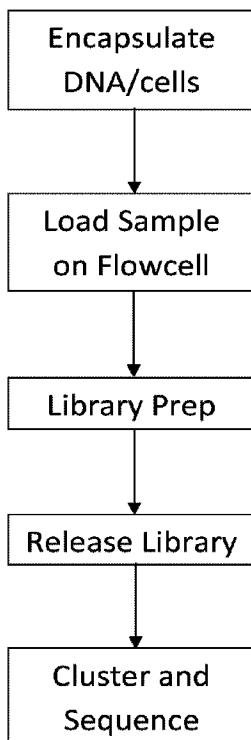




(86) Date de dépôt PCT/PCT Filing Date: 2019/02/11
(87) Date publication PCT/PCT Publication Date: 2019/08/22
(45) Date de délivrance/Issue Date: 2023/03/21
(85) Entrée phase nationale/National Entry: 2019/12/11
(86) N° demande PCT/PCT Application No.: US 2019/017521
(87) N° publication PCT/PCT Publication No.: 2019/160820
(30) Priorité/Priority: 2018/02/13 (US62/630,175)

(51) Cl.Int./Int.Cl. *C12Q 1/6834* (2018.01)
(72) Inventeurs/Inventors:
WU, YIR-SHYUAN, US;
GORPE-YASAR, FILIZ, US;
KHURANA, TARUN KUMAR, US;
POPIC, VICTORIA, US;
JAEGER, ERICH B., US;
RONAGHI, MOSTAFA, US
(73) Propriétaire/Owner:
ILLUMINA, INC., US
(74) Agent: BERESKIN & PARR LLP/S.E.N.C.R.L., S.R.L.

(54) Titre : SEQUENCAGE D'ADN A L'AIDE DE BILLES D'HYDROGEL
(54) Title: DNA SEQUENCING USING HYDROGEL BEADS



(57) Abrégé/Abstract:

Systems, methods, and compositions provided herein relate to preparation of beads encapsulating long DNA fragments for high-throughput spatial indexing. Some embodiments include preparation of nucleic acid libraries within the bead, wherein the bead includes pores that allow diffusion of reagents while retaining genetic material.

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
22 August 2019 (22.08.2019)



(10) International Publication Number
WO 2019/160820 A1

(51) International Patent Classification:
C12Q 1/6834 (2018.01)

(21) International Application Number:
PCT/US2019/017521

(22) International Filing Date:
11 February 2019 (11.02.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/630,175 13 February 2018 (13.02.2018) US

(71) Applicant: **ILLUMINA, INC.** [US/US]; 5200 Illumina Way, San Diego, California 92122 (US).

(72) Inventors: **WU, Yir-Shyuan**; c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US). **GORPE-YASAR, Filiz**; c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US). **KHURANA, Tarun Kumar**; c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US). **POPIC, Victoria**; c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US). **JAEGGER, Erich B.**; c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US). **RONAGHI, Mostafa**;

c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US).

(74) Agent: **FULLER, Michael L.**; 2040 Main Street, Fourteenth Floor, Irvine, California 92614 (US).

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

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

(54) Title: DNA SEQUENCING USING HYDROGEL BEADS

(57) Abstract: Systems, methods, and compositions provided herein relate to preparation of beads encapsulating long DNA fragments for high-throughput spatial indexing. Some embodiments include preparation of nucleic acid libraries within the bead, wherein the bead includes pores that allow diffusion of reagents while retaining genetic material.

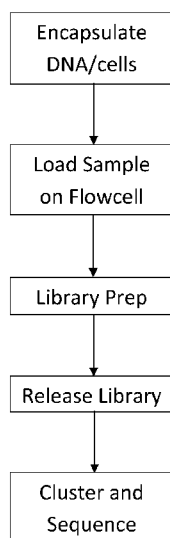


FIG. 2

WO 2019/160820 A1

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

DNA SEQUENCING USING HYDROGEL BEADS

FIELD

[0001] Systems, methods, and compositions provided herein relate to hydrogel beads, and methods of encapsulating long DNA within hydrogel beads, for use in determining the sequence of polynucleotides and related library preparation.

BACKGROUND

[0002] Next generation sequencers are powerful tools that generate large amounts of genomic data per sequencing run. Interpreting and analyzing this large amount of data can be challenging. Sequencing by synthesis (SBS) technology provides high quality sequencing data. However short reads (maximum read length is 2×300 bp) are one limitation of current SBS chemistry. Recently, there is increased emphasis on sequencing longer DNA molecules in order to better capture single nucleotide variants (SNP), insertion/deletion, and structural variants, and for improved genomic identification.

SUMMARY

[0003] Some embodiments provided herein relate to a hydrogel bead for performing DNA reactions. In some embodiments, the hydrogel bead includes a hydrogel polymer precursor, a crosslinker, and DNA disposed within the hydrogel bead. In some embodiments, the bead includes pores that allow diffusion of a reagent through the bead while retaining the DNA. In some embodiments, the DNA is a long DNA molecule of 50,000 base pairs or greater.

[0004] Some embodiments provided herein relate to a flow cell device for performing DNA sequencing. In some embodiments, the flow cell device includes a solid support. In some embodiments, the solid support includes a surface having a degradable hydrogel encapsulating DNA deposited thereon. In some embodiments, the degradable hydrogel includes pores that are sized to allow diffusion of a reagent through the hydrogel, but are too small to allow DNA to traverse the pores.

[0005] Some embodiments provided herein relate to a system for DNA sequencing. In some embodiments, the system includes a stage configured to hold a flow cell device, a flow cell device, and a detector for obtaining sequencing data.

[0006] Some embodiments provided herein relate to a method of sequencing DNA. In some embodiments, the method includes obtaining a bead encapsulating DNA as described herein. In some embodiments, the method includes providing a flow cell device described herein. In some embodiments, the method further includes amplifying DNA encapsulated within the hydrogel, performing a tagmentation reaction on the DNA encapsulated within the hydrogel, or sequencing the DNA. In some embodiments, the method further includes generating a DNA library encapsulated within the hydrogel.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. 1A is a schematic that illustrates an embodiment for spatial indexing of long DNA by on-flow cell library preparation and seeding.

[0008] FIG. 1B is a schematic that illustrates spatial indexing using hydrogel beads that encapsulate long DNA molecules. Reagents may be used on the hydrogel beads to spatially generate a library on a flow cell surface.

[0009] FIG. 2 is a flow diagram that depicts a method of encapsulating long DNA within a hydrogel bead, and preparing a library within the hydrogel bead, which can be clustered and sequenced on a flow cell device.

[0010] FIG. 3 is a schematic that illustrates workflow of DNA sequencing of long DNA encapsulated within hydrogel beads, including DNA fragments of about 100 kb (without a multiple displacement amplification (MDA) step (panel (a)) or with an MDA step (panel (b)) prior to tagmentation) and DNA fragments of about 10-20 kb (panel (c)).

[0011] FIG. 4 is a graph that depicts strobed reads of long DNA hydrogel spatial indexing sequencing data from a 100 kb DNA fragment without MDA.

[0012] FIG. 5 shows a line graph of linked reads of long DNA hydrogel spatial indexing on 100 kb DNA fragments with MDA.

[0013] FIG. 6 shows a line graph of linked reads of long DNA hydrogel spatial indexing on 10 kb DNA fragments with MDA.

[0014] FIGs. 7A and 7B depict line graphs of spatial reads for long DNA encapsulated within a hydrogel bead. FIG. 7A shows the spatial reads for cells encapsulated within a hydrogel bead, and the inset depicts a micrograph showing a cell within the hydrogel bead. FIG. 7B shows the spatial reads for long DNA fragments encapsulated within a hydrogel bead, and the inset depicts a micrograph showing the fragments encapsulated within the beads.

[0015] FIG. 8 depicts a micrograph showing identification of microbial species encapsulated within a hydrogel bead. The hydrogel bead encapsulated various microbial species, and spatial sequencing reads were performed to identify the microbes.

[0016] FIG. 9 illustrate a graph showing the distribution of barcode reads for long DNA encapsulating within hydrogel beads.

[0017] FIG. 10A illustrates a graph showing short reads and linked reads from a single run for an *E. coli* cell encapsulated within a hydrogel bead. As shown in the figure, linked reads span across repeat regions, and can improve de novo sequence assembly. FIG. 10B shows a micrograph depicting spatial linked reads and interstitial short reads.

