METHODS AND SYSTEMS FOR PROCESSING POLYNUCLEOTIDES

CROSS-REFERENCE

[0001] This application claims priority to U.S. Provisional Patent Application No. 62/438,341, filed on December 22, 2016, and also claims priority to U.S. Patent Application 15/720,085, filed on September 29, 2017. Each of the above-referenced applications is herein incorporated by reference in its entirety for all purposes.

BACKGROUND

[0002] Significant advances in analyzing and characterizing biological and biochemical materials and systems have led to unprecedented advances in understanding the mechanisms of life, health, disease and treatment. Among these advances, technologies that target and characterize the genomic make up of biological systems have yielded some of the most groundbreaking results, including advances in the use and exploitation of genetic amplification technologies, and nucleic acid sequencing technologies.

[0003] Nucleic acid sequencing can be used to obtain information in a wide variety of biomedical contexts, including diagnostics, prognostics, biotechnology, and forensic biology. Sequencing may involve basic methods including Maxam-Gilbert sequencing and chain-termination methods, or de novo sequencing methods including shotgun sequencing and bridge PCR, or next-generation methods including polony sequencing, 454 pyrosequencing, Illumina sequencing, SOLiD sequencing, Ion Torrent semiconductor sequencing, HeliScope single molecule sequencing, SMRT® sequencing, and others. Nucleic acid sequencing technologies, including next-generation DNA sequencing, have been useful for genomic and proteomic analysis of cell populations.

SUMMARY

[0004] Recognized herein is the need for methods, compositions and systems for analyzing genomic and proteomic information from individual cells or a small population of cells. Such cells include, but are not limited to, cancer cells, fetal cells, and immune cells involved in immune responses. Provided herein are methods, compositions and systems for analyzing individual cells or a small population of cells, including the analysis and attribution of nucleic acids and proteins from and to these individual cells or cell populations.

[0004a] In one aspect, the present disclosure provides a method for analyte characterization, comprising: (a) contacting a plurality of cells with a labelling agent to provide a cell bound to said labelling agent, wherein said cell comprises a plurality of analytes, wherein said labelling agent is configured to couple to a first analyte of said plurality of analytes, and wherein said labelling agent comprises coupled thereto a reporter oligonucleotide molecule comprising a

reporter sequence that permits identification of said labelling agent; (b) providing a plurality of partitions, wherein a partition of said plurality of partitions comprises said cell and a plurality of nucleic acid barcode molecules coupled to a bead and, wherein (i) a first nucleic acid barcode molecule of said plurality of nucleic acid barcode molecules comprises a nucleic acid barcode sequence and a first capture sequence configured to couple to said reporter oligonucleotide molecule, and (ii) a second nucleic acid barcode molecule of said plurality of nucleic acid barcode molecules comprises said nucleic acid barcode sequence and a second capture sequence configured to couple to a second analyte of said plurality of analytes, wherein said first analyte and said second analyte are different types of analytes; and wherein said first capture sequence is different from said second capture sequence; (c) using said first nucleic acid barcode molecule and said reporter oligonucleotide molecule to generate a first barcoded nucleic acid molecule comprising said nucleic acid barcode sequence or complement thereof and a sequence of said reporter oligonucleotide molecule or complement thereof; and (d) using said second nucleic acid barcode molecule and said second analyte to generate a second barcoded nucleic acid molecule comprising said nucleic acid barcode sequence or complement thereof and a sequence of said second analyte or complement thereof.

[0005] In an example, the present disclosure provides a method of characterizing a cell. The method comprises (a) providing a partition comprising a cell and at least one labelling agent, wherein the at least one labelling agent is (i) capable of binding to a cell surface feature of the cell and (ii) is coupled to a reporter oligonucleotide comprising a nucleic acid barcode sequence that permits identification of the at least one labelling agent, wherein the partition comprises an anchor oligonucleotide that is capable of interacting with the reporter oligonucleotide barcode; (b) in the partition, synthesizing a nucleic acid molecule comprising at least a portion of the nucleic acid barcode sequence or a complement thereof; and (c) subjecting the nucleic acid molecule to sequencing to identify the labelling agent or the cell.

