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 immune cells with a labelling agent to provide an immune cell bound to said labelling agent, wherein said immune cell comprises a plurality of analytes, wherein said labelling agent is configured to couple to a first analyte of said plurality of analytes, wherein said labelling agent comprises coupled thereto a reporter oligonucleotide molecule comprising a

reporter sequence that permits identification of said labelling agent, and wherein said first analyte is a cell surface feature of said immune cell; (b) providing a plurality of partitions, wherein a partition of said plurality of partitions comprises said immune 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 comprising a gene specific sequence which targets a gene or a region of a gene of an immune receptor 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) in said partition, 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) in said partition, 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

CLAIMS

WHAT IS CLAIMED IS

1. A method for analyte characterization, comprising:

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

(b) providing a plurality of partitions, wherein a partition of said plurality of partitions comprises said immune 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 comprising a gene specific sequence which targets a gene or a region of a gene of an immune receptor 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) in said partition, 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) in said partition, 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 an amplification product thereof and said barcoded second nucleic acid molecule or an

amplification product 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 messenger ribonucleic acid (mRNA) molecule.

7. The method of claim 1, wherein said cell surface feature is a protein.

8. The method of claim 1, wherein said cell surface feature 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.

9. 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.

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

11. The method of claim 10, wherein said gel bead is degradable upon application of a stimulus.

12. The method of claim 11, 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.

13. 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.

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 release said first nucleic acid barcode molecule and said second nucleic acid barcode molecule from said bead.

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

16. 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.

17. The method of claim 16, 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.

18. The method of claim 1, wherein the genes or regions of genes of immune receptors are selected from T cell receptor alpha variable genes (TRAVgenes), T cell receptor alpha joining genes (TRAJ genes), T cell receptor alpha constant genes (TRAC genes), T cell receptor beta variable genes (TRBV genes), T cell receptor beta diversity genes (TRBD genes), T cell receptor beta joining genes (TRBJ genes), T cell receptor beta constant genes (TRBC genes), T cell receptor gamma variable genes (TRGV genes), T cell receptor gamma joining genes (TRGJ genes), T cell receptor gamma constant genes (TRGC genes), T cell receptor delta variable genes (TRDV genes), T cell receptor delta diversity genes (TRDD genes), T cell receptor delta joining genes (TRDJ genes), T cell receptor delta constant genes (TRDC genes), B cell receptor heavy chain, and B cell receptor light chain.

19. The method of claim 1 wherein the immune cell is:

(i) a B cell;

(ii) a T cell, such as a cytotoxic T cell, a natural killer T cell, a regulatory T cell, or a T helper cell;

(iii) a natural killer cell;

(iv) a cytokine-induced killer cell;

(v) a myeloid cell, such as a granulocyte, for example a basophil granulocytes, a eosinophil granulocyte, a neutrophil granulocyte or a hypersegmented neutrophil;

(vi) a monocyte or a macrophage;

(vii) a mast cell;

(viii) a thrombocyte or a megakaryocyte; or

(ix) a dendritic cell.