DETAILED DESCRIPTION

[0018] In the following detailed description, reference is made to the accompanying drawings, which form a part hereof. In the drawings, similar symbols typically identify similar components, unless context dictates otherwise. The illustrative embodiments described in the detailed description, drawings, and claims are not meant to be limiting. Other embodiments may be utilized, and other changes may be made, without departing from the spirit or scope of the subject matter presented herein. It will be readily understood that the aspects of the present disclosure, as generally described herein, and illustrated in the Figures, can be arranged, substituted, combined, separated, and designed in a wide variety of different configurations, all of which are explicitly contemplated herein.

[0019] Embodiments relate to compositions, systems, and methods for encapsulating long DNA fragments into beads to determine the nucleotide sequence of the DNA fragments. This allows creation of reliable and high-throughput methods of sequencing relatively long DNA fragments, as described below. The methods and systems described herein relate to sequencing of long DNA fragments encapsulated in a single bead, enabling

improved sequencing and identification of genomic DNA. In some embodiments, the method includes encapsulating a sample of DNA fragments within a hydrogel bead, loading the hydrogel beads encapsulating the sample of DNA fragments on a flow cell device, preparing a library, releasing the prepared library on a surface of the flow cell device, and clustering and sequencing the released library.

[0020] In some embodiments, preparing a library includes tagmentation of the encapsulated DNA. Tagmentation of the encapsulated DNA cleaves longer DNA sequences into shorter tagmentation fragments which are then used to generate clusters of DNA on a surface of the flowcell. A cluster is a product of a tagmentation fragment of the long DNA, each of which can be sequenced using SBS sequencing, for example. A group of clusters from a single long DNA molecule is referred to herein as a long DNA island. In some embodiments, a single hydrogel bead may encapsulate a single long DNA molecule or multiple long DNA molecules. Each long DNA molecule generates a single long DNA island. The clusters of all long DNA islands within a single hydrogel bead is referred to herein as a cluster cloud. Thus, a cluster cloud represents all clusters within a single hydrogel bead, and may include many long DNA islands (each long DNA island representing a single long DNA molecule), and each long DNA island includes multiple clusters.

[0021] The beads may include hydrogel polymers and crosslinkers that are mixed in the presence of a long DNA molecule, or a source containing a long DNA molecule, which then form hydrogel beads encapsulating the DNA molecule. In some embodiments, the long DNA source is a cell. The hydrogel beads may include pores that allow diffusion of reagents through the hydrogel bead while retaining the long DNA within the bead, thereby allowing reactions to take place within each of the beads.

[0022] Some embodiments include methods of using the beads encapsulating long DNA to perform nucleic acid reactions, including for example, high-throughput spatial indexing of long DNA molecules. As shown in FIG. 1A, library preparation from a long DNA molecule may be readily prepared by clustering and seeding the clusters from a single long DNA molecule as a “cluster patch” on the surface, which can then be read and spatially mapped. As used herein, the term “long DNA” can include DNA fragments that are greater than 300 base pairs. Long DNA fragments, as used herein, refers to DNA of a length of great than 1 kb, 2.5 kb, 5 kb, or more, such as 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90,

100, 200, 300, 400, or 500 kb, or more, including an amount within a range defined by any two of the aforementioned values.

[0023] Without being bound by theory, the methods, systems, and compositions provided herein include several advantages over current library preparation techniques. For example, in some embodiments, the methods allow sample preparation in a single bead to be used to fragment a genomic sample into a series of long DNA fragments. That single bead can then be adhered to one special location on a flow cell where the long DNA fragments are deposited such that each of the long DNA fragments are positioned adjacent one another on the flow cell surface. The system then determines the nucleotide sequence from each long DNA fragment. Since they are adjacent one another on the flow cell surface, the system may use this spatial location data to more efficiently reconstruct the final sequence of the original genomic DNA. The system may deposit spatially co-located reads directly from single cells, long DNA fragments, or chromosomes. In some embodiments, the methods allow for low input, PCR-free workflow for library preparation. In some embodiments, the methods may be performed without a need for molecular barcoding. In some embodiments, the methods allow simplified workflow automation. In some embodiments, the methods are compatible with a variety of nucleic acid assays and workflows.

[0024] Some embodiments relate to methods of preparing a hydrogel bead that encapsulates long DNA. In some embodiments, the hydrogel bead encapsulating long DNA can be used to process the cellular genome and perform DNA library preparation inside the bead. In some embodiments, the hydrogel bead encapsulating a long DNA fragment encapsulates a single cell, which can be used to process the cellular genomic DNA, and to perform whole DNA library preparation inside the bead.

[0025] In some embodiments, the pore size of the hydrogel bead can be engineered to allow the diffusion of enzymes, chemicals, and smaller sized primers (< 50bps), while retaining larger nucleic acids (>300bps) such that the long DNA fragments and the produced DNA library may be retained inside the hydrogel beads during processing. In some embodiments, specific primers can be chemically linked within the hydrogel bead matrix to hybridize and process specific genomic DNA. The DNA library from a single cell can then be released to a specific area, for example, on flow cell surface for library seeding. Subsequently, this results in a spatial distribution of “DNA clusters” on the flow cell

originating from the encapsulated long DNA fragments, thus simplifying the read alignment during post processing.

[0026] As used herein, the term “reagent” describes an agent or a mixture of two or more agents useful for reacting with, interacting with, diluting, or adding to a sample, and may include agents used in nucleic acid reactions, including, for example buffers, chemicals, enzymes, polymerase, primers having a size of less than 50 base pairs, template nucleic acids, nucleotides, labels, dyes, or nucleases. In some embodiments, the reagent includes lysozyme, proteinase K, random hexamers, polymerase (for example, Φ 29 DNA polymerase, Taq polymerase, Bsu polymerase), transposase (for example, Tn5), primers (for example, P5 and P7 adaptor sequences), ligase, catalyzing enzyme, deoxynucleotide triphosphates, buffers, or divalent cations.

Hydrogel Beads Encapsulating Genetic Material

[0027] One embodiment includes a bead including a hydrogel polymer and genetic material. As used herein, the term “hydrogel” refers to a substance formed when an organic polymer (natural or synthetic) is cross-linked via covalent, ionic, or hydrogen bonds to create a three-dimensional open-lattice structure that entraps water molecules to form a gel. In some embodiments, the hydrogel may be a biocompatible hydrogel. As used herein, the term “biocompatible hydrogel” refers to a polymer that forms a gel that is not toxic to living cells and allows sufficient diffusion of oxygen and nutrients to entrapped cells to maintain viability. In some embodiments, the hydrogel polymer includes 60-90% fluid, such as water, and 10-30% polymer. In certain embodiments, the water content of hydrogel is about 70-80%.

[0028] Hydrogels may be prepared by cross-linking hydrophilic biopolymers or synthetic polymers. Thus, in some embodiments, the hydrogel may include a crosslinker. As used herein, the term “crosslinker” refers to a molecule that can form a three-dimensional network when reacted with the appropriate base monomers. Examples of the hydrogel polymers, which may include one or more crosslinkers, include but are not limited to, hyaluronans, chitosans, agar, heparin, sulfate, cellulose, alginates (including alginate sulfate), collagen, dextrans (including dextran sulfate), pectin, carrageenan, polylysine, gelatins (including gelatin type A), agarose, (meth)acrylate-oligolactide-PEO-oligolactide-(meth)acrylate, PEO-PPO-PEO copolymers (Pluronics), poly(phosphazene),

poly(methacrylates), poly(N-vinylpyrrolidone), PL(G)A-PEO-PL(G)A copolymers, poly(ethylene imine), polyethylene glycol (PEG)-thiol, PEG-acrylate, acrylamide, N,N'-bis(acryloyl)cystamine, PEG, polypropylene oxide (PPO), polyacrylic acid, poly(hydroxyethyl methacrylate) (PHEMA), poly(methyl methacrylate) (PMMA), poly(N-isopropylacrylamide) (PNIPAAm), poly(lactic acid) (PLA), poly(lactic-co-glycolic acid) (PLGA), polycaprolactone (PCL), poly(vinylsulfonic acid) (PVSA), poly(L-aspartic acid), poly(L-glutamic acid), bisacrylamide, diacrylate, diallylamine, triallylamine, divinyl sulfone, diethyleneglycol diallyl ether, ethyleneglycol diacrylate, polymethyleneglycol diacrylate, polyethyleneglycol diacrylate, trimethylopropylene trimethacrylate, ethoxylated trimethylol triacrylate, or ethoxylated pentaerythritol tetracrylate, or combinations thereof. Thus, for example, a combination may include a polymer and a crosslinker, for example polyethylene glycol (PEG)-thiol/PEG-acrylate, acrylamide/N,N'-bis(acryloyl)cystamine (BACy), or PEG/polypropylene oxide (PPO).