[0006] In some examples, in (a), the at least one labelling agent is bound to the cell surface feature. In some examples, prior to (a), the at least one labelling agent is subjected to conditions suitable for binding the at least one labelling agent to the cell surface feature. In some examples, subjecting the at least one labelling agent to the conditions suitable for binding the at least one labelling agent to the cell surface feature is performed when the cell and the at least one labelling agent are free from the partition. In some examples, prior to (a), the at least one labelling agent is coupled to the reporter oligonucleotide.

[0007] In some examples, in (b), the reporter oligonucleotide is subjected to a primer extension reaction that generates the nucleic acid molecule. In some examples, the primer

extension reaction comprises subjecting the reporter oligonucleotide to conditions suitable to hybridize the anchor oligonucleotide to the reporter oligonucleotide and extend the anchor oligonucleotide using the reporter oligonucleotide as a template.

[0008] In some examples, in (b), the anchor oligonucleotide is coupled to a bead. In some examples, in (b), the anchor oligonucleotide is coupled to a bead and the method further comprises releasing the anchor oligonucleotide from the bead prior to the synthesizing. In some examples, the bead is a gel bead. In some examples, the releasing comprises subjecting the bead to a stimulus that degrades the bead. In some examples, the stimulus is a chemical stimulus. In some examples, the bead comprises at least about 1,000 copies of the anchor oligonucleotide. In some examples, the bead comprises at least about 10,000 copies of the anchor oligonucleotide. In some examples, the bead comprises at least about 100,000 copies of the anchor oligonucleotide.

[0009] In some examples, prior to (c), the nucleic acid molecule is released from the partition. In some examples, (c) comprises identifying the at least one labelling agent. In some examples, (c) comprises identifying the cell surface feature from identifying the at least one labelling agent. In some examples, (c) comprises determining an abundance of the given cell surface feature on the cell. In some examples, (c) comprises identifying the cell. In some examples, (c) comprises identifying the at least one labelling agent and the cell.

[0010] In some examples, the reporter oligonucleotide comprises a unique molecular identification (UMI) sequence. In some examples, the UMI sequence permits identification of the cell. In some examples, (c) comprises determining a sequence of the UMI sequence and identifying the cell.

[0011] In some examples, the partition is a droplet in an emulsion. In some examples, the at least one labelling agent is selected from the group comprising of an antibody, or an epitope binding fragment thereof, a cell surface receptor binding molecule, a receptor ligand, a small molecule, a bi-specific antibody, a bi-specific T-cell engager, a T-cell receptor engager, a B-cell receptor engager, a pro-body, an aptamer, a monobody, an affimer, a darpin, a protein scaffold, an antigen, an antigen presenting particle and a major histocompatibility complex (MHC). In some examples, the cell surface feature is selected from the group comprising of a receptor, an antigen, a surface protein, a transmembrane protein, a cluster of differentiation protein, a protein channel, a protein pump, a carrier protein, a phospholipid, a glycoprotein, a glycolipid, a cell-cell interaction protein complex, an antigen-presenting complex, a major histocompatibility complex, an engineered T-cell receptor, a T-cell receptor, a B-cell receptor, a chimeric antigen receptor, a gap junction, and an adherens junction. In some examples, the partition comprises only one cell.

[0012] In some examples, the cell is bound to at least one of the at least one labelling agent. In some examples, the at least one of the at least one labelling agent comprises at least two of the same labelling agent. In some examples, the at least one of the at least one labelling agent comprises at least two different labelling agents. In some examples, the cell is bound to at least about 5 different labelling agents. In some examples, the cell is bound to at least about 10 different labelling agents. In some examples, the cell is bound to at least about 50 different labelling agents. In some examples, the cell is bound to at least about 100 different labelling agents. In some examples, the (c) comprises determining an identity of at least a subset of the different labelling agents.

[0013] In some examples, the method further comprises (i) liberating nucleic acid from the cell and (ii) subjecting the nucleic acid or a derivative thereof to sequencing. In some examples, the nucleic acid is liberated from the cell into the partition.

