ligation product dissociates from the target nucleic acid.

[00150] In some embodiments, sequencing involves washing to remove unbound oligonucleotides and unligated probes, thereafter revealing a fluorescent product for imaging. In certain exemplary embodiments, a detectable fluorescent label is used to detect one or more nucleotides and/or oligonucleotides described herein. In certain embodiments, a detectable fluorescent label such as a fluorescent protein, fluorescent dye, or fluorescent quantum dot is used to label probes.

[00151] Fluorescent labels and their attachment to nucleotides and/or oligonucleotides are described in many reviews, including Haugland, Handbook of Fluorescent Probes and Research Chemicals, Ninth Edition (Molecular Probes, Inc., Eugene, 2002); Keller and Manak, DNA Probes, 2nd Edition (Stockton Press, New York, 1993); Eckstein, editor, Oligonucleotides and Analogues: A Practical Approach (IRL Press, Oxford, 1991); and Wetmur, Critical Reviews in Biochemistry and Molecular Biology, 26:227-259 (1991). Particular methodologies applicable to the invention are disclosed in the following sample of references: U.S. Pat. Nos. 4,757,141, 5,151,507 and 5,091,519. In one aspect, one or more fluorescent dyes are used as labels for labeled target sequences, e.g., as disclosed by U.S. Pat. No. 5,188,934 (4,7-dichlorofluorescein dyes); U.S. Pat. No. 5,366,860 (spectrally resolvable rhodamine dyes); U.S. Pat. No. 5,847,162 (4,7-dichlororhodamine dyes); U.S. Pat. No. 4,318,846 (ether-substituted fluorescein dyes); U.S. Pat. No. 5,800,996 (energy transfer dyes); Lee et al.; U.S. Pat. No. 5,066,580 (xanthine dyes); U.S. Pat. No. 5,688,648 (energy transfer dyes); and the like. Labelling can also be carried out with quantum dots, as disclosed in the following patents and patent publications: U.S. Pat. Nos. 6,322,901, 6,576,291, 6,423,551, 6,251,303, 6,319,426, 6,426,513, 6,444,143, 5,990,479, 6,207,392, 2002/0045045 and 2003/0017264. As used herein, the term "fluorescent label" includes a signaling moiety that conveys information through the fluorescent absorption and/or emission properties of one or more molecules. Such fluorescent properties include fluorescence intensity, fluorescence lifetime, emission spectrum characteristics, energy transfer, and the like.

[00152] Commercially available fluorescent nucleotide analogues readily incorporated into nucleotide and/or oligonucleotide sequences include, but are not limited to, Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy5-dUTP (Amersham Biosciences, Piscataway, N.J.), fluorescein-12-dUTP, tetramethylrhodamine-6-dUTP, TEXAS RED™-5-dUTP, CASCADE BLUE™-7-dUTP, BODIPY TMFL-14-dUTP, BODIPY TMR-14-dUTP, BODIPY TMTR-14-dUTP, RHODAMINE GREEN™-5-dUTP, OREGON GREENR™ 488-5-dUTP, TEXAS RED™-12-dUTP, BODIPY™ 630/650-14-dUTP, BODIPY™ 650/665-14-dUTP, ALEXA FLUOR™ 488-5-dUTP, ALEXA FLUOR™ 532-5-dUTP, ALEXA FLUOR™ 568-5-dUTP, ALEXA FLUOR™ 594-5-dUTP, ALEXA FLUOR™ 546-14-dUTP,

fluorescein-12-UTP, tetramethylrhodamine-6-UTP, TEXAS RED™-5-UTP, mCherry, CASCADE BLUE™-7-UTP, BODIPY™ FL-14-UTP, BODIPY TMR-14-UTP, BODIPY™ TR-14-UTP, RHODAMINE GREEN™-5-UTP, ALEXA FLUOR™ 488-5-UTP, LEXA FLUOR™ 546-14-UTP (Molecular Probes, Inc. Eugene, Oreg.) and the like. Protocols are known in the art for custom synthesis of nucleotides having other fluorophores (See, Henegariu et al. (2000) *Nature Biotechnol.* 18:345).

[00153] Other fluorophores available for post-synthetic attachment include, but are not limited to, ALEXA FLUOR™ 350, ALEXA FLUOR™ 532, ALEXA FLUOR™ 546, ALEXA FLUOR™ 568, ALEXA FLUOR™ 594, ALEXA FLUOR™ 647, BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethyl rhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, Oreg.), Cy2, Cy3.5, Cy5.5, Cy7 (Amersham Biosciences, Piscataway, N.J.) and the like. FRET tandem fluorophores may also be used, including, but not limited to, PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-Texas Red, APC-Cy7, PE-Alexa dyes (610, 647, 680), APC-Alexa dyes and the like.

[00154] Examples of fluorescent proteins include, but are not limited to, green fluorescent protein, superfolder green fluorescent protein, enhanced green fluorescent protein, Dronpa (a photoswitchable green fluorescent protein), yellow-green fluorescent protein, yellow fluorescent protein, red fluorescent protein, orange fluorescent protein, blue fluorescent protein, cyan fluorescent protein, violet fluorescent protein, mApple, mNectarine, mNeptune, mCherry, mStrawberry, mPlum, mRaspberry, mCrimson3, mCarmine, mCardinal, mScarlet, mRuby2, FusionRed, mNeonGreen, TagRFP675, and mRFP1. and the like.

[00155] Metallic silver or gold particles may be used to enhance signal from fluorescently labeled nucleotide and/or oligonucleotide sequences (Lakowicz et al. (2003) *Bio Techniques* 34:62).

[00156] Biotin, or a derivative thereof, may also be used as a label on a nucleotide and/or an oligonucleotide sequence, and subsequently bound by a detectably labeled avidin/streptavidin derivative (e.g. phycoerythrin-conjugated streptavidin), or a detectably labeled anti-biotin antibody. Digoxigenin may be incorporated as a label and subsequently bound by a detectably labeled anti-digoxigenin antibody (e.g. fluoresceinated anti-digoxigenin). An aminoallyl-dUTP residue may be incorporated into an oligonucleotide sequence and subsequently coupled to an N-hydroxy succinimide (NHS) derivatized fluorescent dye. In general, any member of a conjugate pair may be incorporated into a detection oligonucleotide provided that a detectably labeled conjugate partner

can be bound to permit detection. As used herein, the term antibody refers to an antibody molecule of any class, or any sub-fragment thereof, such as an Fab.

[00157] Other suitable labels for an oligonucleotide sequence may include fluorescein (FAM), digoxigenin, dinitrophenol (DNP), dansyl, biotin, bromodeoxyuridine (BrdU), hexahistidine (6×His), phosphor-amino acids (e.g. P-tyr, P-ser, P-thr) and the like. In one embodiment the following hapten/antibody pairs are used for detection, in which each of the antibodies is derivatized with a detectable label: biotin/α-biotin, digoxigenin/α-digoxigenin, dinitrophenol (DNP)/α-DNP, 5-Carboxyfluorescein (FAM)/α-FAM.

[00158] In certain exemplary embodiments, a nucleotide and/or an oligonucleotide sequence can be indirectly labeled, especially with a hapten that is then bound by a capture agent, e.g., as disclosed in U.S. Pat. Nos. 5,344,757, 5,702,888, 5,354,657, 5,198,537 and 4,849,336, PCT publication WO 91/17160 and the like. Many different hapten-capture agent pairs are available for use. Exemplary haptens include, but are not limited to, biotin, des-biotin and other derivatives, dinitrophenol, dansyl, fluorescein, CY5, digoxigenin and the like. For biotin, a capture agent may be avidin, streptavidin, or antibodies. Antibodies may be used as capture agents for the other haptens (many dye-antibody pairs being commercially available, e.g., Molecular Probes, Eugene, Oreg.).

[00159] In some embodiments, an antioxidant compound is included in the washing and imaging buffers (i.e., "anti-fade buffers") to reduce photobleaching during fluorescence imaging. Exemplary antioxidants include, without limitation, Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid) and Trolox-quinone, propyl-gallate, tertiary butylhydroquinone, butylated hydroxyanisole, butylated hydroxytoluene, glutathione, ascorbic acid, and tocopherols. Such antioxidants have an antifade effect on fluorophores. That is, the antioxidant reduces photobleaching during tiling, greatly enhances the signal-to-noise ratio (SNR) of sensitive fluorophores, and enables higher SNR imaging of thicker samples. For a fixed exposure time, including an antioxidant increases the SNR by increasing the concentration of the non-bleached fluorophore during exposure to light. Including an antioxidant also removes the diminishing returns of longer exposure times (caused by the limited fluorophore lifetime before photobleaching), providing for increased SNR by allowing increased exposure times.