[0029] In some embodiments, a crosslinker forms a disulfide bond in the hydrogel polymer, thereby linking hydrogel polymers. In some embodiments, the hydrogel polymers form a hydrogel matrix having pores (for example, a porous hydrogel matrix). These pores are capable of retaining sufficiently large genetic material within the hydrogel bead, for example, long DNA fragments, but allow small materials, such as reagents, to pass through the pores, thereby passing in and out of the hydrogel beads. In some embodiments, the pore size is finely tuned by varying the ratio of the concentration of polymer to the concentration of crosslinker. In some embodiments, the ratio of polymer to crosslinker is 30:1, 25:1, 20:1, 19:1, 18:1, 17:1, 16:1, 15:1, 14:1, 13:1, 12:1, 11:1, 10:1, 9:1, 8:1, 7:1, 6:1, 5:1, 4:1, 3:1, 2:1, 1:1, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:15, 1:20, or 1:30, or a ratio within a range defined by any two of the aforementioned ratios. In some embodiments, additional functions such as DNA primer, or charged chemical groups can be grafted to polymer matrix to meet the requirements of different applications.

[0030] As used herein, the term “porosity” means the fractional volume (dimension-less) of a hydrogel that is composed of open space, for example, pores or other openings. Therefore, porosity measures void spaces in a material and is a fraction of volume of voids over the total volume, as a percentage between 0 and 100% (or between 0 and 1).

Porosity of the hydrogel may range from 0.5 to 0.99, from about 0.75 to about 0.99, or from about 0.8 to about 0.95.

[0031] The hydrogels can have any pore size. As used herein, the term “pore size” refers to a diameter or an effective diameter of a cross-section of the pores. The term “pore size” can also refer to an average diameter or an average effective diameter of a cross-section of the pores, based on the measurements of a plurality of pores. The effective diameter of a cross-section that is not circular equals the diameter of a circular cross-section that has the same cross-sectional area as that of the non-circular cross-section. In some embodiments, the hydrogel can be swollen when the hydrogel is hydrated. The sizes of the pores size can then change depending on the water content in the hydrogel. In some embodiments, the pores of the hydrogel can have a pore of sufficient size to retain genetic material within the hydrogel but allow reagents to pass through.

[0032] In some embodiments, the crosslinker is a reversible crosslinker. In some embodiments, a reversible crosslinker is capable of reversibly crosslinking the hydrogel polymer and is capable of being un-crosslinked in the presence of a cleaver. In some embodiments, a crosslinker can be cleaved by the presence of a reducing agent, by elevated temperature, or by an electric field. In some embodiments, the reversible crosslinker may be N,N'-bis(acryloyl)cystamine, a reversible crosslinker for polyacrylamide gels, wherein a disulfide linkage may be broken in the presence of a suitable reducing agent. In some embodiments, contacting the crosslinker with a reducing agent cleaves the disulfide bonds of the crosslinker, breaking down the hydrogel beads. The hydrogel beads degrade, and release the contents, such as nucleic acids that were retained therein. In some embodiments, the crosslinker is cleaved by increasing the temperature to greater than 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100°C. In some embodiments, the crosslinker is cleaved by contacting the hydrogel beads with a reducing agent. In some embodiments, the reducing agents include phosphine compounds, water soluble phosphines, nitrogen containing phosphines and salts and derivatives thereof, dithioerythritol (DTE), dithiothreitol (DTT) (cis and trans isomers, respectively, of 2,3-dihydroxy-1,4-dithiolbutane), 2-mercaptoethanol or β -mercaptoethanol (BME), 2-mercaptoethanol or aminoethanethiol, glutathione, thioglycolate or thioglycolic acid, 2,3-dimercaptopropanol, tris(2-carboxyethyl)phosphine (TCEP),

tris(hydroxymethyl)phosphine (THP), or P-[tris(hydroxymethyl)phosphine] propionic acid (THPP).

[0033] In some embodiments, elevating the temperature to increase diffusion or contacting with a reducing agent degrades the crosslinker, thereby releasing encapsulated genetic material from the hydrogel bead.

[0034] In some embodiments, the crosslinking of the crosslinker establishes pores within the hydrogel bead. In some embodiments, the size of the pores in the hydrogel beads are regulatable and are formulated to encapsulate genetic material, such as DNA fragments of greater than about 5000 base pairs, but to allow smaller particles, such as reagents, or smaller sized nucleic acids of less than about 50 base pairs, such as primers, to pass through the pores, as shown in FIG. 1B. In some embodiments, the reagents including reagents for processing genetic material, such as reagents for isolating nucleic acids from a cell, for amplifying, barcoding, or sequencing nucleic acids, or for preparation of nucleic acid libraries. In some embodiments, reagents include, for example, lysozyme, proteinase K, random hexamers, polymerase (for example, Φ 29 DNA polymerase, Taq polymerase, Bsu polymerase), transposase (for example, Tn5), primers (for example, P5 and P7 adaptor sequences), ligase, catalyzing enzyme, deoxynucleotide triphosphates, buffers, or divalent cations.

[0035] In some embodiments, the long DNA includes genomic DNA, viral nucleic acids, bacterial nucleic acids, or mammalian nucleic acids. In some embodiments, the hydrogel beads include a source of long DNA, including, for example a cell. In some embodiments, the cell is a single cell, including a prokaryotic or a eukaryotic cell. In some embodiments, the cell is a mammalian cell. In some embodiments, the cell is a human cell. In some embodiments, the cell is a bacterial cell. Thus, as shown in FIGs. 7A and 7B, the method may be performed on long DNA fragments or on cells, either or both of which is encapsulated with a hydrogel bead.

Methods of Making Beads

[0036] Some embodiments provided herein relate to methods of making beads that encapsulate long DNA fragments. In some embodiments, a hydrogel bead is prepared by vortex assisted emulsion. As used herein, vortex assisted emulsion refers to vortexing a hydrogel polymer with long DNA fragments or a source of long DNA fragments in a

container, such as in a tube, vial, or reaction vessel. The components can be mixed, for example by manual or mechanical vortexing or shaking. In some embodiments, manual mixing results in hydrogel beads that encapsulate genetic material having a size of 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, or 150 μm in diameter, or a size within a range defined by any two of the aforementioned values. In some embodiments, the size of the beads is non-uniform, and thus, the size of the beads includes beads of various diameters.

[0037] In some embodiments, the beads are prepared by microfluidic droplet generation. As shown in FIG. 1B, microfluidic droplet generation includes use of a microfluidic device for assisted gel emulsion generation. In some embodiments, the microfluidic device includes microchannels configured to produce a hydrogel bead of a desired size and configured to encapsulate a selected amount of genetic material per bead. In some embodiments, the microfluidic device has a height of 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, or 200 μm , or a height within a range defined by any two of the aforementioned values. In some embodiments, the microfluidic device includes one or more channels. In some embodiments, the microfluidic device includes a channel for an aqueous stream and a channel for an immiscible fluid. In some embodiments, the width of the one or more channels is identical. In some embodiments, the width of the one or more channels is different. In some embodiments, the width of the one or more channels is 20, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 μm , or a width within a range defined by any two of the aforementioned values. In some embodiments, the width of the aqueous channel is 75 μm . In some embodiments, the width of the immiscible fluid channel is 78 μm . One of skill in the art will recognize that the width can vary to finely tune the size of the bead. In addition to the size of the microfluidic device and the width of the channels, the flow rate of the aqueous channel and the immiscible fluid channel may also affect the size of the hydrogel beads.