[0014] In an example, the present disclosure provides a system for characterizing a cell. The system comprises an electronic display screen comprising a user interface that displays a graphical element that is accessible by a user to execute a protocol to characterize the cell; and a computer processor coupled to the electronic display screen and programmed to execute the protocol upon selection of the graphical element by the user, which protocol comprises: (a) providing a partition comprising a cell and at least one labelling agent, wherein the at least one labelling agent is (i) capable of binding to a cell surface feature of the cell and (ii) is coupled to a reporter oligonucleotide comprising a nucleic acid barcode sequence that permits identification of the at least one labelling agent, wherein the partition comprises an anchor oligonucleotide that is capable of interacting with the reporter oligonucleotide barcode; (b) in the partition, synthesizing a nucleic acid molecule comprising at least a portion of the nucleic acid barcode sequence or a complement thereof; and (c) subjecting the nucleic acid molecule to sequencing to identify the labelling agent or the cell.

[0015] In some examples, in protocol (a), the at least one labelling agent is bound to the cell surface feature. In some examples, prior to protocol (a), the at least one labelling agent is subjected to conditions suitable for binding the at least one labelling agent to the cell surface feature. In some examples, subjecting the at least one labelling agent to the conditions suitable for binding the at least one labelling agent to the cell surface feature is performed when the cell and the at least one labelling agent are free from the partition. In some examples, prior to protocol (a), the at least one labelling agent is coupled to the reporter oligonucleotide.

[0016] In some examples, in protocol (b), the reporter oligonucleotide is subjected to a primer extension reaction that generates the nucleic acid molecule. In some examples, the

primer extension reaction comprises subjecting the reporter oligonucleotide to conditions suitable to hybridize the anchor oligonucleotide to the reporter oligonucleotide and extend the anchor oligonucleotide using the reporter oligonucleotide as a template.

[0017] In some examples, in protocol (b), the anchor oligonucleotide is coupled to a bead. In some examples, in (b), the anchor oligonucleotide is coupled to a bead and the method further comprises releasing the anchor oligonucleotide from the bead prior to the synthesizing. In some examples, the bead is a gel bead. In some examples, the releasing comprises subjecting the bead to a stimulus that degrades the bead. In some examples, the stimulus is a chemical stimulus. In some examples, the bead comprises at least about 1,000 copies of the anchor oligonucleotide. In some examples, the bead comprises at least about 10,000 copies of the anchor oligonucleotide. In some examples, the bead comprises at least about 100,000 copies of the anchor oligonucleotide.

[0018] In some examples, prior to protocol (c), the nucleic acid molecule is released from the partition. In some examples, protocol (c) comprises identifying the at least one labelling agent. In some examples, protocol (c) comprises identifying the cell surface feature from identifying the at least one labelling agent. In some examples, protocol (c) comprises determining an abundance of the given cell surface feature on the cell. In some examples, protocol (c) comprises identifying the cell. In some examples, protocol (c) comprises identifying the at least one labelling agent and the cell.

[0019] In some examples, the reporter oligonucleotide comprises a unique molecular identification (UMI) sequence. In some examples, the UMI sequence permits identification of the cell. In some examples, protocol (c) comprises determining a sequence of the UMI sequence and identifying the cell.

[0020] In some examples, the partition is a droplet in an emulsion. In some examples, the at least one labelling agent is selected from the group comprising of an antibody, or an epitope binding fragment thereof, a cell surface receptor binding molecule, a receptor ligand, a small molecule, a bi-specific antibody, a bi-specific T-cell engager, a T-cell receptor engager, a B-cell receptor engager, a pro-body, an aptamer, a monobody, an affimer, a darpin, a protein scaffold, an antigen, an antigen presenting particle and a major histocompatibility complex (MHC). In some examples, the cell surface feature is selected from the group comprising of a receptor, an antigen, a surface protein, a transmembrane protein, a cluster of differentiation protein, a protein channel, a protein pump, a carrier protein, a phospholipid, a glycoprotein, a glycolipid, a cell-cell interaction protein complex, an antigen-presenting complex, a major histocompatibility complex,

an engineered T-cell receptor, a T-cell receptor, a B-cell receptor, a chimeric antigen receptor, a gap junction, and an adherens junction. In some examples, the partition comprises only one cell.

[0021] In some examples, the cell is bound to at least one of the at least one labelling agent. In some examples, the at least one of the at least one labelling agent comprises at least two of the same labelling agent. In some examples, the at least one of the at least one labelling agent comprises at least two different labelling agents. In some examples, the cell is bound to at least about 5 different labelling agents. In some examples, the cell is bound to at least about 10 different labelling agents. In some examples, the cell is bound to at least about 50 different labelling agents. In some examples, the cell is bound to at least about 100 different labelling agents. In some examples, protocol (c) comprises determining an identity of at least a subset of the different labelling agents.