[00160] An exemplary sequencing cycle optionally begins with a brief sample wash, before proceeding to the first signal addition. Depending on whether sequential or combinatorial encoding is being used for a particular round, the corresponding set of read oligonucleotides, fluorescently labeled decoding probes, and their round-specific competitors are added and ligated. In combinatorial encodings, the read oligonucleotide for a given position *x* is added, plus a set of fluorescently labeled dibase-encoding oligonucleotides, plus a competitor oligonucleotide for the

previous position that was labeled (unless it is the first round of labeling, in which case competitor oligonucleotide is omitted). In sequential encodings, the read oligonucleotide for a given round x , a 4-channel fluorophore mixture, and a round $x-1$ competitor oligonucleotide are added, except if it is the first round of labeling. The presence of PEG in the sequencing ligation mixture substantially accelerates the signal addition onto the target. Following incubation of the sample in imaging buffer, the sample is imaged, and briefly rinsed before proceeding to the next sequencing cycle.

[00161] In addition, fluorophore cleavage from probes or probe stripping can be used to eliminate signal carryover from one round to the next when multiple sequencing cycles are used. For example, fluorophores can be stripped off with formamide. Alternatively, thiol-linked dyes can be used having a disulfide linkage between the fluorophore and an oligonucleotide probe, which enables cleavage of the fluorophore from the oligonucleotide probe in a reducing environment. Exemplary disulfide reducing agents, which can be used for cleaving disulfide bonds include, without limitation, tris(2-carboxyethyl)phosphine (TCEP), dithiothreitol (DTT), and β -mercaptoethanol (BME). Following fluorescence imaging during a sequencing round, a stripping agent and/or a reducing agent is added, and subsequent washing steps remove the diffusive fluorescent signal before performing another round of sequencing.

[00162] The methods disclosed herein also provide for a method of screening a candidate agent to determine whether the candidate agent modulates gene expression of a nucleic acid in a cell in an intact tissue by performing a method described herein to determine the gene sequence of a target nucleic acid in the cell in the intact tissue, and detecting the level of gene expression of the target nucleic acid, wherein an alteration in the level of expression of the target nucleic acid in the presence of the candidate agent relative to the level of expression of the target nucleic acid in the absence of the candidate agent indicates that the candidate agent modulates gene expression of the nucleic acid in the cell in the intact tissue.

[00163] In certain aspects, the methods disclosed herein provide for a faster processing time, higher multiplexity, higher efficiency, higher sensitivity, lower error rate, and more spatially resolved cell types, as compared to existing gene expression analysis tools. The methods provide improved sequencing-by-ligation techniques (SCAL and SEDAL2) for in situ sequencing with error reduction. In some other aspects, the methods disclosed herein include spatially sequencing (e.g. reagents, chips or services) for biomedical research and clinical diagnostics (e.g. cancer, bacterial infection, viral infection, etc.) with single-cell and/or single-molecule sensitivity.

Specific Amplification of Nucleic Acids via Intramolecular Ligation (SNAIL)

[00164] An efficient approach for generating cDNA libraries from cellular RNAs in situ may be utilized, which is referred to herein as SNAIL, for Specific Amplification of Nucleic Acids via Intramolecular Ligation. In certain embodiments, the method includes contacting a fixed and permeabilized intact tissue with at least a pair of oligonucleotide primers under conditions to allow for specific hybridization, wherein the pair of primers includes a first oligonucleotide and a second oligonucleotide.

[00165] More generally, the nucleic acid present in a cell of interest in a tissue serves as a scaffold for an assembly of a complex that includes a pair of primers, referred to herein as a first oligonucleotide and a second oligonucleotide. In some embodiments, the contacting the fixed and permeabilized intact tissue includes hybridizing the pair of primers to the same target nucleic acid. In some embodiments, the target nucleic acid is RNA. In such embodiments, the target nucleic acid may be mRNA. In other embodiments, the target nucleic acid is DNA.

[00166] As used herein, the terms “hybridize” and “hybridization” refer to the formation of complexes between nucleotide sequences which are sufficiently complementary to form complexes via Watson-Crick base pairing. Where a primer “hybridizes” with target (template), such complexes (or hybrids) are sufficiently stable to serve the priming function required by, e.g., the DNA polymerase to initiate DNA synthesis. It will be appreciated that the hybridizing sequences need not have perfect complementarity to provide stable hybrids. In many situations, stable hybrids will form where fewer than about 10% of the bases are mismatches, ignoring loops of four or more nucleotides. Accordingly, as used herein the term “complementary” refers to an oligonucleotide that forms a stable duplex with its “complement” under assay conditions, generally where there is about 90% or greater homology.

SNAIL Oligonucleotide Primers

[00167] In the subject methods, the SNAIL oligonucleotide primers include at least a first oligonucleotide and a second oligonucleotide; wherein each of the first oligonucleotide and the second oligonucleotide includes a first complementarity region, a second complementarity region, and a third complementarity region; wherein the second oligonucleotide further includes a barcode sequence; wherein the first complementarity region of the first oligonucleotide is complementary to a first portion of the target nucleic acid, wherein the second complementarity region of the first oligonucleotide is complementary to the first complementarity region of the second oligonucleotide, wherein the third complementarity region of the first oligonucleotide is complementary to the third complementarity region of the second oligonucleotide, wherein the second complementarity region of

the second oligonucleotide is complementary to a second portion of the target nucleic acid, and wherein the first complementarity region of the first oligonucleotide is adjacent to the second complementarity region of the second oligonucleotide. In an alternative embodiment, the second oligonucleotide is a closed circular molecule, and a ligation step is omitted.

[00168] The present disclosure provides methods where the contacting a fixed and permeabilized tissue includes hybridizing a plurality of oligonucleotide primers having specificity for different target nucleic acids. In some embodiments, the methods include a plurality of first oligonucleotides, including, but not limited to, 5 or more first oligonucleotides, e.g., 8 or more, 10 or more, 12 or more, 15 or more, 18 or more, 20 or more, 25 or more, 30 or more, 35 or more that hybridize to target nucleotide sequences. In some embodiments, a method of the present disclosure includes a plurality of first oligonucleotides, including, but not limited to, 15 or more first oligonucleotides, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different first oligonucleotides that hybridize to 15 or more, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different target nucleotide sequences. In some embodiments, the methods include a plurality of second oligonucleotides, including, but not limited to, 5 or more second oligonucleotides, e.g., 8 or more, 10 or more, 12 or more, 15 or more, 18 or more, 20 or more, 25 or more, 30 or more, 35 or more. In some embodiments, a method of the present disclosure includes a plurality of second oligonucleotides including, but not limited to, 15 or more second oligonucleotides, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different first oligonucleotides that hybridize to 15 or more, e.g., 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, and up to 80 different target nucleotide sequences. A plurality of oligonucleotide pairs can be used in a reaction, where one or more pairs specifically bind to each target nucleic acid. For example, two primer pairs can be used for one target nucleic acid in order to improve sensitivity and reduce variability. It is also of interest to detect a plurality of different target nucleic acids in a cell, e.g. detecting up to 2, up to 3, up to 4, up to 5, up to 6, up to 7, up to 8, up to 9, up to 10, up to 12, up to 15, up to 18, up to 20, up to 25, up to 30, up to 40 or more distinct target nucleic acids. The primers are typically denatured prior to use, typically by heating to a temperature of at least about 50°C, at least about 60°C, at least about 70°C, at least about 80°C, and up to about 99°C, up to about 95°C, up to about 90°C.

[00169] In some embodiments, the primers are denatured by heating before contacting the sample. In certain aspects, the melting temperature (T_m) of oligonucleotides is selected to minimize ligation in solution. The “melting temperature” or “ T_m ” of a nucleic acid is defined as the temperature at which half of the helical structure of the nucleic acid is lost due to heating or other dissociation of the hydrogen bonding between base pairs, for example, by acid or alkali treatment, or the like. The T_m

of a nucleic acid molecule depends on its length and on its base composition. Nucleic acid molecules rich in GC base pairs have a higher T_m than those having an abundance of AT base pairs. Separated complementary strands of nucleic acid spontaneously reassociate or anneal to form duplex nucleic acid when the temperature is lowered below the T_m . The highest rate of nucleic acid hybridization occurs approximately 25 degrees C below the T_m . The T_m may be estimated using the following relationship: $T_m = 69.3 + 0.41(\text{GC})\%$ (Marmur et al. (1962) J. Mol. Biol. 5:109-118).