[0038] In some embodiments, the flow rate of the solution in the aqueous phase channel is 1, 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 $\mu\text{L}/\text{min}$, or a rate within a range defined by any two of the aforementioned values. In some embodiments, the flow rate of the immiscible fluid in the immiscible fluid channel is 20, 30, 50, 80, 100, 150, 160, 170, 180, 190, 200, 225, 250, 275, 300, 325, 350, 375, or 400 $\mu\text{L}/\text{min}$,

or a rate within a range defined by any two of the aforementioned values. In some embodiments, the solution in the aqueous phase includes a hydrogel polymer, a crosslinker, and genetic material, which flows through an aqueous channel into an immiscible fluid, such as a carrier oil, at a flow rate less than the flow rate of the immiscible fluid, thereby forming droplets. In some embodiments, the immiscible fluid is oil, such as mineral oil, a hydrocarbon oil, a silicon oil, or a polydimethylsiloxane oil, or mixtures thereof. In some embodiments, the hydrogel droplets containing genetic material are formulated in a uniform size distribution. In some embodiments, the size of the hydrogel beads is finely tuned by adjusting the size of the microfluidic device, the size of the one or more channels, or the flow rate of either or both of the aqueous solution or immiscible fluid. In some embodiments, the resulting hydrogel bead has a diameter ranging from 2 to 150 μm , for example, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, or 150 μm , or a diameter within a range defined by any two of the aforementioned values.

[0039] In some embodiments, the size and uniformity of the hydrogel bead encapsulating genetic material can be further controlled by contacting hydrogel polymer prior to bead formation with a fluidic modifier, such as with an alcohol, including isopropyl alcohol.

[0040] In some embodiments, the amount of long DNA fragments encapsulated within a bead can be controlled by diluting or concentrating the long DNA fragments within the inputted sample. The sample including the long DNA fragments is mixed with hydrogel polymer, and the hydrogel polymer containing the long DNA fragments is submitted to vortex assisted emulsion or microfluidic droplet generation, as described herein.

[0041] In some embodiments, the hydrogel beads are functionalized with a nucleotide. In some embodiments, the nucleotide is an oligonucleotide or polyT nucleotide. In some embodiments, the nucleotide is bound to the hydrogel bead, and the functionalized bead can be used for targeted capture of a nucleotide of interest.

Methods of Processing Long DNA Fragments within Hydrogel Beads

[0042] Some embodiments include methods of processing long DNA fragments within a bead as shown in FIG. 2, which depicts a flow diagram for preparing and processing

long DNA molecules in a hydrogel bead. In a first step, a DNA sample, such as from genomic data or a cell is encapsulated within a hydrogel bead. In some embodiments, the long DNA fragment is retained within the hydrogel beads, and reagents are able to pass through the pores of the hydrogel beads. In some embodiments, reagents can include lysis agents, nucleic acid purification agents, tagmentation agents, PCR agents, or other agents used in processing of genetic materials. Thus, the hydrogel beads provide a microenvironment for controlled reactions of long DNA fragments within the hydrogel beads by allowing a barrier for reagents to pass in and out of the hydrogel beads, while retaining the long DNA fragments within the beads. Once the DNA is encapsulated into the beads, the process moves to the next step where the sample can be loaded into a flow cell to create the long DNA fragments through the library preparation process.

[0043] As used herein, the term “tagmentation” refers to the modification of DNA by a transposome complex comprising transposase enzyme complexed with adaptors comprising transposon end sequence. Tagmentation results in the simultaneous fragmentation of the DNA and ligation of the adaptors to the 5' ends of both strands of duplex fragments. Following a purification step to remove the transposase enzyme, additional sequences can be added to the ends of the adapted fragments, for example by PCR, ligation, or any other suitable methodology known to those of skill in the art.

[0044] In some embodiments, entire DNA library preparation can be accomplished seamlessly inside the hydrogel beads bound to the flow cell with multiple reagent exchanges by passing through the porous hydrogel while retaining the gDNA and its library products within the hydrogel matrix. The hydrogel may be resistant to high temperatures up to 95°C for several hours to support different biochemical reactions.

[0045] In the next step in the process, the hydrogel bead encapsulating the long DNA fragments from the prior library preparation is treated to release, purify and isolate the long DNA fragments from the bead. Thus, for example the hydrogel bead is contacted with a lysis buffer. As used herein, “lysis” means perturbation or alteration to a cell wall or viral particle facilitating access to or release of the cellular RNA or DNA. Neither complete disruption nor breakage of the cell wall is an essential requirement for lysis. By the term “lysis buffer” is meant a buffer that contains at least one lysing agent. Typical enzymatic lysing agents include, but are not limited to, lysozyme, glucolase, zymolose, lyticase,

proteinase K, proteinase E, and viral endolysins and exolysins. Thus, for example, lysis of cells in the beads may be performed by introducing lysing agents, such as lysozyme and proteinase K into the hydrogel beads. The gDNA from the cells is now contained within the beads. In some embodiments, following lysis treatment, isolated nucleic acid is retained within the hydrogel bead, and may be used for further processing.

[0046] As used herein, the terms “isolated,” “to isolate,” “isolation,” “purified,” “to purify,” “purification,” and grammatical equivalents thereof as used herein, unless specified otherwise, refer to the reduction in the amount of at least one contaminant (such as protein and/or nucleic acid sequence) from a sample or from a source (e.g., a cell) from which the material is isolated. Thus purification results in an “enrichment,” for example, an increase in the amount of a desirable protein and/or nucleic acid sequence in the sample.

[0047] In some embodiments, the encapsulated nucleic acids are sequenced in full or in part within the hydrogel beads. The encapsulated nucleic acids can be sequenced according to any suitable sequencing methodology, such as direct sequencing, including sequencing by synthesis, sequencing by ligation, sequencing by hybridization, nanopore sequencing and the like.

[0048] Some embodiments provided herein relate to sequencing-by-synthesis (SBS) enabled for long DNA fragments. In SBS, extension of a nucleic acid primer along a nucleic acid template (e.g. a target nucleic acid or amplicon thereof) is monitored to determine the sequence of nucleotides in the template. The underlying chemical process can be polymerization (e.g. as catalyzed by a polymerase enzyme). In a particular polymerase-based SBS embodiment, fluorescently labeled nucleotides are added to a primer (thereby extending the primer) in a template dependent fashion such that detection of the order and type of nucleotides added to the primer can be used to determine the sequence of the template.

[0049] One or more amplified encapsulated nucleic acids can be subjected to an SBS or other detection technique that involves repeated delivery of reagents in cycles. For example, to initiate a first SBS cycle, one or more labeled nucleotides, DNA polymerase, etc., can be flowed into/through a hydrogel bead that houses one or more amplified nucleic acid molecules. Those sites where primer extension causes a labeled nucleotide to be incorporated can be detected. Optionally, the nucleotides can further include a reversible

termination property that terminates further primer extension once a nucleotide has been added to a primer. For example, a nucleotide analog having a reversible terminator moiety can be added to a primer such that subsequent extension cannot occur until a deblocking agent is delivered to remove the moiety. Thus, for embodiments that use reversible termination, a deblocking reagent can be delivered to the flow cell (before or after detection occurs). Washes can be carried out between the various delivery steps. The cycle can then be repeated *n* times to extend the primer by *n* nucleotides, thereby detecting a sequence of length *n*. Exemplary SBS procedures, fluidic systems and detection platforms that can be readily adapted for use with amplicons produced by the methods of the present disclosure are described, for example, in Bentley et al., *Nature* 456:53-59 (2008), WO 04/018497; U.S. Pat. No. 7,057,026; WO 91/06678; WO 07/123744; U.S. Pat. No. 7,329,492; U.S. Pat. No. 7,211,414; U.S. Pat. No. 7,315,019; U.S. Pat. No. 7,405,281, and US 2008/0108082.

[0050] Other sequencing procedures that use cyclic reactions can be used, such as pyrosequencing. Pyrosequencing detects the release of inorganic pyrophosphate (PPi) as particular nucleotides are incorporated into a nascent nucleic acid strand (Ronaghi, et al., *Analytical Biochemistry* 242(1), 84-9 (1996); Ronaghi, *Genome Res.* 11(1), 3-11 (2001); Ronaghi et al. *Science* 281(5375), 363 (1998); U.S. Pat. No. 6,210,891; U.S. Pat. No. 6,258,568 and U.S. Pat. No. 6,274,320).