[0022] In some examples, protocol comprises (i) liberating nucleic acid from the cell and (ii) subjecting the nucleic acid or a derivative thereof to sequencing. In some examples, the nucleic acid is liberated from the cell into the partition.

[0023] An additional example of the disclosure provides a method for analyte characterization. The method includes: (a) providing a plurality of partitions, where a given partition of the plurality of partitions comprises a plurality of barcode molecules and a plurality of analytes. In some cases, the plurality of barcode molecules comprises at least 1,000 barcode molecules. In addition, (i) a first individual barcode molecule of the plurality of barcode molecules can comprise a first nucleic acid barcode sequence that is capable of coupling to a first analyte of the plurality of analytes, and (ii) a second individual barcode molecule of the plurality of barcoded molecules can comprise a second nucleic acid barcode sequence that is capable of coupling to a second analyte of the plurality of analytes where the first analyte and the second analyte are different types of analytes. The method also includes (b) in the given partition, (i) synthesizing a first nucleic acid molecule comprising at least a portion of the first nucleic acid barcode sequence or complement thereof, and (ii) synthesizing a second nucleic acid molecule comprising at least a portion of the second nucleic acid barcode sequence or complement thereof; and (c) removing the first nucleic acid molecule and the second nucleic acid molecule from the given partition.

[0024] In some examples, the method further comprises subjecting the first nucleic acid molecule and the second nucleic acid molecule, or a derivative of the first nucleic acid molecule and/or the second nucleic acid molecule, to sequencing to characterize the first analyte and/or the second analyte. In some examples, the method further comprises repeating (a)-(c) based on a characterization of the first analyte or the second analyte from the sequencing. In some

examples, the method further comprises selecting the first analyte or the second analyte based on a characterization of the first analyte or the second analyte obtained from the sequencing or a subsequent sequencing upon repeating (a)-(c).

[0025] In some examples, (b) further comprises: (1) synthesizing the first nucleic acid molecule comprising at least a portion of the first nucleic acid barcode sequence or complement thereof, and (2) synthesizing the second nucleic acid molecule comprising at least a portion of the second nucleic acid barcode sequence or complement thereof.

[0026] In some examples, the first analyte is a nucleic acid molecule, such as genomic deoxyribonucleic acid (gDNA) or messenger RNA (mRNA).

[0027] In some examples, the first analyte is a labelling agent capable of coupling to a cell surface feature of a cell. In some examples, the first individual barcode molecule or the second individual barcode molecule is capable of coupling to the labelling agent via a third nucleic acid molecule coupled to the labelling agent. In some examples, the cell surface feature is a receptor, an antigen, or a protein. In some examples, the labelling agent is an antibody, an antibody fragment or a major histocompatibility complex (MHC). In some examples, the given partition comprises the cell or one or more components of the cell. In some examples, the given partition comprises a single cell. In some examples, the first nucleic acid molecule or the second nucleic molecule comprises a third barcode sequence. In some examples, the third barcode sequence is derived from a third nucleic acid molecule. In some examples, the third nucleic acid molecule is coupled to a labelling agent capable of binding to a cell surface feature of a cell.

[0028] In some examples, the first analyte and second analyte are different types of nucleic acid molecules. In some examples, the first analyte is a ribonucleic acid molecule and the second analyte is a deoxyribonucleic acid molecule. In some examples, (i) the first individual barcode molecule comprises a first priming sequence capable of hybridizing to the first analyte; or (ii) the second individual barcode molecule comprises a second priming sequence capable of hybridizing to the second analyte. In some examples, the first barcode molecule or the second barcode molecule comprises a unique molecular identification (UMI) sequence.

[0029] In some examples, the first analyte is a nucleic acid molecule and the second analyte is a labelling agent capable of coupling to a cell surface feature. In some examples, the first analyte is a messenger ribonucleic acid molecule. In some examples, (i) the first individual barcode molecule comprises a first priming sequence capable of hybridizing to the first analyte; or (ii) the second individual barcode molecule comprises a second priming sequence capable of hybridizing to the labelling agent. In some examples, the labelling agent is an antibody, or an epitope binding fragment thereof, or a major histocompatibility complex (MHC). In some

examples, the cell surface feature is selected from the group consisting of a receptor, an antigen, or a protein.