[00170] In certain embodiments, the plurality of second oligonucleotides includes a padlock probe. In some embodiments, the probe includes a detectable label that can be measured and quantitated. The terms "label" and "detectable label" refer to a molecule capable of detection, including, but not limited to, radioactive isotopes, fluorescers, chemiluminescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, chromophores, dyes, metal ions, metal sols, ligands (e.g., biotin or haptens) and the like. The term "fluorescer" refers to a substance or a portion thereof that is capable of exhibiting fluorescence in the detectable range. Particular examples of labels that may be used with the invention include, but are not limited to phycoerythrin, Alexa dyes, fluorescein, YPet, CyPet, Cascade blue, allophycocyanin, Cy3, Cy5, Cy7, rhodamine, dansyl, umbelliferone, Texas red, luminol, acradimum esters, biotin, green fluorescent protein (GFP), enhanced green fluorescent protein (EGFP), yellow fluorescent protein (YFP), enhanced yellow fluorescent protein (EYFP), blue fluorescent protein (BFP), red fluorescent protein (RFP), firefly luciferase, Renilla luciferase, NADPH, beta-galactosidase, horseradish peroxidase, glucose oxidase, alkaline phosphatase, chloramphenicol acetyl transferase, and urease.

[00171] In some embodiments, the one or more first oligonucleotides and second oligonucleotides bind to a different region of the target nucleic acid, or target site. In a pair, each target site is different, and the target sites are adjacent sites on the target nucleic acid, e.g. usually not more than 15 nucleotides distant, e.g. not more than 10, 8, 6, 4, or 2 nucleotides distant from the other site, and may be contiguous sites. Target sites are typically present on the same strand of the target nucleic acid in the same orientation. Target sites are also selected to provide a unique binding site, relative to other nucleic acids present in the cell. Each target site is generally from about 19 to about 25 nucleotides in length, e.g. from about 19 to 23 nucleotides, from about 19 to 21 nucleotides, or from about 19 to 20 nucleotides. The pair of first and second oligonucleotides are selected such that each oligonucleotide in the pair has a similar melting temperature for binding to its cognate target site, e.g. the T_m may be from about 50°C, from about 52°C, from about 55°C, from about 58°, from about 62°C, from about 65°C, from about 70°C, or from about 72°C. The GC content of the target site is generally selected to be no more than about 20%, no more than about 30%, no more than about 40%, no more than about 50%, no more than about 60%, no more than about 70%,

[00172] In some embodiments, the first oligonucleotide includes a first, second, and third complementarity region. The target site of the first oligonucleotide may refer to the first complementarity region. As summarized above, the first complementarity region of the first oligonucleotide may have a length of 19-25 nucleotides. In certain aspects, the second complementarity region of the first oligonucleotide has a length of 3-10 nucleotides, including, e.g., 4-8 nucleotides or 4-7 nucleotides. In some aspects, the second complementarity region of the first oligonucleotide has a length of 6 nucleotides. In some embodiments, the third complementarity region of the first oligonucleotide likewise has a length of 6 nucleotides. In such embodiments, the third complementarity region of the first oligonucleotide has a length of 3-10 nucleotides, including, e.g., 4-8 nucleotides or 4-7 nucleotides.

[00173] In some embodiments, second first oligonucleotide includes a first, second, and third complementarity region. The target site of the second oligonucleotide may refer to the second complementarity region. As summarized above, the second complementarity region of the second oligonucleotide may have a length of 19-25 nucleotides. In certain aspects, the first complementarity region of the first oligonucleotide has a length of 3-10 nucleotides, including, e.g., 4-8 nucleotides or 4-7 nucleotides. In some aspects, the first complementarity region of the first oligonucleotide has a length of 6 nucleotides. In some aspects, the first complementarity region of the second oligonucleotide includes the 5' end of the second oligonucleotide. In some embodiments, the third complementarity region of the second oligonucleotide likewise has a length of 6 nucleotides. In such embodiments, the third complementarity region of the second oligonucleotide has a length of 3-10 nucleotides, including, e.g., 4-8 nucleotides or 4-7 nucleotides. In further embodiments, the third complementarity region of the second oligonucleotide includes the 3' end of the second oligonucleotide. In some embodiments, the first complementarity region of the second oligonucleotide is adjacent to the third complementarity region of the second oligonucleotide.

[00174] In some aspects, the second oligonucleotide includes a barcode sequence, wherein the barcode sequence of the second oligonucleotide provides barcoding information for identification of the target nucleic acid. The term "barcode" refers to a nucleic acid sequence that is used to identify a single cell or a subpopulation of cells. Barcode sequences can be linked to a target nucleic acid of interest during amplification and used to trace back the amplicon to the cell from which the target nucleic acid originated. A barcode sequence can be added to a target nucleic acid of interest during amplification by carrying out amplification with an oligonucleotide that contains a region including the barcode sequence and a region that is complementary to the target nucleic acid such that the barcode sequence is incorporated into the final amplified target nucleic acid product (i.e., amplicon).

Tissue

[00175] As described herein, the methods disclosed include in situ sequencing technology of an intact tissue by at least contacting a fixed and permeabilized intact tissue with at least a pair of oligonucleotide primers under conditions to allow for specific hybridization. Tissue specimens suitable for use with the methods described herein generally include any type of tissue specimens collected from living or dead subjects, such as, e.g., biopsy specimens and autopsy specimens, of which include, but are not limited to, epithelium, muscle, connective, and nervous tissue. Tissue specimens may be collected and processed using the methods described herein and subjected to microscopic analysis immediately following processing, or may be preserved and subjected to microscopic analysis at a future time, e.g., after storage for an extended period of time. In some embodiments, the methods described herein may be used to preserve tissue specimens in a stable, accessible and fully intact form for future analysis. In some embodiments, the methods described herein may be used to analyze a previously-preserved or stored tissue specimen. In some embodiments, the intact tissue includes brain tissue such as visual cortex slices. In some embodiments, the intact tissue is a thin slice with a thickness of 5-20 μm , including, but not limited to, e.g., 5-18 μm , 5-15 μm , or 5-10 μm . In other embodiments, the intact tissue is a thick slice with a thickness of 20-200 μm , including, but not limited to, e.g., 20-150 μm , 50-100 μm , or 50-80 μm .

[00176] Aspects of the invention include fixing intact tissue. The term "fixing" or "fixation" as used herein is the process of preserving biological material (e.g., tissues, cells, organelles, molecules, etc.) from decay and/or degradation. Fixation may be accomplished using any convenient protocol. Fixation can include contacting the sample with a fixation reagent (i.e., a reagent that contains at least one fixative). Samples can be contacted by a fixation reagent for a wide range of times, which can depend on the temperature, the nature of the sample, and on the fixative(s). For example, a sample can be contacted by a fixation reagent for 24 or less hours, 18 or less hours, 12 or less hours, 8 or less hours, 6 or less hours, 4 or less hours, 2 or less hours, 60 or less minutes, 45 or less minutes, 30 or less minutes, 25 or less minutes, 20 or less minutes, 15 or less minutes, 10 or less minutes, 5 or less minutes, or 2 or less minutes.

[00177] A sample can be contacted by a fixation reagent for a period of time in a range of from 5 minutes to 24 hours, e.g., from 10 minutes to 20 hours, from 10 minutes to 18 hours, from 10 minutes to 12 hours, from 10 minutes to 8 hours, from 10 minutes to 6 hours, from 10 minutes to 4 hours, from 10 minutes to 2 hours, from 15 minutes to 20 hours, from 15 minutes to 18 hours, from 15 minutes to 12 hours, from 15 minutes to 8 hours, from 15 minutes to 6 hours, from 15 minutes to 4 hours, from 15 minutes to 2 hours, from 15 minutes to 1.5 hours, from 15 minutes to 1 hour, from 10

minutes to 30 minutes, from 15 minutes to 30 minutes, from 30 minutes to 2 hours, from 45 minutes to 1.5 hours, or from 55 minutes to 70 minutes.

[00178] A sample can be contacted by a fixation reagent at various temperatures, depending on the protocol and the reagent used. For example, in some instances a sample can be contacted by a fixation reagent at a temperature ranging from -22°C to 55°C, where specific ranges of interest include, but are not limited to 50 to 54°C, 40 to 44°C, 35 to 39°C, 28 to 32°C, 20 to 26°C, 0 to 6°C, and -18 to -22°C. In some instances a sample can be contacted by a fixation reagent at a temperature of -20°C, 4°C, room temperature (22-25°C), 30°C, 37°C, 42°C, or 52°C.

[00179] Any convenient fixation reagent can be used. Common fixation reagents include crosslinking fixatives, precipitating fixatives, oxidizing fixatives, mercurials, and the like. Crosslinking fixatives chemically join two or more molecules by a covalent bond and a wide range of cross-linking reagents can be used. Examples of suitable cross-linking fixatives include but are not limited to aldehydes (e.g., formaldehyde, also commonly referred to as "paraformaldehyde" and "formalin"; glutaraldehyde; etc.), imidoesters, NHS (N-Hydroxysuccinimide) esters, and the like. Examples of suitable precipitating fixatives include but are not limited to alcohols (e.g., methanol, ethanol, etc.), acetone, acetic acid, etc. In some embodiments, the fixative is formaldehyde (i.e., paraformaldehyde or formalin). A suitable final concentration of formaldehyde in a fixation reagent is 0.1 to 10%, 1-8%, 1-4%, 1-2%, 3-5%, or 3.5-4.5%, including about 1.6% for 10 minutes. In some embodiments the sample is fixed in a final concentration of 4% formaldehyde (as diluted from a more concentrated stock solution, e.g., 38%, 37%, 36%, 20%, 18%, 16%, 14%, 10%, 8%, 6%, etc.). In some embodiments the sample is fixed in a final concentration of 10% formaldehyde. In some embodiments the sample is fixed in a final concentration of 1% formaldehyde. In some embodiments, the fixative is glutaraldehyde. A suitable concentration of glutaraldehyde in a fixation reagent is 0.1 to 1%. A fixation reagent can contain more than one fixative in any combination. For example, in some embodiments the sample is contacted with a fixation reagent containing both formaldehyde and glutaraldehyde.