In pyrosequencing, released PPi can be detected by being immediately converted to adenosine triphosphate (ATP) by ATP sulfurylase, and the level of ATP generated can be detected via luciferase-produced photons. Thus, the sequencing reaction can be monitored via a luminescence detection system. Excitation radiation sources used for fluorescence based detection systems are not necessary for pyrosequencing procedures. Useful fluidic systems, detectors and procedures that can be adapted for application of pyrosequencing to amplicons produced according to the present disclosure are described, for example, in WIPO Pat. App. Ser. No. PCT/US11/57111, US 2005/0191698 A1, U.S. Pat. No. 7,595,883, and U.S. Pat. No. 7,244,559.

[0051] Some embodiments can utilize methods involving the real-time monitoring of DNA polymerase activity. For example, nucleotide incorporations can be detected through fluorescence resonance energy transfer (FRET) interactions between a

fluorophore-bearing polymerase and γ -phosphate-labeled nucleotides, or with zero mode waveguides (ZMWs). Techniques and reagents for FRET-based sequencing are described, for example, in Levene et al. *Science* 299, 682-686 (2003); Lundquist et al. *Opt. Lett.* 33, 1026-1028 (2008); Korlach et al. *Proc. Natl. Acad. Sci. USA* 105, 1176-1181 (2008).

[0052] Some SBS embodiments include detection of a proton released upon incorporation of a nucleotide into an extension product. For example, sequencing based on detection of released protons can use an electrical detector and associated techniques that are commercially available. Examples of such sequencing systems are pyrosequencing (e.g. commercially available platform from 454TM Life Sciences a subsidiary of Roche), sequencing using γ -phosphate-labeled nucleotides (e.g. commercially available platform from Pacific Biosciences) and sequencing using proton detection (e.g. commercially available platform from Ion Torrent subsidiary of Life Technologies) or sequencing methods and systems described in US 2009/0026082 A1; US 2009/0127589 A1; US 2010/0137143 A1; or US 2010/0282617 A1.

Methods set forth herein for amplifying target nucleic acids using kinetic exclusion can be readily applied to substrates used for detecting protons. More specifically, methods set forth herein can be used to produce clonal populations of amplicons that are used to detect protons.

[0053] Another sequencing technique is nanopore sequencing (see, for example, Deamer et al. *Trends Biotechnol.* 18, 147-151 (2000); Deamer et al. *Acc. Chem. Res.* 35:817-825 (2002); Li et al. *Nat. Mater.* 2:611-615 (2003)).

In some nanopore embodiments, the target nucleic acid or individual nucleotides removed from a target nucleic acid pass through a nanopore. As the nucleic acid or nucleotide passes through the nanopore, each nucleotide type can be identified by measuring fluctuations in the electrical conductance of the pore. (U.S. Pat. No. 7,001,792; Soni et al. *Clin. Chem.* 53, 1996-2001 (2007); Healy, *Nanomed.* 2, 459-481 (2007); Cockroft et al. *J. Am. Chem. Soc.* 130, 818-820 (2008)).

[0054] Exemplary methods for array-based expression and genotyping analysis that can be applied to detection according to the present disclosure are described in U.S. Pat. No. 7,582,420; 6,890,741; 6,913,884 or 6,355,431 or US Pat. Pub. Nos. 2005/0053980 A1,

[0055] In the methods of isolating nucleic acids, amplification, and sequencing as described herein, various reagents are used for nucleic acid isolation and preparation. Such reagents may include, for example, lysozyme, proteinase K, random hexamers, polymerase (for example, Φ 29 DNA polymerase, Taq polymerase, Bsu polymerase), transposase (for example, Tn5), primers (for example, P5 and P7 adaptor sequences), ligase, catalyzing enzyme, deoxynucleotide triphosphates, buffers, or divalent cations. These reagents pass through the pores of the hydrogel beads, whereas the genetic material is retained within the hydrogel beads. An advantage of the methods set forth herein is that they provide for an encapsulated microenvironment for the processing of nucleic acids within a hydrogel bead. This enables single cell processing for rapid and efficient processing of a target nucleic acid.

[0056] Adaptors can include sequencing primer sites, amplification primer sites, and indexes. As used herein an “index” can include a sequence of nucleotides that can be used as a molecular identifier and/or barcode to tag a nucleic acid, and/or to identify the source of a nucleic acid. In some embodiments, an index can be used to identify a single nucleic acid, or a subpopulation of nucleic acids. In some embodiments, nucleic acid libraries can be prepared within a hydrogel bead. In some embodiments, a single cell encapsulated within a hydrogel bead may be used for combinatorial indexing of the single cells, for example, using a contiguity preserving transposition (CPTSeq) approach. In some embodiments, DNA from a single cell may be barcoded by encapsulation of single cells after WGA amplification with another bead carrying barcoded transposons and dissolving the gel matrix by contacting it with a reducing agent, for example, to release genomic DNA for barcoding.

[0057] Embodiments of the “spatial indexing” methods and techniques described herein shortens data analysis and simplifies the process of library preparation from single cells and long DNA molecules. Existing protocols for single cell sequencing requires efficient physical separation of the cells and uniquely barcoding each isolated cell and pooling everything back together to do sequencing. Current protocols for synthetic long reads also requires cumbersome barcoding steps, and pooling each barcoded fragments together for sequencing and letting data analysis to distinguish genetic information coming from each

barcoded cell. During these long processes there is also loss of genetic material which causes dropouts in the sequences. Embodiments described herein not only shorten the process but also increase data resolution for single cells. Furthermore, embodiments provided herein simplify the assembly of genomes of new organisms. Embodiments described herein may be used to reveal rare genetic variations and co-occurrence of mutations. In some embodiments, DNA library confined in the hydrogel beads until release provide the opportunity to control the size of the fragments that is released on the surface by controlling the release process and hydrogel formulation.

[0058] In some embodiments, the surface is a flow cell device. In some embodiments, the flow cell is a custom flow cell device having wells, grooves, or patterns. In some embodiments, the flow cell comprises a patterned surface. In some embodiments, the patterned surface comprises wells. In some embodiments, the wells are from about 10 μm to about 50 μm in diameter, such as 10 μm , 15 μm , 20 μm , 25 μm , 30 μm , 35 μm , 40 μm , 45 μm , or 50 μm in diameter, or within a range defined by any two of the aforementioned values, and wherein the wells are about 0.5 μm to about 1 μm in depth, such as 0.5 μm , 0.6 μm , 0.7 μm , 0.8 μm , 0.9 μm , or 1 μm in depth, or within a range defined by any two of the aforementioned values. In some embodiments, the wells are comprised of hydrophobic material. In some embodiments, the hydrophobic material comprises an amorphous fluoropolymer, such as CYTOP, Fluoropel®, or Teflon®.

[0059] In some embodiments, the library may be amplified using primer sites in the adaptor sequences, and sequenced using sequencing primer sites in the adaptor sequences. In some embodiments the adaptor sequences can include indexes to identify the source of the nucleic acids. The efficiency of subsequent amplification steps can be reduced by the formation of primer-dimers. To increase the efficiency of subsequent amplification steps, non-ligated single-stranded adaptors can be removed from ligation products.

Preparing Nucleic Acid Libraries with Hydrogel Beads

[0060] Some embodiments of the systems, methods and compositions provided herein include methods in which adaptors are ligated to target nucleic acids. Adaptors can include sequencing primer binding sites, amplification primer binding sites, and indexes. For example, an adaptor can include a P5 sequence, a P7 sequence, or a complement thereof. As used herein a P5 sequence comprises a sequence defined by SEQ ID NO: 1

(AATGATACGGCGACCACCGA) and a P7 sequence comprises a sequence defined by SEQ ID NO: 2 (CAAGCAGAAGACGGCATACGA). In some embodiments, the P5 or P7 sequence can further include a spacer polynucleotide, which may be from 1 to 20, such as 1 to 15, or 1 to 10, nucleotides, such as 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length. In some embodiments, the spacer includes 10 nucleotides. In some embodiments, the spacer is a polyT spacer, such as a 10T spacer. Spacer nucleotides may be included at the 5' ends of polynucleotides, which may be attached to a suitable support via a linkage with the 5' end of the polynucleotide. Attachment can be achieved through a sulphur-containing nucleophile, such as phosphorothioate, present at the 5' end of the polynucleotide. In some embodiments, the polynucleotide will include a polyT spacer and a 5' phosphorothioate group. Thus, in some embodiments, the P5 sequence is 5'phosphorothioate-TTTTTTTTTTAATGATACGGCGACCACCGA-3', and in some embodiments, the P7 sequence is 5'phosphorothioate-TTTTTTTTTTCAAGCAGAAGACGGCATACGA-3'.