[0030] In some examples, the first analyte comprises a nucleic acid molecule with a nucleic acid sequence or complement thereof that encodes at least a portion of a V(D)J sequence of an immune cell receptor. In some examples, the nucleic acid molecule is a messenger ribonucleic acid. In some examples, the nucleic acid molecule is complementary DNA (cDNA) derived from reverse transcription of an mRNA encoding the at least a portion of the V(D)J sequence.

[0031] In some examples, the first analyte comprises a nucleic acid molecule with a nucleic acid sequence that is capable of functioning as a component of a gene editing reaction. In some examples, the gene editing reaction comprises clustered regularly interspaced short palindromic repeats (CRISPR)-based gene editing.

[0032] In some examples, at least one of the first individual barcode molecule and the second individual barcode molecule is coupled to a bead, such as a gel bead. The bead can be degradable. In some examples, the method further comprises, after (a), releasing the first individual barcode molecule or the second individual barcode from the bead. In some examples, the given partition further comprises an agent capable of releasing the first individual barcode molecule or the second individual barcode from the bead.

[0033] In some examples, the given partition selected is a droplet among a plurality of droplets or a well among a plurality of wells. In some examples, the first nucleic acid barcode sequence and the second nucleic barcode sequence are identical. In some examples, the method further comprises performing one or more reactions subsequent to removing the first nucleic acid molecule and the second nucleic acid molecule from the given partition.

[0034] Another example of the disclosure provides a composition for characterizing a plurality of analytes. The composition comprises a partition comprising a plurality of barcode molecules and the plurality of analytes. The plurality of barcode molecules can comprise at least 1,000 barcode molecules. In addition, (i) a first individual barcode molecule of the plurality of barcode molecules can comprise a first nucleic acid barcode sequence that is capable of coupling to a first analyte of the plurality of analytes; and (ii) a second individual barcode molecule of the plurality of barcoded molecules can comprise a second nucleic acid barcode sequence that is capable of coupling to a second analyte of the plurality of analytes, where the first analyte and the second analyte are different types of analytes.

[0035] In some examples, the first analyte is a nucleic acid molecule, such as genomic deoxyribonucleic acid (gDNA) or is messenger RNA (mRNA).

[0036] In some examples, the first analyte is a labelling agent capable of coupling to a cell surface feature of a cell. In some examples, the first individual barcode molecule or the second individual barcode molecule is capable of coupling to the labelling agent via a third nucleic acid molecule coupled to the labelling agent. In some examples, the cell surface feature is a receptor, an antigen, or a protein. In some examples, the labelling agent is an antibody, or an epitope binding fragment thereof, or a major histocompatibility complex (MHC). In some examples, the partition comprises the cell or one or more components of the cell. In some examples, the partition comprises a single cell. In some examples, the first nucleic acid molecule or the second nucleic molecule comprises a third barcode sequence. In some examples, the third barcode sequence is derived from a third nucleic acid molecule. In some examples, the third nucleic acid molecule is coupled to a labelling agent capable of binding to a cell surface feature of a cell.

[0037] In some examples, the first analyte and second analyte are different types nucleic acid molecules. In some examples, the first analyte is a ribonucleic acid molecule and the second analyte is a deoxyribonucleic acid molecule. In some examples, (i) the first individual barcode molecule comprises a first priming sequence capable of hybridizing to the first analyte; or (ii) the second individual barcode molecule comprises a second priming sequence capable of hybridizing to the second analyte. In some examples, the first barcode molecule or the second barcode molecule comprises a unique molecular identification (UMI) sequence.

[0038] In some examples, the first analyte is a nucleic acid molecule and the second analyte is a labelling agent capable of coupling to a cell surface feature. In some examples, the first analyte is a ribonucleic acid molecule. In some examples, (i) the first individual barcode molecule comprises a first priming sequence capable of hybridizing to the first analyte; or (ii) the second individual barcode molecule comprises a second priming sequence capable of hybridizing to a third nucleic acid molecule coupled to the labelling agent. In some examples, the labelling agent is an antibody, or an epitope binding fragment thereof, or a major histocompatibility complex (MHC). In some examples, the cell surface feature is a receptor, an antigen, or a protein.