[00180] The terms "permeabilization" or "permeabilize" as used herein refer to the process of rendering the cells (cell membranes etc.) of a sample permeable to experimental reagents such as nucleic acid probes, antibodies, chemical substrates, etc. Any convenient method and/or reagent for permeabilization can be used. Suitable permeabilization reagents include detergents (e.g., Saponin, Triton X-100, Tween-20, etc.), organic fixatives (e.g., acetone, methanol, ethanol, etc.), enzymes, etc. Detergents can be used at a range of concentrations. For example, 0.001%-1% detergent, 0.05%-0.5% detergent, or 0.1%-0.3% detergent can be used for permeabilization (e.g., 0.1 %

Saponin, 0.2% tween-20, 0.1-0.3% triton X-100, etc.). In some embodiments methanol on ice for at least 10 minutes is used to permeabilize.

[00181] In some embodiments, the same solution can be used as the fixation reagent and the permeabilization reagent. For example, in some embodiments, the fixation reagent contains 0.1%-10% formaldehyde and 0.001%-1% saponin. In some embodiments, the fixation reagent contains 1% formaldehyde and 0.3% saponin.

[00182] A sample can be contacted by a permeabilization reagent for a wide range of times, which can depend on the temperature, the nature of the sample, and on the permeabilization reagent(s). For example, a sample can be contacted by a permeabilization reagent for 24 or more hours, 24 or less hours, 18 or less hours, 12 or less hours, 8 or less hours, 6 or less hours, 4 or less hours, 2 or less hours, 60 or less minutes, 45 or less minutes, 30 or less minutes, 25 or less minutes, 20 or less minutes, 15 or less minutes, 10 or less minutes, 5 or less minutes, or 2 or less minutes. A sample can be contacted by a permeabilization reagent at various temperatures, depending on the protocol and the reagent used. For example, in some instances a sample can be contacted by a permeabilization reagent at a temperature ranging from -82°C to 55°C, where specific ranges of interest include, but are not limited to: 50 to 54°C, 40 to 44°C, 35 to 39°C, 28 to 32°C, 20 to 26°C, 0 to 6°C, -18 to -22 °C, and -78 to -82°C. In some instances a sample can be contacted by a permeabilization reagent at a temperature of -80°C, -20°C, 4°C, room temperature (22-25°C), 30°C, 37°C, 42°C, or 52°C.

[00183] In some embodiments, a sample is contacted with an enzymatic permeabilization reagent. Enzymatic permeabilization reagents that permeabilize a sample by partially degrading extracellular matrix or surface proteins that hinder the permeation of the sample by assay reagents. Contact with an enzymatic permeabilization reagent can take place at any point after fixation and prior to target detection. In some instances the enzymatic permeabilization reagent is proteinase K, a commercially available enzyme. In such cases, the sample is contacted with proteinase K prior to contact with a post-fixation reagent. Proteinase K treatment (i.e., contact by proteinase K; also commonly referred to as "proteinase K digestion") can be performed over a range of times at a range of temperatures, over a range of enzyme concentrations that are empirically determined for each cell type or tissue type under investigation. For example, a sample can be contacted by proteinase K for 30 or less minutes, 25 or less minutes, 20 or less minutes, 15 or less minutes, 10 or less minutes, 5 or less minutes, or 2 or less minutes. A sample can be contacted by 1 µg/ml or less, 2 µg/ml or less, 4 µg/ml or less, 8 µg/ml or less, 10 µg/ml or less, 20 µg/ml or less, 30 µg/ml or less, 50 µg/ml or less, or 100µg/ml or less proteinase K. A sample can be contacted by proteinase K at a temperature ranging from 2°C to 55°C, where specific ranges of interest include, but are not limited to: 50 to 54°C, 40 to

44°C, 35 to 39°C, 28 to 32°C, 20 to 26°C, and 0 to 6°C. In some instances a sample can be contacted by proteinase K at a temperature of 4°C, room temperature (22-25°C), 30°C, 37°C, 42°C, or 52°C. In some embodiments, a sample is not contacted with an enzymatic permeabilization reagent. In some embodiments, a sample is not contacted with proteinase K. Contact of an intact tissue with at least a fixation reagent and a permeabilization reagent results in the production of a fixed and permeabilized tissue.

Ligase

[00184] In some embodiments, the methods disclosed include adding ligase to ligate the second oligonucleotide and generate a closed nucleic acid circle. In some embodiments, the adding ligase includes adding DNA ligase. In alternative embodiments, the second oligonucleotide is provided as a closed nucleic acid circle, and the step of adding ligase is omitted. In certain embodiments, ligase is an enzyme that facilitates the sequencing of a target nucleic acid molecule.

[00185] The term "ligase" as used herein refers to an enzyme that is commonly used to join polynucleotides together or to join the ends of a single polynucleotide. Ligases include 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 include bacterial ligases such as E. coli DNA ligase and Taq DNA ligase, Ampligase® thermostable DNA ligase (Epicentre®Technologies Corp., part of Illumina®, Madison, Wis.) and phage ligases such as T3 DNA ligase, T4 DNA ligase and T7 DNA ligase and mutants thereof.

Rolling Circle Amplification

[00186] In some embodiments, the methods of the invention include the step of performing rolling circle amplification in the presence of a nucleic acid molecule, wherein the performing includes using the second oligonucleotide as a template and the first oligonucleotide as a primer for a polymerase to form one or more amplicons. In such embodiments, a single-stranded, circular polynucleotide template is formed by ligation of the second nucleotide, which circular polynucleotide includes a region that is complementary to the first oligonucleotide. Upon addition of a DNA polymerase in the presence of appropriate dNTP precursors and other cofactors, the first oligonucleotide is elongated by replication of multiple copies of the template. This amplification product can be readily detected by binding to a detection probe. In some embodiments, the polymerase is preincubated without

dNTPs to allow the polymerase to penetrate the sample uniformly before performing rolling circle amplification.

[00187] In some embodiments, only when a first oligonucleotide and second oligonucleotide hybridize to the same target nucleic acid molecule, the second oligonucleotide can be circularized and rolling-circle amplified to generate a cDNA nanoball (i.e., amplicon) containing multiple copies of the cDNA. The term “amplicon” refers to the amplified nucleic acid product of a PCR reaction or other nucleic acid amplification process. In some embodiments, amine-modified nucleotides are spiked into the rolling circle amplification reaction.

[00188] Techniques for rolling circle amplification are known in the art (see, e.g., Baner et al, Nucleic Acids Research, 26:5073-5078, 1998; Lizardi et al, Nature Genetics 19:226, 1998; 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:e1 18, 2001 ; Dean et al. Genome Res. 1 1 :1095- 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). In some embodiments the polymerase is phi29 DNA polymerase.

[00189] In certain aspects, the nucleic acid molecule includes an amine-modified nucleotide. In such embodiments, the amine-modified nucleotide includes an acrylic acid N-hydroxysuccinimide moiety modification. Examples of other amine-modified nucleotides include, 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.

Amplicon Embedding in a Tissue-hydrogel Setting

[00190] In some embodiments, the methods disclosed include embedding one or more amplicons in the presence of hydrogel subunits to form one or more hydrogel-embedded amplicons. The hydrogel-tissue chemistry described includes covalently attaching nucleic acids to in situ synthesized hydrogel for tissue clearing, enzyme diffusion, and multiple-cycle sequencing while an existing hydrogel-tissue chemistry method cannot. In some embodiments, to enable amplicon embedding in the tissue-hydrogel setting, amine-modified nucleotides are spiked into the rolling circle amplification reaction, functionalized with an acrylamide moiety using acrylic acid N-hydroxysuccinimide esters, and copolymerized with acrylamide monomers to form a hydrogel.