[0061] Indexes can be useful to identify the source of a nucleic acid molecule. In some embodiments, an adaptor can be modified to prevent the formation of concatemers, for example by the addition of blocking groups that prevent extension of the adaptor at one or both ends. Examples of 3' blocking groups include a 3'-spacer C3, a dideoxynucleotide, and attachment to a substrate. Examples of 5' blocking groups include a dephosphorylated 5' nucleotide, and attachment to a substrate.

[0062] Adaptors include nucleic acids, such as single-stranded nucleic acids. Adaptors can include short nucleic acids having a length less than, greater than, or equal to about 5 nucleotides, 10 nucleotides, 20 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 60 nucleotides, 70 nucleotides, 80 nucleotides, 90 nucleotides, 100 nucleotides, or a range between any two of the foregoing sizes. In some embodiments, the adaptors are of sufficient size to pass through the pores of the hydrogel beads. Target nucleic acids include DNA, such as genomic or cDNA; RNA, such as mRNA, sRNA or rRNA; or a hybrid of DNA and RNA. The nucleic acid can be isolated from a single cell encapsulated within a hydrogel bead. A nucleic acid can contain phosphodiester bonds, and can include other types of backbones, comprising, for example, phosphoramidate, phosphorothioate, phosphorodithioate, O-methylphosphoroamidite and peptide nucleic acid backbones and linkages. A nucleic acid can contain any combination of deoxyribo- and ribonucleotides, and

any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthanine, hypoxanthanine, isocytosine, isoguanine, and base analogs such as nitropyrrole (including 3-nitropyrrole) and nitroindole (including 5-nitroindole). In some embodiments, a nucleic acid can include at least one promiscuous base. A promiscuous base can base-pair with more than one different type of base and can be useful, for example, when included in oligonucleotide primers or inserts that are used for random hybridization in complex nucleic acid samples such as genomic DNA samples. An example of a promiscuous base includes inosine that may pair with adenine, thymine, or cytosine. Other examples include hypoxanthine, 5-nitroindole, acyclic 5-nitroindole, 4-nitropyrazole, 4-nitroimidazole and 3-nitropyrrole. Promiscuous bases that can base-pair with at least two, three, four or more types of bases can be used.

[0063] An example method includes dephosphorylating the 5' ends of target nucleic acids to prevent the formation of concatemers in subsequent ligation steps; ligating first adaptors to the 3' ends of the dephosphorylated targets using a ligase, in which the 3' ends of the first adaptors are blocked; re-phosphorylating of the 5' ends of the ligated targets; ligating a second adaptor to the 5' ends of the dephosphorylated targets using the single-stranded ligase, in which the 5' ends of the second adaptors are non-phosphorylated.

[0064] Another example includes partial digestion of the nucleic acid with a 5' exonuclease to form a double-stranded nucleic acid with single-stranded 3' overhangs. An adaptor containing a 3' blocking group can be ligated to the 3' ends of double-stranded nucleic acid with 3' overhangs. The double-stranded nucleic acid with 3' overhangs with ligated adaptors can be dehybridized to form single-stranded nucleic acids. An adaptor containing a non-phosphorylated 5' end can be ligated to the 5' end of the single-stranded nucleic acid.

[0065] Methods to dephosphorylate nucleic acids, such as the 5' nucleotide of a nucleic acid include contacting a nucleic acid with a phosphatase. Examples of phosphatases include calf intestinal phosphatase, shrimp alkaline phosphatase, antarctic phosphatase, and APEX alkaline phosphatase (Epicentre).

[0066] Methods to ligate nucleic acids include contacting nucleic acids with a ligase. Examples of ligases include T4 RNA ligase 1, T4 RNA ligase 2, RtcB ligase, Methanobacterium RNA ligase, and TS2126 RNA ligase (CIRCLIGASE).

[0067] Methods to phosphorylate nucleic acids, such as the 5' nucleotide of a nucleic acid include contacting a nucleic acid with a kinase. Examples of kinases include T4 polynucleotide kinase.

[0068] Embodiments provided herein relate to preparing nucleic acids libraries in a hydrogel bead, such that the nucleic acid library is prepared in a single reaction volume.

[0069] Embodiments of the systems and methods provided herein include kits, containing any one or more of the hydrogel polymers, crosslinkers, or microfluidic devices for preparing hydrogel beads that encapsulate genetic material, and further including components useful for processing of the genetic material, including reagents for cell lysis, and nucleic acid amplification and sequencing, or for nucleic acid library preparation, including lysozyme, proteinase K, random hexamers, polymerase (for example, Φ 29 DNA polymerase, Taq polymerase, Bsu polymerase), transposase (for example, Tn5), primers (for example, P5 and P7 adaptor sequences), ligase, catalyzing enzyme, deoxynucleotide triphosphates, buffers, or divalent cations as described herein, and as used for the respective processing of genetic material.

EXAMPLES

Example 1—Preparation of Hydrogel Beads

[0070] The following example demonstrates an embodiment of preparing hydrogel beads encapsulating long DNA fragments using microfluidic droplet generators.

[0071] A droplet generator was used to generate the hydrogel beads. Samples containing long DNA fragments were mixed with polymer precursor and the mixture was loaded into a sample reservoir on a cartridge. Within 2 minutes, around 50,000 hydrogel beads containing long DNA were generated from each channel (8 channels for 8 independent sample processing each cartridge). The long DNA hydrogel beads were loaded onto a flow cell, where hydrogel beads stuck inside (100 μ m high channel and 120 μ m hydrogel beads diameter) for hands-free library preparation. The Nextera enzymes and reagents contact the flow cell, contacting the long DNA embedded inside the hydrogel bead, forming a library. The library was then seeded on the flow cell. During library seeding, oil was loaded to fill the void between beads and the flow cell was heated to accelerate diffusion of the library. In the presence of the oil, seeding occurred in close proximity to the footprint of each hydrogel

bead (from 120 μm diameter hydrogel beads, library seeding is limited to a roughly 120 μm diameter area).

[0072] Long DNA molecules were loaded and trapped in hydrogel beads (about 120 μm in diameter) and library preparation was directly performed on these long DNA molecules embedded inside the hydrogel beads. As a result, all DNA libraries from a specific long DNA molecule were stored within the same hydrogel beads. The library was then released from the hydrogel beads to the flow cell surface to seed them as a group on the flow cell surface. The clusters released from a long DNA molecule grouped together as a “cluster patch” on the flow cell. Clusters inside a single patch from a single long DNA molecule simplifies re-construction of the genome with higher accuracy and fewer scaffolding gaps.

Example 2—Long DNA Spatial Indexing

[0073] The following example demonstrates an embodiment of strobed reads of long DNA fragment of 100 kb encapsulated within a hydrogel bead with or without MDA.

[0074] Hydrogel beads were prepared by mixing a polymer in the presence of Coriell genomic DNA of about 100 kb and forming hydrogel beads using a microdroplet generator. The DNA was subjected to spatial indexing sequencing by placing the formed hydrogel beads encapsulating the DNA fragments on a flow cell device, and contacting the flow cell with reagents. No MDA was performed. The beads were degraded and clusters formed on the flow cell device. As shown in FIG. 5, the average clusters per long DNA island was about 33, the average long DNA island size was 64000 base pairs, and there were about 405 long DNA islands per bead.

[0075] A second set of hydrogel beads were prepared by mixing a polymer in the presence of Coriell genomic DNA of about 100 kb and forming hydrogel beads using a microdroplet generator. The DNA was subjected to spatial indexing sequencing by placing the formed hydrogel beads encapsulating the DNA fragments on a flow cell device, and contacting the flow cell with reagents. MDA was performed prior to tagmentation. The beads were degraded and clusters formed on the flow cell device. As shown in FIG. 6, the average clusters per long DNA island increased to about 85, the average long DNA island size was 58000 base pairs, and there were about 166 long DNA islands per bead.

[0076] A third set of hydrogel beads were prepared by mixing a polymer in the presence of Coriell genomic DNA of about 10 kb and forming hydrogel beads using a microdroplet generator. The DNA was subjected to spatial indexing sequencing by placing the formed hydrogel beads encapsulating the DNA fragments on a flow cell device, and contacting the flow cell with reagents. MDA was performed prior to tagmentation. The beads were degraded and clusters formed on the flow cell device. As shown in FIG. 7, the average clusters per long DNA island was about 57, the average long DNA island size was 10461 base pairs, and there were about 85 long DNA islands per bead.