[0039] In some examples, the first analyte comprises a nucleic acid molecule with a nucleic acid sequence or complement thereof that encodes at least a portion of a V(D)J sequence of an immune cell receptor. In some examples, the nucleic acid sequence is a ribonucleic acid molecule. In some examples, the nucleic acid molecule comprises a nucleic acid sequence that is complementary DNA (cDNA) derived from reverse transcription of an mRNA encoding the at least a portion of the V(D)J sequence.

[0040] In some examples, the first analyte comprises a nucleic acid molecule with a nucleic acid sequence that is capable of functioning as a component of a gene editing reaction. In some examples, the gene editing reaction comprises clustered regularly interspaced short palindromic repeats (CRISPR)-based gene editing. In some examples, at least one of the first individual barcode molecule and the second individual barcode molecule is coupled to a bead, such as a gel bead. The bead may be degradable. In some examples, the given partition further comprises an agent capable of releasing the first individual barcode molecule or the second individual barcode from the bead. In some examples, the given partition is a droplet among a plurality of droplets or a well among a plurality of wells. In some examples, the first nucleic acid barcode sequence and the second nucleic barcode sequence are identical.

[0041] An additional example of the disclosure provides a system for characterizing a plurality of analytes. The system comprises a partitioning unit for providing a partition comprising a plurality of barcode molecules and the plurality of analytes, where: (i) a first individual barcode molecule of the plurality of barcode molecules comprises a first nucleic acid barcode sequence and is capable of coupling to a first analyte of the plurality of analytes; and (ii) a second individual barcode molecule of the plurality of barcode molecules comprises a second nucleic acid barcode sequence and is capable of coupling to a second analyte of the plurality of analytes, where the first analyte and the second analyte are different types of analytes. The system also includes a controller coupled to the partitioning unit, where the controller is programmed to (i) direct the partitioning unit to provide the partition; (ii) subject the partition to conditions that are sufficient to: (1) synthesize a first nucleic acid molecule comprising at least a portion of the first nucleic acid barcode sequence or complement thereof; and (2) synthesize a second nucleic acid molecule comprising at least a portion of the second nucleic acid barcode sequence or complement thereof, where sequencing of the first nucleic acid molecule and the second nucleic acid molecule, or derivatives thereof, characterizes the first analyte or the second analyte.

[0042] In some examples, the partitioning unit comprises a plurality of channels. In some examples, the partitioning unit further comprises at least one channel junction, where the at least one channel junction is configured to provide the partition. In some examples, the system also includes (i) a first channel fluidically connected to the at least one channel junction and configured to provide a first fluid to the at least one channel junction; (ii) and a second channel fluidically connected to the at least one channel junction and configured to provide a second fluid, immiscible with the first fluid, to the at least one channel junction.

[0043] In some examples, the first analyte is a nucleic acid molecule, such as genomic deoxyribonucleic acid (gDNA) or messenger RNA (mRNA).

[0044] In some examples, the first analyte is a labelling agent capable of coupling to a cell surface feature of a cell. In some examples, the first individual barcode molecule or the second individual barcode molecule is capable of coupling to the labelling agent via a third nucleic acid molecule coupled to the labelling agent. In some examples, the cell surface feature is a receptor, an antigen, or a protein. In some examples, the labelling agent is an antibody, or an epitope binding fragment thereof, or a major histocompatibility complex (MHC). In some examples, the partition comprises the cell or one or more components of the cell. In some examples, the partition comprises a single cell. In some examples, the first nucleic acid molecule or the second nucleic molecule comprises a third barcode sequence. In some examples, the third barcode sequence is derived from a third nucleic acid molecule. In some examples, the third nucleic acid molecule is coupled to a labelling agent capable of binding to a cell surface feature of a cell.

[0045] In some examples, the first analyte and second analyte are different types nucleic acid molecules. In some examples, the first analyte is a ribonucleic acid molecule and the second analyte is a deoxyribonucleic acid molecule. In some examples, (i) the first individual barcode molecule comprises a first priming sequence capable of hybridizing to the first analyte; or (ii) the second individual barcode molecule comprises a second priming sequence capable of hybridizing to the second analyte. In some examples, the first barcode molecule or the second barcode molecule comprises a unique molecular identification (UMI) sequence.