[00191] As used herein, the terms “hydrogel” or “hydrogel network” mean a network of polymer chains that are water-insoluble, sometimes found as a colloidal gel in which water is the dispersion medium. In other words, hydrogels are a class of polymeric materials that can absorb large amounts of water without dissolving. Hydrogels can contain over 99% water and may include natural or synthetic polymers, or a combination thereof. Hydrogels also possess a degree of flexibility very similar to

natural tissue, due to their significant water content. A detailed description of suitable hydrogels may be found in published U.S. patent application 20100055733, herein specifically incorporated by reference. As used herein, the terms "hydrogel subunits" or "hydrogel precursors" mean hydrophilic monomers, prepolymers, or polymers that can be crosslinked, or "polymerized", to form a three-dimensional (3D) hydrogel network. Without being bound by any scientific theory, it is believed that this fixation of the biological specimen in the presence of hydrogel subunits crosslinks the components of the specimen to the hydrogel subunits, thereby securing molecular components in place, preserving the tissue architecture and cell morphology.

[00192] In some embodiments, the embedding includes copolymerizing the one or more amplicons with acrylamide. As used herein, the term "copolymer" describes a polymer which contains more than one type of subunit. The term encompasses polymer which include two, three, four, five, or six types of subunits.

[00193] In certain aspects, the embedding includes clearing the one or more hydrogel-embedded amplicons wherein the target nucleic acid is substantially retained in the one or more hydrogel-embedded amplicons. In such embodiments, the clearing includes substantially removing a plurality of cellular components from the one or more hydrogel-embedded amplicons. In some other embodiments, the clearing includes substantially removing lipids and/or proteins from the one or more hydrogel-embedded amplicons. As used herein, the term "substantially" means that the original amount present in the sample before clearing has been reduced by approximately 70% or more, such as by 75% or more, such as by 80% or more, such as by 85% or more, such as by 90% or more, such as by 95% or more, such as by 99% or more, such as by 100%.

[00194] In some embodiments, clearing the hydrogel-embedded amplicons includes performing electrophoresis on the specimen. In some embodiments, the amplicons are electrophoresed using a buffer solution that includes an ionic surfactant. In some embodiments, the ionic surfactant is sodium dodecyl sulfate (SDS). In some embodiments, the specimen is electrophoresed using a voltage ranging from about 10 to about 60 volts. In some embodiments, the specimen is electrophoresed for a period of time ranging from about 15 minutes up to about 10 days. In some embodiments, the methods further involve incubating the cleared specimen in a mounting medium that has a refractive index that matches that of the cleared tissue. In some embodiments, the mounting medium increases the optical clarity of the specimen. In some embodiments, the mounting medium includes glycerol.

Cells

[00195] Methods disclosed herein include a method for in situ gene sequencing of a target nucleic acid in a cell in an intact tissue. In certain embodiments, the cell is present in a population of cells. In certain other embodiments, the population of cells includes a plurality of cell types including, but not limited to, excitatory neurons, inhibitory neurons, and non-neuronal cells. Cells for use in the assays of the invention can be an organism, a single cell type derived from an organism, or can be a mixture of cell types. Included are naturally occurring cells and cell populations, genetically engineered cell lines, cells derived from transgenic animals, etc. Virtually any cell type and size can be accommodated. Suitable cells include bacterial, fungal, plant and animal cells. In one embodiment of the invention, the cells are mammalian cells, e.g. complex cell populations such as naturally occurring tissues, for example blood, liver, pancreas, neural tissue, bone marrow, skin, and the like. Some tissues may be disrupted into a monodisperse suspension. Alternatively, the cells may be a cultured population, e.g. a culture derived from a complex population, a culture derived from a single cell type where the cells have differentiated into multiple lineages, or where the cells are responding differentially to stimulus, and the like.

[00196] Cell types that can find use in the subject invention include stem and progenitor cells, e.g. embryonic stem cells, hematopoietic stem cells, mesenchymal stem cells, neural crest cells, etc., endothelial cells, muscle cells, myocardial, smooth and skeletal muscle cells, mesenchymal cells, epithelial cells; hematopoietic cells, such as lymphocytes, including T-cells, such as Th1 T cells, Th2 T cells, ThO T cells, cytotoxic T cells; B cells, pre- B cells, etc.; monocytes; dendritic cells; neutrophils; and macrophages; natural killer cells; mast cells, etc.; adipocytes, cells involved with particular organs, such as thymus, endocrine glands, pancreas, brain, such as neurons, glia, astrocytes, dendrocytes, etc. and genetically modified cells thereof. Hematopoietic cells may be associated with inflammatory processes, autoimmune diseases, etc., endothelial cells, smooth muscle cells, myocardial cells, etc. may be associated with cardiovascular diseases; almost any type of cell may be associated with neoplasias, such as sarcomas, carcinomas and lymphomas; liver diseases with hepatic cells; kidney diseases with kidney cells; etc.

[00197] The cells may also be transformed or neoplastic cells of different types, e.g. carcinomas of different cell origins, lymphomas of different cell types, etc. The American Type Culture Collection (Manassas, VA) has collected and makes available over 4,000 cell lines from over 150 different species, over 950 cancer cell lines including 700 human cancer cell lines. The National Cancer Institute has compiled clinical, biochemical and molecular data from a large panel of human tumor cell lines, these are available from ATCC or the NCI (Phelps et al. (1996) Journal of Cellular Biochemistry Supplement 24:32-91). Included are different cell lines derived spontaneously, or

selected for desired growth or response characteristics from an individual cell line; and may include multiple cell lines derived from a similar tumor type but from distinct patients or sites.

[00198] Cells may be non-adherent, e.g. blood cells including monocytes, T cells, B-cells; tumor cells, etc., or adherent cells, e.g. epithelial cells, endothelial cells, neural cells, etc. In order to profile adherent cells, they may be dissociated from the substrate that they are adhered to, and from other cells, in a manner that maintains their ability to recognize and bind to probe molecules.

[00199] Such cells can be acquired from an individual using, e.g., a draw, a lavage, a wash, surgical dissection etc., from a variety of tissues, e.g., blood, marrow, a solid tissue (e.g., a solid tumor), ascites, by a variety of techniques that are known in the art. Cells may be obtained from fixed or unfixed, fresh or frozen, whole or disaggregated samples. Disaggregation of tissue may occur either mechanically or enzymatically using known techniques.

Examples of Non-Limiting Aspects of the Disclosure

[00200] Aspects, including embodiments, of the present subject matter described above may be beneficial alone or in combination, with one or more other aspects or embodiments. Without limiting the foregoing description, certain non-limiting aspects of the disclosure numbered 1-34 are provided below. As will be apparent to those of skill in the art upon reading this disclosure, each of the individually numbered aspects may be used or combined with any of the preceding or following individually numbered aspects. This is intended to provide support for all such combinations of aspects and is not limited to combinations of aspects explicitly provided below:

1. A sequencing device comprising:

(a) an illumination and detection module comprising a spinning disk confocal component comprising a plurality of laser lines for illumination with flat illumination correction, wherein the plurality of laser lines are used to illuminate a sample with excitation light at one or more wavelengths, a bandpass emission filter, a long-pass image splitter, a first camera that detects fluorescence emissions in a first wavelength range and a second camera that detects fluorescence emissions in a second wavelength range, wherein the first camera and the second camera can detect emissions simultaneously;

(b) a microscope module comprising a motorized stage capable of multi-axis positioning along x, y, and z axes, an objective Z drive, an objective turret wheel comprising multiple objectives, wherein each objective provides a different magnification, wherein one or more objectives are immersion objectives, wherein each immersion objective has an objective immersion collar, and optics, wherein the optics route light from the objectives to the illumination and detection module;

(c) an automated immersion media module comprising i) a container comprising immersion media, ii) fluidic lines coupled to the container and to the objective immersion collars of the immersion objectives of the microscope module, wherein the fluidic lines carry immersion media to and from the objective immersion collars, wherein the immersion collars capture excess immersion media, and iii) a series of pumps connected to the fluidic lines and to a microcontroller, wherein the microcontroller controls the pumps addition and removal of the immersion media through the fluidic lines, wherein the automated immersion media module provides controlled volumes of the immersion media to the objective immersion collars at the tops of the immersion objectives during imaging;

(d) a multi-well plate, wherein the motorized stage can be moved to position a well of the multi-well plate under the objective used for imaging;

(e) a fluidic coupling tower, wherein the fluidic coupling tower is on top of the motorized stage and positions the fluidic lines in wells of the multi-well plate;

(f) a fluidic management module comprising a symmetrical rotary valve comprising a rotary valve mechanism, a pump, wherein the pump is connected to the fluidic lines, and bubble detectors, wherein the bubble detectors are positioned on either side of the fluidic lines leading to the pump, wherein the fluidic management module allows unidirectional or bidirectional movement of reagents, buffers, and waste through the fluidic lines;

(g) a reagent, buffer, and waste module comprising a i) sliding tray, wherein reagent cartridges and buffer cartridges can be positioned in the sliding tray and coupled to the fluidic management module, ii) a waste module comprising a waste container, wherein the waste container is coupled to a fluidic line from the fluid management pump, and iii) a capping mechanism, wherein the capping mechanism closes the waste container when the waste container is removed from the system for waste disposal and opens the waste container when the waste container is placed back into the system;

(h) an electrical module comprising: i) a first firmware board controlling media dispensing from the automated immersion media module and ii) a second firmware board controlling the fluid management module and the reagent, buffer, and waste module, wherein the electrical module regulates power to the other modules of the system; and

(i) a processor programmed to provide a user interface and operate the modules of the sequencing device.