Example 3—Metagenomics on Complex Mixture of Microbial Species

[0077] The following example demonstrates an embodiment of identifying single cell microbes encapsulated within a hydrogel.

[0078] Hydrogel beads were prepared as described herein using a microfluidics microdroplet generator. The polymer material was mixed with a sample containing a number of microbes, including *L. gasseri*, *S. aureus*, *B. cereus*, *B. vulgatus*, *A. baumannii*, *S. agalactiae*, and *P. acnes*. The encapsulated cells were then lysed and subjected to library preparation, whereupon the hydrogel beads were degraded and the libraries deposited on a surface. As shown in FIG. 9, each microbe was capable of being identified due to its spatial compartmentalization on the flow cell device. Thus, the encapsulating and subsequent nucleic acid reactions enable strain-level identification of microbial species in complex mixtures using reads compartmentalization in a mini-metagenomics assay.

Example 4—On Flow Cell Spatial Indexing

[0079] The following example demonstrates an embodiment for on-flow cell spatial indexing.

[0080] A flow cell device was obtained and washed with 200 μ l PR2. Beads for processing were also washed with PR2. A diluted hydrogel was prepared in PR2. Increased dilution results in increased spacing between hydrogels. The hydrogel was embedded on the flow cell, and the introduction of air bubbles to the flow cell was avoided. 200 μ l PR2 was flowed through the flow cell to ensure beads remained fixed to go through the process. 100 μ l RSB was flowed through the flow cell.

[0081] A tagmentation mix was prepared by mixing 25 μ l tagmentation reagent, 23 μ l RSB, and 2 μ l enzyme. The tagmentation mix was introduced to the narrow channel to remove any possible air bubble on the inlet. The tagmentation mix was then flowed slowly to the inlet. The flow cell was sealed and incubated for 10 min at 55°C.

[0082] A stop buffer mix was prepared by mixing 25 μ l tagmentation buffer, 25 μ l RSB, and 10 μ l stop buffer. The stop buffer mix was slowly flowed onto the flow cell without introducing any bubbles, and incubated at room temperature for 5 mins. After incubation, 200 μ l of PR2 was flowed through the device.

[0083] NPM was prepared by mixing 175 μ l RSB and 75 μ l NPM. The NPM mix was slowly flowed onto the flow cell device without introducing any air bubble, and incubated for 3 mins at room temperature. 200 μ l of oil with surfactant was flowed onto the flow cell device. Micrographs revealed that the hydrogels were surrounded with NPM mix and oil. The flow cell was sealed and incubated for 3 mins at 72°C for gap filling reaction.

[0084] 20-30 μ l of oil with surfactant and oil with DTT (29/2 ratio) were flowed onto the flow cell device, and the device was sealed. The start temperature release process was 90°C for 3 mins, 60°C for 5 mins, 40°C for 2 mins, and 20°C for 2 mins. The flow cell was washed with 400 μ l PR2, and 200 μ l CLM. The flow cell was then washed with 400 μ l PR2. Where phix seeding is desired, a Phix was prepared with 2-3 pM concentration, and the phix library was flowed onto the device, and incubated at room temperature for 5mins. The flow cell was washed with 200 μ l PR2. 100-200 μ l AMX for 1st extension was flowed, and incubated for 5 min at 50°C. The flowcell was washed with PR2, and a 24 or 30 cycle amplification was performed

[0085] The embodiments, examples, and figures described herein provide compositions, methods, and systems for retaining genetic material in physically confined space during the process from lysis to library generation. Some embodiments provide libraries originated from single long DNA molecule or a single cell to be released on a surface of a flow cell in confined space. Once the library from a single DNA molecule or single cell in the individual compartments are released to the surface of the flow cell, the library from each compartment gets seeded at close proximity to each other.

[0086] The term “comprising” as used herein is synonymous with “including,” “containing,” or “characterized by,” and is inclusive or open-ended and does not exclude additional, unrecited elements or method steps.

[0087] The above description discloses several methods and materials of the present invention. This invention is susceptible to modifications in the methods and materials, as well as alterations in the fabrication methods and equipment. Such modifications will become apparent to those skilled in the art from a consideration of this disclosure or practice of the invention disclosed herein. Consequently, it is not intended that this invention be limited to the specific embodiments disclosed herein, but that it cover all modifications and alternatives coming within the true scope and spirit of the invention.

[0088]

WHAT IS CLAIMED IS:

1. A flow cell device for performing DNA sequencing, comprising:
a solid support comprising a surface having a degradable hydrogel encapsulating DNA deposited thereon, wherein the degradable hydrogel comprises pores that are sized to allow diffusion of a reagent through the hydrogel, but are too small to allow DNA to traverse the pores.
2. The flow cell device of Claim 1, wherein the solid support is functionalized with a surface polymer.
3. The flow cell device of Claim 2, wherein the surface polymer is poly(N-(5-azidoacetamidylpentyl) acrylamide-co-acrylamide) (PAZAM) or silane free acrylamide (SFA).
4. The flow cell device of any one of Claims 1-3, wherein the flow cell comprises a patterned surface.
5. The flow cell device of Claim 4, wherein the patterned surface comprises wells.
6. The flow cell device of Claim 5, wherein the wells are from 10 μm to 50 μm in diameter, or within a range defined by any two of the aforementioned values, and wherein the wells are 0.5 μm to 1 μm in depth, or within a range defined by any two of the aforementioned values.
7. The flow cell device of Claim 6, wherein the wells are 10 μm , 15 μm , 20 μm , 25 μm , 30 μm , 35 μm , 40 μm , 45 μm or 50 μm in diameter.
8. The flow cell device of Claim 6, wherein the wells are 0.5 μm , 0.6 μm , 0.7 μm , 0.8 μm , 0.9 μm , or 1 μm in depth.
9. The flow cell device of Claim 5, wherein the wells are comprised of hydrophobic material.
10. The flow cell device of Claim 9, wherein the hydrophobic material comprises an amorphous fluoropolymer, such as CYTOP, Fluoropel®, or Teflon®.
11. The flow cell device of any one of Claims 1-10, wherein the degradable hydrogel is a hydrogel bead or a hydrogel layer.
12. The flow cell device of Claim 11, wherein the hydrogel bead has a diameter of 50 μm to 150 μm .

13. The flow cell device of any one of Claims 1-12, wherein the hydrogel comprises polyethylene glycol (PEG)-thiol, PEG-acrylate, acrylamide, N,N'-bis(acryloyl)cystamine (BACy), PEG, polypropylene oxide (PPO), polyacrylic acid, poly(hydroxyethyl methacrylate) (PHEMA), poly(methyl methacrylate) (PMMA), poly(N-isopropylacrylamide) (PNIPAAm), poly(lactic acid) (PLA), poly(lactic-co-glycolic acid) (PLGA), polycaprolactone (PCL), poly(vinylsulfonic acid) (PVSA), poly(L-aspartic acid), poly(L-glutamic acid), polylysine, agar, agarose, alginate, heparin, alginate sulfate, dextran sulfate, hyaluronan, pectin, carrageenan, gelatin, chitosan, cellulose, collagen, bisacrylamide, diacrylate, diallylamine, triallylamine, divinyl sulfone, diethyleneglycol diallyl ether, ethyleneglycol diacrylate, polymethyleneglycol diacrylate, polyethyleneglycol diacrylate, trimethylopropylene trimethacrylate, ethoxylated trimethylol triacrylate, or ethoxylated pentaerythritol tetracrylate, or combinations or mixtures thereof.

14. The flow cell device of Claim 13, wherein the hydrogel comprises PEG-thiol/PEG-acrylate, acrylamide/N,N'-bis(acryloyl)cystamine (BACy), or PEG/PPO.

15. The flow cell device of any one of Claims 1-14, wherein the DNA is 50,000 base pairs or greater.

16. The flow cell device of any one of Claims 1-15, wherein the reagent comprises enzymes, chemicals, and primers having a size of less than 50 base pairs.

17. The flow cell device of any one of Claims 1-16, wherein the reagent comprises lysozyme, proteinase K, random hexamers, polymerase, transposase, primers, ligase, catalyzing enzyme, deoxynucleotide triphosphates, buffers, or divalent cations.