[0046] In some examples, the first analyte is a nucleic acid molecule and the second analyte is a labelling agent capable of coupling to a cell surface feature. In some examples, the first analyte is a ribonucleic acid molecule. In some examples, (i) the first individual barcode molecule comprises a first priming sequence capable of hybridizing to the first analyte; or (ii) the second individual barcode molecule comprises a second priming sequence capable of hybridizing to a third nucleic acid molecule coupled to the labelling agent. In some examples, the labelling agent is an antibody, or an epitope binding fragment thereof, or a major histocompatibility complex (MHC). In some examples, the cell surface feature is a receptor, an antigen, or a protein.

[0047] In some examples, the first analyte comprises a nucleic acid molecule with a nucleic acid sequence or complement thereof that encodes at least a portion of a V(D)J sequence of an immune cell receptor. In some examples, the nucleic acid sequence is a messenger ribonucleic acid molecule. In some examples, the nucleic acid molecule is complementary DNA (cDNA)

derived from reverse transcription of an mRNA encoding the at least a portion of the V(D)J sequence.

[0048] In some examples, the first analyte comprises a nucleic acid molecule with a nucleic acid sequence that is capable of functioning as a component of a gene editing reaction. In some examples, the gene editing reaction comprises clustered regularly interspaced short palindromic repeats (CRISPR)-based gene editing.

[0049] In some examples, at least one of the first individual barcode molecule and the second individual barcode molecule is coupled to a bead, such as a gel bead. The bead may be degradable. In some examples, the partition further comprises an agent capable of releasing the first individual barcode molecule or the second individual barcode from the bead. In some examples, the partition is a droplet among a plurality of droplets or a well among a plurality of wells. In some examples, the nucleic acid barcode sequence and the second nucleic barcode sequence are identical. In some examples, the partition comprises at least 1,000 barcode molecules.

[0050] Additional examples and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative examples of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different examples, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0051] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0052] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings (also "Figure" and "FIG." herein), of which:

[0053] FIG. 1 schematically illustrates a microfluidic channel structure for partitioning individual or small groups of cells;

[0054] FIG. 2 schematically illustrates a microfluidic channel structure for co-partitioning cells and microcapsules (*e.g.*, beads) comprising additional reagents;

[0055] FIGS. 3A-3F schematically illustrate an example process for amplification and barcoding of cell's nucleic acids;

[0056] FIG. 4 provides a schematic illustration of use of barcoding of cell's nucleic acids in attributing sequence data to individual cells or groups of cells for use in their characterization;

[0057] FIG. 5 provides a schematic illustration of cells associated with labeled cell-binding ligands;

[0058] FIG. 6 provides a schematic illustration of an example workflow for performing RNA analysis using the methods described herein;

[0059] FIG. 7 provides a schematic illustration of an example barcoded oligonucleotide structure for use in analysis of ribonucleic (RNA) using the methods described herein;

[0060] FIG. 8 provides an image of individual cells co-partitioned along with individual barcode bearing beads;

[0061] FIGS. 9A-9E provide schematic illustrations of example barcoded oligonucleotide structures for use in analysis of RNA and example operations for performing RNA analysis;

[0062] FIG. 10 provides a schematic illustration of example barcoded oligonucleotide structure for use in example analysis of RNA and use of a sequence for in vitro transcription;

CLAIMS

WHAT IS CLAIMED IS

1. A method for analyte characterization, comprising:

(a) contacting a plurality of cells with a labelling agent to provide a cell bound to said labelling agent, wherein said cell comprises a plurality of analytes, wherein said labelling agent is configured to couple to a first analyte of said plurality of analytes, and wherein said labelling agent comprises coupled thereto a reporter oligonucleotide molecule comprising a reporter sequence that permits identification of said labelling agent;

(b) providing a plurality of partitions, wherein a partition of said plurality of partitions comprises said cell and a plurality of nucleic acid barcode molecules coupled to a bead and, wherein (i) a first nucleic acid barcode molecule of said plurality of nucleic acid barcode molecules comprises a nucleic acid barcode sequence and a first capture sequence configured to couple to said reporter oligonucleotide molecule, and (ii) a second nucleic acid barcode molecule of said plurality of nucleic acid barcode molecules comprises said nucleic acid barcode sequence and a second capture sequence configured to couple to a second analyte of said plurality of analytes, wherein said first analyte and said second analyte are different types of analytes; and wherein said first capture sequence is different from said second capture sequence;

(c) using said first nucleic acid barcode molecule and said reporter oligonucleotide molecule to generate a first barcoded nucleic acid molecule comprising said nucleic acid barcode sequence or complement thereof and a sequence of said reporter oligonucleotide molecule or complement thereof; and

(d) using said second nucleic acid barcode molecule and said second analyte to generate a second barcoded nucleic acid molecule comprising said nucleic acid barcode sequence or complement thereof and a sequence of said second analyte or complement thereof.