2. The sequencing device of aspect 1, wherein the plurality of laser lines comprises at least 5 laser lines.

3. The sequencing device of aspect 2, wherein the bandpass emission filter is a penta-bandpass emission filter.

4. The sequencing device of any one of aspects 1-3, wherein the motorized stage has a piezo z-axis.

5. The sequencing device of any one of aspects 1-4, wherein the immersion media is water.

6. The sequencing device of any one of aspects 1-5, wherein the immersion media is filtered and bubble-free.

7. The sequencing device of any one of aspects 1-6, further comprising an O-ring and a shrink-wrapped coating over each objective.

8. The sequencing device of any one of aspects 1-7, further comprising a pressure monitor to monitor pressure in the fluidic lines, wherein increases in pressure in a fluidic line can be used to detect a potential blockage of the fluidic line.

9. The sequencing device of any one of aspects 1-8, further comprising a plurality of light-emitting diodes (LEDs), wherein each LED can emit light to provide a status indication for the system.

10. The sequencing device of any one of aspects 1-9, further comprising a display component for displaying information and providing a user interface.

11. The sequencing device of any one of aspects 1-10, wherein the processor is further programmed to perform steps comprising:

- (a) locating a selected sample in the multi-well plate;
- (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning;
- (c) using the signal to segment an XY bounding box around the sample;
- (d) imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera

binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent;

(e) displaying the image produced in step (d);

(f) providing an interface for a user to select a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample;

(g) imaging the sample in the selected XY region of interest across the previously sampled Z extents;

(h) calculating a volume of the region of interest in the sample and displaying the calculated sample volume of the region of interest to the user;

(i) segmenting the image of the sample in the region of interest along the Z extents;

(j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and

(k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.

12. The sequencing device of any one of aspects 1-11, wherein the processor is further programmed to perform steps comprising:

providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents that are available and the selected sequencing protocol;

providing constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and

suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints.

13. The sequencing device of any one of aspects 1-12, wherein the processor is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.

14. The sequencing device of any one of aspects 1-13, wherein the processor is further programmed to perform steps comprising:

performing a rapid confocal sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position;

determining the sample top and bottom interface using a segmentation method; and

setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.

15. The sequencing device of any one of aspects 1-14, wherein the sequencing is in situ sequencing of a target nucleic acid in a tissue sample.

16. The sequencing device of aspect 15, wherein the tissue sample is a tissue slice having a thickness of 20 μm to 200 μm .

17. The sequencing device of any one of aspects 1-16, wherein the in situ sequencing is sequential or combinatorial in situ sequencing.

18. The sequencing device of anyone of aspects 1-17, wherein the microscope module comprises an epifluorescent microscope, a confocal microscope, a structured illumination microscope, or a light sheet or oblique-plane light sheet microscope.

19. The sequencing device of aspect 18, wherein the confocal microscope is a spinning disk or point scanning confocal microscope.

20. A method of using the sequencing device of any one of aspects 1-19, the method comprising:

loading samples into the multi-well plate;

selecting which samples in the multi-well plate are sequenced;

selecting a sequencing protocol; and

sequencing nucleic acids in the selected samples using the sequencing device of any one of aspects 1-19.

21. The method of aspect 20, wherein the sequencing is in situ volumetric sequencing of tissue samples.

22. The method of aspect 20 or 21, wherein the tissue samples are tissue slices having a thickness of 20-200 μm .

23. The method of any one of aspects 20-22, wherein the in situ sequencing is sequential or combinatorial in situ sequencing.

24. A computer implemented method, the computer performing steps comprising:

- (a) locating a selected sample in the multi-well plate;
- (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning;
- (c) using the signal to segment an XY bounding box around the sample;
- (d) imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent;
- (e) displaying the image produced in step (d);
- (f) providing an interface for a user to select a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample;
- (g) imaging the sample in the selected XY region of interest across the previously sampled Z extents;
- (h) calculating a sample volume of the region of interest and displaying the calculated sample volume of the region of interest to the user;
- (i) segmenting the image of the sample in the region of interest along Z extents;
- (j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and
- (k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.

25. A non-transitory computer-readable medium comprising program instructions that, when executed by a processor in a computer, causes the processor to perform the method of aspect 24.

26. A computer implemented method, the computer performing steps comprising:
providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents available and the selected sequencing protocol;
providing constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and
suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints.

27. The computer implemented method of aspect 26, wherein the computer is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.

28. A non-transitory computer-readable medium comprising program instructions that, when executed by a processor in a computer, causes the processor to perform the method of aspect 26 or 27.

29. A computer implemented method, the computer performing steps comprising:
performing a rapid confocal sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position;
determining the sample top and bottom interface using a segmentation method; and
setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.

30. A non-transitory computer-readable medium comprising program instructions that, when executed by a processor in a computer, causes the processor to perform the method of aspect 29.

31. An automated immersion media module comprising:

(a) a container comprising immersion media;

(b) fluidic lines coupled to the container and to the objective immersion collars of the objectives of the microscope module, wherein the fluidic lines carry immersion media to and from an objective immersion collar on an immersion objective, wherein the immersion collar captures excess immersion media; and

(c) a series of pumps connected to the fluidic lines and to a microcontroller, wherein the microcontroller controls the pumps addition and removal of the immersion media through the fluidic lines, wherein the automated immersion media module provides controlled volumes of the immersion media to the objective immersion collars at the tops of the objectives during imaging.

32. A method of using the automated immersion media module of aspect 31, the method comprising using the automated immersion media module of aspect 31 to deliver immersion media to an objective immersion collar attached to an immersion objective of a microscope.

33. A fluidic management module comprising a symmetrical rotary valve comprising a rotary valve mechanism, a pump, wherein the pump is connected to the fluidic lines, and bubble detectors, wherein the bubble detectors are positioned on either side of the fluidic lines leading to the pump, wherein the fluidic management module allows bidirectional or unidirectional movement of reagents, buffers, and waste through the fluidic lines.

34. A reagent, buffer, and waste module comprising:

(a) a sliding tray, wherein reagent cartridges and buffer cartridges can be positioned in the sliding tray and coupled to the fluidic management module;

(b) a waste module comprising a waste container, wherein the waste container is coupled to a fluidic line from the fluid management pump; and

(c) a capping mechanism, wherein the capping mechanism closes the waste container when the waste container is removed from the system for waste disposal and opens the waste container when the waste container is placed back into the system.

EXPERIMENTAL

[00201] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not

intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, *etc.*) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

[00202] All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[00203] The present invention has been described in terms of particular embodiments found or proposed by the present inventor to comprise preferred modes for the practice of the invention. It will be appreciated by those of skill in the art that, in light of the present disclosure, numerous modifications and changes can be made in the particular embodiments exemplified without departing from the intended scope of the invention. For example, due to codon redundancy, changes can be made in the underlying DNA sequence without affecting the protein sequence. Moreover, due to biological functional equivalency considerations, changes can be made in protein structure without affecting the biological action in kind or amount. All such modifications are intended to be included within the scope of the appended claims.

Example 1

Volumetric Next-Generation *In Situ* Sequencer

Overview

- [00204]
1. Integrated fluidics and rapid confocal platform and up to 5 channel imaging
 2. Automated immersion media module
 - Routines for robust bubble prevention
 3. Sample/fluidic coupling tower
 4. Custom waste container
 5. Sequencer software
 - a. GUI
 - b. Rapid sample find algorithms (general for 3D sequencing samples)

- c. Sequencing time, parallelism, reagent consumption, and disk space joint optimization (general for 3D sequencing samples)
- d. Closed-loop detection of sample interface and xyz position for minimal drift of data across multiple rounds.

Technical Description

[00205] Samples were sequenced on an automated, integrated fluidic and imaging platform capable of conducting multiple rounds of sequencing and imaging cycles on many samples in parallel. The sequencer has several key modules:

[00206] The illumination and detection module, integrating components from Andor Technologies, consists of a spinning disk confocal component, 5 laser line illumination, Borealis flat illumination correction, a penta-bandpass emission filter, a long-pass image splitter, and two detection cameras for simultaneous detection of emissions generated by short and long wavelength illumination.

[00207] The custom microscope module combines a motorized XYZ stage (with a piezo Z), objective Z drive, turret wheel, multiple objectives ranging low to high magnification, and optics routing light from the objectives to the illumination and detection module.

[00208] A novel automated immersion media module, which couples to objectives requiring immersion media between the objective and the coverglass, provides precisely controlled volumes of filtered, bubble-free media (such as water) to the tops of objectives, managing both inflow and outflow. This module enables sequencing in parallel of samples with a high-magnification immersion objective that otherwise would require manual addition of immersion media. The addition and removal of the immersion media is managed by a series of pumps and a microcontroller such that surface tension forces on the coverglass are minimized (minimize shifts in the Z dimension of the sample), bubbles are avoided, liquid overflow is prevented, and sufficient volume is present on the objective for prolonged imaging. This involves algorithmic coordination of the liquid flowrate, position of the objective in x, y, and z relative to the sample, liquid replacement steps to remove bubbles, and fluid replacement timing. For instance, after feeding a calibrated volume of immersion media (typically water) onto the objective through the collar, diagonal stage movements are performed that preclude bubble formation from occurring. The immersion collar captures excess immersion media, preventing media access to the objectives through multiple O-rings and a shrink-wrapped coating over the objectives. Fluidic lines bringing immersion media to and from the objective immersion collars are managed by a central tower that integrates into the objective turret.

[00209] A novel fluidic coupling tower placed on the XYZ stage positions fluidic lines into a multi-well sample plate. This coupler enables easy addition and removal of the sample from the microscope

stage and mating of fluidic lines to the samples, while minimizing the structural load on the automated microscope stage (excess weight on the stage inhibits precise positioning and rapid movements). Beam breaks detect whether the sample is fully coupled into the system and structural elements prevent damage to the fluidic lines or harm to the user during operation.

[00210] The fluid management module allows bidirectional movement of reagents, buffers, and waste via a symmetrical rotary valve, pump, rotary valve mechanism. Pump movements are controlled and line pressures monitored by custom firmware. Liquids moving towards the pump are primed for precise volume movements via the use of bubble detectors on either side of the lines leading to the pump (before the rotary valves). Pressures in the lines are monitored for nominal fluid flow and to detect potential blockages of the lines.

[00211] The reagent, buffer, and waste module organizes the addition of sequencing reagents and buffers to the system and the removal of waste from the system. It consists of a sliding tray into which custom reagent and buffer cartridges are positioned and subsequently coupled to the fluidic management module in an automated fashion. The waste module receives an outflow line from the fluid management pump via a capping mechanism design that ensures that the waste container is closed when removed from the system for waste disposal, but open when placed back into the system.

[00212] An electrical module regulates power for the various components of the system. One firmware board controls the automated immersion media dispensers and a second firmware board controls the fluid management module and the reagent, buffer, and waste modules, in addition to LEDs displaying relevant status indications for the system.

[00213] A custom computer software program, including both backend and GUI modules, provides interface between a user and the sequencer firmware and hardware, including sequencing run set up, sequencing run options, sample region of interest (ROI) definition, the high level operations required for sequencing and imaging, high level control of the automated immersion system, parallelization of sequencing across samples, logging, error monitoring, data acquisition, management and transfer, and run progress monitoring. The computer software program contains several novel algorithms for run set up and for consistent imaging of sequencing rounds. An automated, 3D sample-find and ROI specification algorithm rapidly detects signal in the XY plane from a sample via low magnification, widefield imaging mode acquisition with maximal camera binning. This signal is used to segment an XY bounding box around the sample. Subsequently, this XY bound is rapidly sub-sampled in confocal imaging mode in Z, with higher magnification but maximal camera binning, in order to determine the approximate Z extent of the sample. A single Z plane is rapidly collected through the midpoint of the Z extents previously determined, across the XY

extent, and displayed to the user. At this point, the user can select through an interface the desired XY ROI to be imaged during sequencing for a given sample. Subsequently, a more detailed acquisition of the XY ROI is acquired across the previously sampled Z extent. This volume is displayed to the user and is additionally used to segment in Z the sample extents in the given ROI. The user may more finely adjust the Z extent of the volume via the interface before sequencing begins. The imaging extents derived from this ROI definition are automatically converted into the appropriate montaged fields of view for a given imaging objective, and further into microscope stage positions, objective Z positioning, and piezo bounds for optimal imaging of the XYZ ROI during sequencing. Using this ROI definition algorithm affords rapid, semi-automatic ROI definition in 3D and minimizes the amount of user interaction (for example, arbitrary manual control and search of a well for a sample and montage definition and testing, which is excessively time consuming and difficult to an inexperienced user). This automated ROI procedure is performed for each sample well that the user intends to sequence.

[00214] A second algorithm guides the user in sequencing run set up to ensure optimal sample parallelism and sequencer time efficiency, while preventing collection of a prohibitive (to store or transfer) amount of data or use of more buffer and reagent than is available. This algorithm is essential for spatial sequencing approaches, especially volumetric ones, since users are interested in ROIs of varying XYZ extents across different sample types, and excess imaging volume beyond the ROI is irrelevant to the user. Moreover, a single sample acquisition can produce many terabytes of raw data, which need to be stored and/or transferred. Further, because a sequencing kit has a maximum capacity of buffer and reagent, the user is limited in how many samples can be defined for use, especially depending on the exact sequencing protocol and number of rounds that are required. Thus, an optimization algorithm is required that balances multiple sample sequencing time, buffer/reagent availability, sequencing protocols, total data collection and transfer capacity, and ROI extents in XYZ. The algorithm may place hard constraints on the total sequencing time (for example, three days), the total data acquired and rate of acquisition (related to available data transfer / off-loading rate), maximum ROI budget across all samples, and maximum available buffer/reagents, while suggesting combinations of ROIs and sample protocols that pack the maximum amount of desired sequencing into these constraints. This optimization also involves a subroutine optimizer for sample sequencing parallelization, as different numbers of samples and imaging types may yield optimal parallelization solutions of varying time.

[00215] Another algorithm applied during sequencing of the sample obviates the need for so-called perfect focus system hardware, which utilize angled infrared laser light reflected from the coverglass to find, in closed loop, a Z position set point. This algorithm uses a rapid confocal sweep in Z, at the

starting XY position of a given sample montage, to determine the characteristic Z profile of the sample at that position. Segmentation methods are used to determine the sample top and bottom interface, and the objective Z position is set at a fixed distance from this interface at the beginning of a montage. This ensures that, even if there has been drift in Z of the sample relative to the stage and objective across rounds, that the sample Z drift is reduced to below some tolerance, facilitating downstream subpixel registration across rounds during post-acquisition processing. This algorithm is especially important in the absence of a perfect focus hardware system for combinatorial sequencing, in which the precise position of diffraction limited spots must be aligned across rounds, and is additionally important when the sample thickness is small, in which case drift in Z that may occur otherwise on any round is large relative to the total sample extent in Z, resulting in a larger percentage of data loss on each round at the edges of the sample extents. Such an algorithm may also be performed for alignment in XY, though due to the aspect ratio of acquisitions, the percentage of data loss from unanticipated sample movement in XY is generally minimal.

WHAT IS CLAIMED IS:

1. A sequencing device comprising:

(a) an illumination and detection module comprising a spinning disk confocal component comprising a plurality of laser lines for illumination with flat illumination correction, wherein the plurality of laser lines are used to illuminate a sample with excitation light at one or more wavelengths, a bandpass emission filter, a long-pass image splitter, a first camera that detects fluorescence emissions in a first wavelength range and a second camera that detects fluorescence emissions in a second wavelength range, wherein the first camera and the second camera can detect emissions simultaneously;

(b) a microscope module comprising a motorized stage capable of multi-axis positioning along x, y, and z axes, an objective Z drive, an objective turret wheel comprising multiple objectives, wherein each objective provides a different magnification, wherein one or more objectives are immersion objectives, wherein each immersion objective has an objective immersion collar, and optics, wherein the optics route light from the objectives to the illumination and detection module;

(c) an automated immersion media module comprising i) a container comprising immersion media, ii) fluidic lines coupled to the container and to the objective immersion collars of the immersion objectives of the microscope module, wherein the fluidic lines carry immersion media to and from the objective immersion collars, wherein the immersion collars capture excess immersion media, and iii) a series of pumps connected to the fluidic lines and to a microcontroller, wherein the microcontroller controls the pumps addition and removal of the immersion media through the fluidic lines, wherein the automated immersion media module provides controlled volumes of the immersion media to the objective immersion collars at the tops of the immersion objectives during imaging;

(d) a multi-well plate, wherein the motorized stage can be moved to position a well of the multi-well plate under the objective used for imaging;

(e) a fluidic coupling tower, wherein the fluidic coupling tower is on top of the motorized stage and positions the fluidic lines in wells of the multi-well plate;

(f) a fluidic management module comprising a symmetrical rotary valve comprising a rotary valve mechanism, a pump, wherein the pump is connected to the fluidic lines, and bubble detectors, wherein the bubble detectors are positioned on either side of the fluidic lines leading to the pump, wherein the fluidic management module allows unidirectional or bidirectional movement of reagents, buffers, and waste through the fluidic lines;

(g) a reagent, buffer, and waste module comprising a i) sliding tray, wherein reagent cartridges and buffer cartridges can be positioned in the sliding tray and coupled to the fluidic

management module, ii) a waste module comprising a waste container, wherein the waste container is coupled to a fluidic line from the fluid management pump, and iii) a capping mechanism, wherein the capping mechanism closes the waste container when the waste container is removed from the system for waste disposal and opens the waste container when the waste container is placed back into the system;

(h) an electrical module comprising: i) a first firmware board controlling media dispensing from the automated immersion media module and ii) a second firmware board controlling the fluid management module and the reagent, buffer, and waste module, wherein the electrical module regulates power to the other modules of the system; and

(i) a processor programmed to provide a user interface and operate the modules of the sequencing device.