2. The method of claim 1, wherein (c) and (d) are performed in said partition.

3. The method of claim 1, further comprising subjecting said first barcoded nucleic acid molecule or a derivative thereof and said barcoded second nucleic acid molecule or a derivative thereof to sequencing to generate sequencing information corresponding to said first barcoded nucleic acid molecule and said second barcoded nucleic acid molecule.

4. The method of claim 3, further comprising using said sequencing information to identify said barcode sequence and associate said first analyte and said second analyte with said cell.

5. The method of claim 1, wherein said first nucleic acid barcode molecule or said second nucleic acid barcode molecule comprises one or more functional sequences selected from the group consisting of an adapter sequence, a primer sequence, a primer binding sequence, a unique molecular identification (UMI) sequence, and a sequence configured to couple to a flow cell of a sequencer.

6. The method of claim 1, wherein said second analyte is selected from the group consisting of a deoxyribonucleic acid (DNA) molecule and a ribonucleic acid (RNA) molecule.

7. The method of claim 6, wherein said RNA molecule is a messenger RNA (mRNA), a clustered regularly interspaced short palindromic repeat (CRISPR) RNA (crRNA), or a single guide RNA (sgRNA).

8. The method of claim 1, wherein said first analyte is a protein.

9. The method of claim 1, wherein said first analyte is a cell surface feature of said cell.

10. The method of claim 9, wherein said cell surface feature of said cell is selected from the group consisting of a receptor, an antigen, a surface protein, a transmembrane protein, a cluster of differentiation protein, a protein channel, a protein pump, a carrier protein, a phospholipid, a glycoprotein, a glycolipid, a cell-cell interaction protein complex, an antigen-presenting complex, a major histocompatibility complex, a T-cell receptor, a B-cell receptor, a chimeric antigen receptor, a gap junction, and an adherens junction.

11. The method of claim 1, wherein said labelling agent is selected from the group comprising of an antibody or an epitope binding fragment thereof, a cell surface receptor binding molecule, a receptor ligand, a small molecule, a bi-specific antibody, a bi-specific T-cell engager, a T-cell receptor engager, a B-cell receptor engager, a pro-body, an aptamer, a monobody, an affimer, a darpin, a protein scaffold, an antigen, an antigen presenting particle, and a major histocompatibility complex (MHC) molecule.

12. The method of claim 1, wherein said bead is a gel bead.

13. The method of claim 12, wherein said gel bead is degradable upon application of a stimulus.

14. The method of claim 13, wherein said stimulus is selected from a photo-stimulus, a thermal stimulus, and chemical stimulus, and further comprising, after (b), using said stimulus to degrade said gel bead.

15. The method of claim 1, wherein said first nucleic acid barcode molecule and said second nucleic acid barcode molecule are releasable from said bead upon application of a stimulus.

16. The method of claim 15, wherein said stimulus is selected from a photo-stimulus, a thermal stimulus, and chemical stimulus, and further comprising, after (b), using said stimulus to release said first nucleic acid barcode molecule and said second nucleic acid barcode molecule from said bead.

17. The method of claim 1, wherein said partition is a droplet among a plurality of droplets or a well among a plurality of wells.

18. The method of claim 1, wherein said second capture sequence is a poly-T sequence, a random N-mer sequence, or a targeted priming sequence.

19. The method of claim 1, wherein said reporter oligonucleotide molecule comprises an adapter sequence complementary to said first capture sequence, wherein (c) comprises hybridizing said adapter sequence to said first capture sequence and performing a nucleic acid extension reaction or a ligation reaction to generate said first barcoded nucleic acid molecule.

20. The method of claim 19, wherein said second capture sequence is complementary to a sequence of a ribonucleic acid (RNA) molecule of said cell and wherein (d) comprises hybridizing said RNA molecule to said second capture sequence and performing a reverse transcription reaction to generate said second barcoded nucleic acid molecule.