2. The sequencing device of claim 1, wherein the plurality of laser lines comprises at least 5 laser lines.

3. The sequencing device of claim 2, wherein the bandpass emission filter is a penta-bandpass emission filter.

4. The sequencing device of any one of claims 1-3, wherein the motorized stage has a piezo z-axis.

5. The sequencing device of any one of claims 1-4, wherein the immersion media is water.

6. The sequencing device of any one of claims 1-5, wherein the immersion media is filtered and bubble-free.

7. The sequencing device of any one of claims 1-6, further comprising an O-ring and a shrink-wrapped coating over each objective.

8. The sequencing device of any one of claims 1-7, further comprising a pressure monitor to monitor pressure in the fluidic lines, wherein increases in pressure in a fluidic line can be used to detect a potential blockage of the fluidic line.

9. The sequencing device of any one of claims 1-8, further comprising a plurality of light-emitting diodes (LEDs), wherein each LED can emit light to provide a status indication for the system.

10. The sequencing device of any one of claims 1-9, further comprising a display component for displaying information and providing a user interface.

11. The sequencing device of any one of claims 1-10, wherein the processor is further programmed to perform steps comprising:

- (a) locating a selected sample in the multi-well plate;
- (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning;
- (c) using the signal to segment an XY bounding box around the sample;
- (d) imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent;
- (e) displaying the image produced in step (d);
- (f) providing an interface for a user to select a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample;
- (g) imaging the sample in the selected XY region of interest across the previously sampled Z extents;
- (h) calculating a volume of the region of interest in the sample and displaying the calculated sample volume of the region of interest to the user;
- (i) segmenting the image of the sample in the region of interest along the Z extents;
- (j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and
- (k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.

12. The sequencing device of any one of claims 1-11, wherein the processor is further programmed to perform steps comprising:

providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents that are available and the selected sequencing protocol;

providing constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and

suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints.

13. The sequencing device of any one of claims 1-12, wherein the processor is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.

14. The sequencing device of any one of claims 1-13, wherein the processor is further programmed to perform steps comprising:

performing a rapid confocal sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position;

determining the sample top and bottom interface using a segmentation method; and

setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.

15. The sequencing device of any one of claims 1-14, wherein the sequencing is in situ sequencing of a target nucleic acid in a tissue sample.

16. The sequencing device of claim 15, wherein the tissue sample is a tissue slice having a thickness of 20 μm to 200 μm .

17. The sequencing device of any one of claims 1-16, wherein the in situ sequencing is sequential or combinatorial in situ sequencing.

18. The sequencing device of anyone of claims 1-17, wherein the microscope module comprises an epifluorescent microscope, a confocal microscope, a structured illumination microscope, or a light sheet or oblique-plane light sheet microscope.

19. The sequencing device of claim 18, wherein the confocal microscope is a spinning disk or point scanning confocal microscope.

20. A method of using the sequencing device of any one of claims 1-19, the method comprising:

- loading samples into the multi-well plate;
- selecting which samples in the multi-well plate are sequenced;
- selecting a sequencing protocol; and
- sequencing nucleic acids in the selected samples using the sequencing device of any one of claims 1-19.

21. The method of claim 20, wherein the sequencing is in situ volumetric sequencing of tissue samples.

22. The method of claim 20 or 21, wherein the tissue samples are tissue slices having a thickness of 20-200 μm .

23. The method of any one of claims 20-22, wherein the in situ sequencing is sequential or combinatorial in situ sequencing.

24. A computer implemented method, the computer performing steps comprising:

- (a) locating a selected sample in the multi-well plate;
- (b) detecting a signal in the XY plane from the selected sample at low magnification using widefield imaging mode acquisition with camera binning;
- (c) using the signal to segment an XY bounding box around the sample;
- (d) imaging the sample within the XY bounding box to produce an image, wherein imaging is performed in confocal imaging mode in Z at higher magnification than used in step (b) with camera binning in order to determine the approximate Z extent of the sample, wherein a single Z plane is collected through the midpoint of the Z extent previously determined and across the XY extent;

- (e) displaying the image produced in step (d);
- (f) providing an interface for a user to select a desired XY region of interest in the sample to be further imaged during sequencing of the selected sample;
- (g) imaging the sample in the selected XY region of interest across the previously sampled Z extents;
- (h) calculating a sample volume of the region of interest and displaying the calculated sample volume of the region of interest to the user;
- (i) segmenting the image of the sample in the region of interest along Z extents;
- (j) providing an interface to the user for the user to adjust the Z extents of the sample volume before beginning sequencing, wherein the imaging extents derived from the region of interest defined by the user are automatically converted into appropriate montaged fields of view for a given imaging objective and to adjust microscope stage positions, objective Z positioning, and piezo bounds for imaging of the region of interest along XYZ axes during sequencing; and
- (k) reiterating steps (a)-(j) to define regions of interest for each sample in the multi-well plate that the user intends to sequence.

25. A non-transitory computer-readable medium comprising program instructions that, when executed by a processor in a computer, causes the processor to perform the method of claim 24.

26. A computer implemented method, the computer performing steps comprising:
providing an interface to the user for the user to select one or more samples for sequencing and a sequencing protocol, wherein the user is limited in how many samples can be selected depending on amounts of buffer and reagents available and the selected sequencing protocol;
providing constraints on total sequencing time, total data acquired, rate of acquisition, and maximum total volume of regions of interest across all samples that are to be sequenced and imaged, and
suggesting protocols that maximize sequencing of desired regions of interest in samples within the constraints.

27. The computer implemented method of claim 26, wherein the computer is further programmed to optimize sample sequencing parallelization depending on number of samples to be sequenced and imaging types to be used in sequencing.

28. A non-transitory computer-readable medium comprising program instructions that, when executed by a processor in a computer, causes the processor to perform the method of claim 26 or 27.

29. A computer implemented method, the computer performing steps comprising:
performing a rapid confocal sweep in Z at a starting XY position of a given sample montage to determine a Z profile of the sample at the starting XY position;
determining the sample top and bottom interface using a segmentation method; and
setting the objective Z position at a fixed distance from the interface at the beginning of the sample montage, wherein drift in Z of the sample relative to the stage and the objective across rounds is reduced to below a selected tolerance to facilitate downstream subpixel registration across rounds during post-acquisition processing.

30. A non-transitory computer-readable medium comprising program instructions that, when executed by a processor in a computer, causes the processor to perform the method of claim 29.

31. An automated immersion media module comprising:
(a) a container comprising immersion media;
(b) fluidic lines coupled to the container and to the objective immersion collars of the objectives of the microscope module, wherein the fluidic lines carry immersion media to and from an objective immersion collar on an immersion objective, wherein the immersion collar captures excess immersion media; and
(c) a series of pumps connected to the fluidic lines and to a microcontroller, wherein the microcontroller controls the pumps addition and removal of the immersion media through the fluidic lines, wherein the automated immersion media module provides controlled volumes of the immersion media to the objective immersion collars at the tops of the objectives during imaging.

32. A method of using the automated immersion media module of claim 31, the method comprising using the automated immersion media module of claim 31 to deliver immersion media to an objective immersion collar attached to an immersion objective of a microscope.

33. A fluidic management module comprising a symmetrical rotary valve comprising a rotary valve mechanism, a pump, wherein the pump is connected to the fluidic lines, and bubble

detectors, wherein the bubble detectors are positioned on either side of the fluidic lines leading to the pump, wherein the fluidic management module allows bidirectional or unidirectional movement of reagents, buffers, and waste through the fluidic lines.

34. A reagent, buffer, and waste module comprising:

(a) a sliding tray, wherein reagent cartridges and buffer cartridges can be positioned in the sliding tray and coupled to the fluidic management module;

(b) a waste module comprising a waste container, wherein the waste container is coupled to a fluidic line from the fluid management pump; and

(c) a capping mechanism, wherein the capping mechanism closes the waste container when the waste container is removed from the system for waste disposal and opens the waste container when the waste container is placed back into the system.