18. A system for DNA sequencing, comprising:

a stage configured to hold a flow cell device of any one of Claims 1-17;

a flow cell device of any one of Claims 1-17; and

a detector for obtaining sequencing data.

19. A method of sequencing DNA comprising:

obtaining a bead encapsulating DNA within a hydrogel;

performing a tagmentation reaction on the DNA encapsulated within the hydrogel, thereby generating a DNA library encapsulated within the hydrogel;

seeding the DNA library on a flow cell device of any one of claims 1-17; and

sequencing the DNA.

20. The method of Claim 19, wherein the DNA is 50,000 base pairs or greater.
21. The method of Claim 19 or 20, further comprising performing a DNA amplification reaction on DNA encapsulated within the hydrogel prior to performing the tagmentation reaction.
22. The method of Claim 21, wherein the DNA amplification reaction comprises multiple displacement amplification (MDA).
23. The method of any one of Claims 19-22, wherein the tagmentation reaction comprises contacting genetic material with a transposase mixture comprising adapter sequences and transposomes.
24. The method of any one of Claims 19-23, wherein seeding comprises cleaving the hydrogel bead to release the DNA library from the bead.
25. The method of Claim 24, wherein the bead is cleaved by contacting the bead with a cleavage mix or by heating the bead to about 90°C to release the DNA library.
26. The method of Claim 25, wherein the cleavage mix comprises dithiothreitol (DTT), tris(2-carboxyethyl)phosphine (TCEP), or tris(3-hydroxypropyl)phosphine (THP).

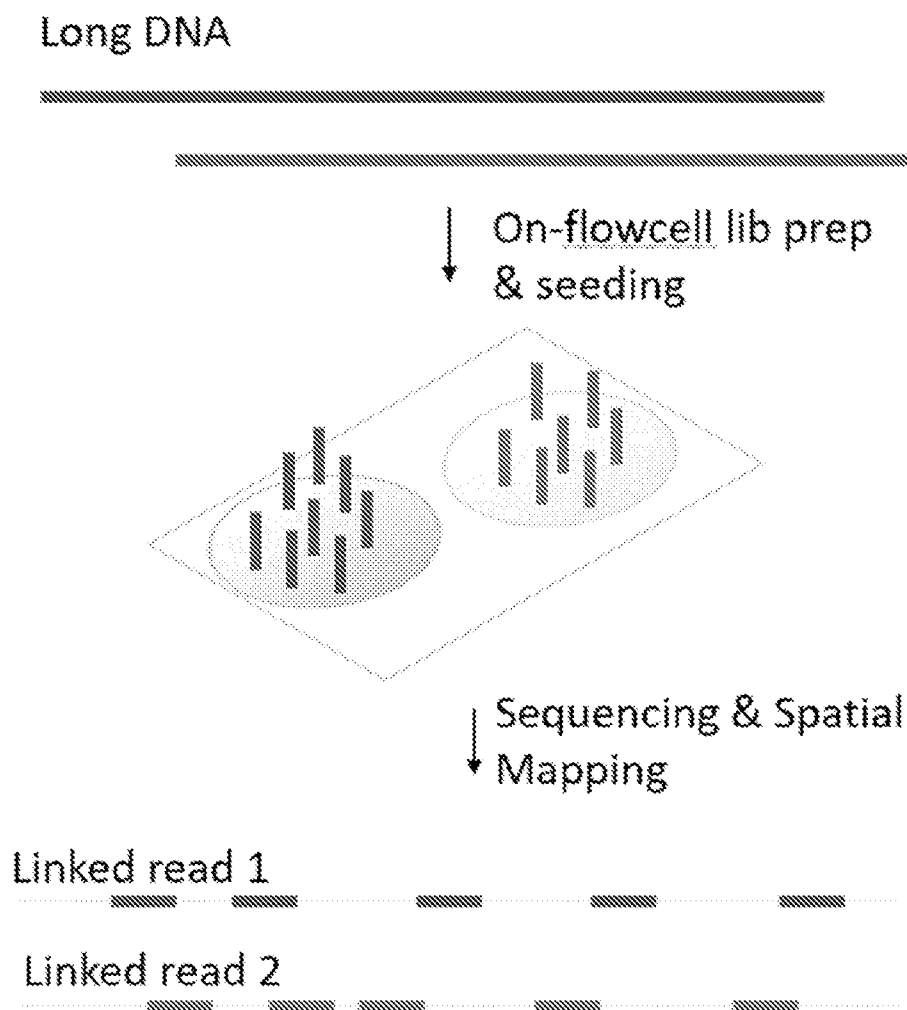


FIG. 1A

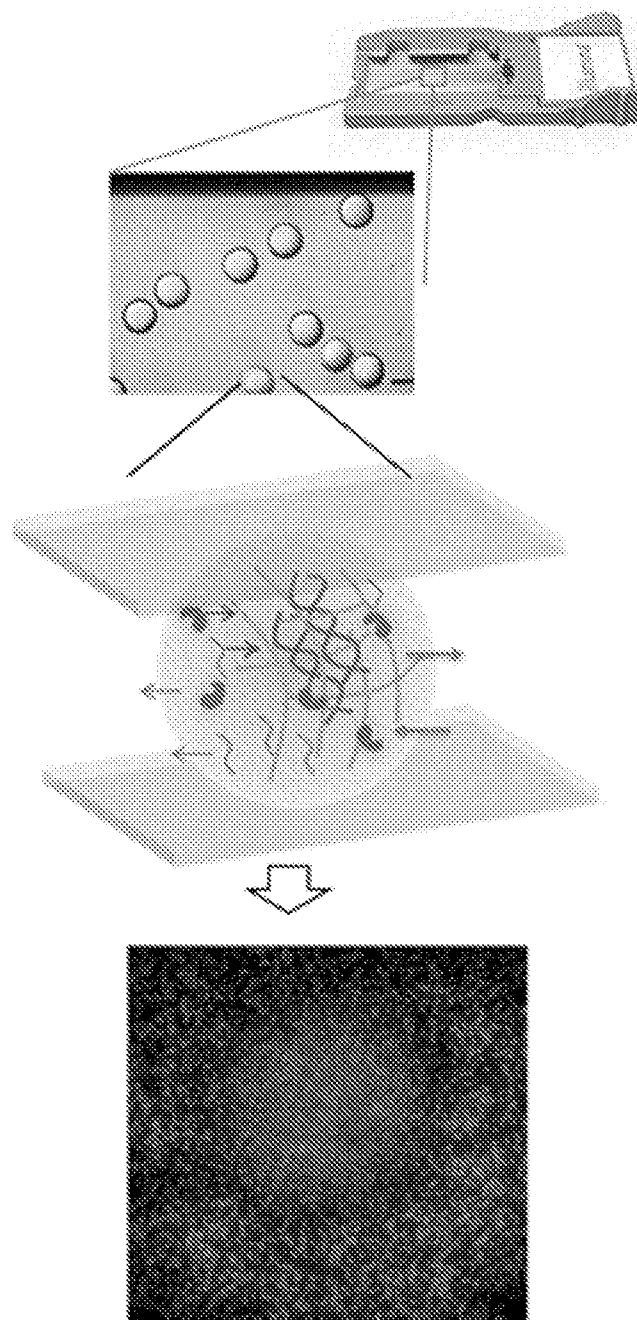


FIG. 1B

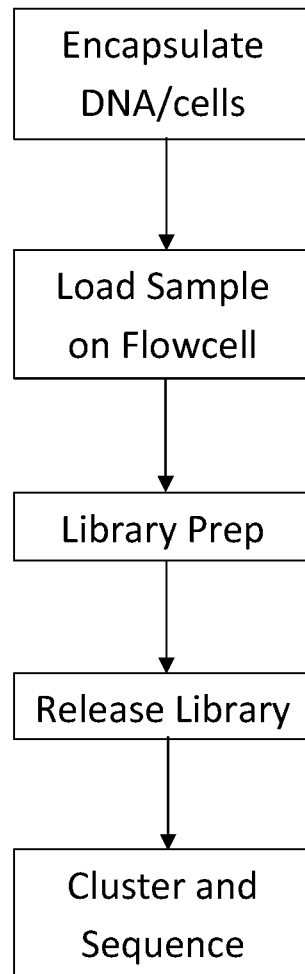


FIG. 2

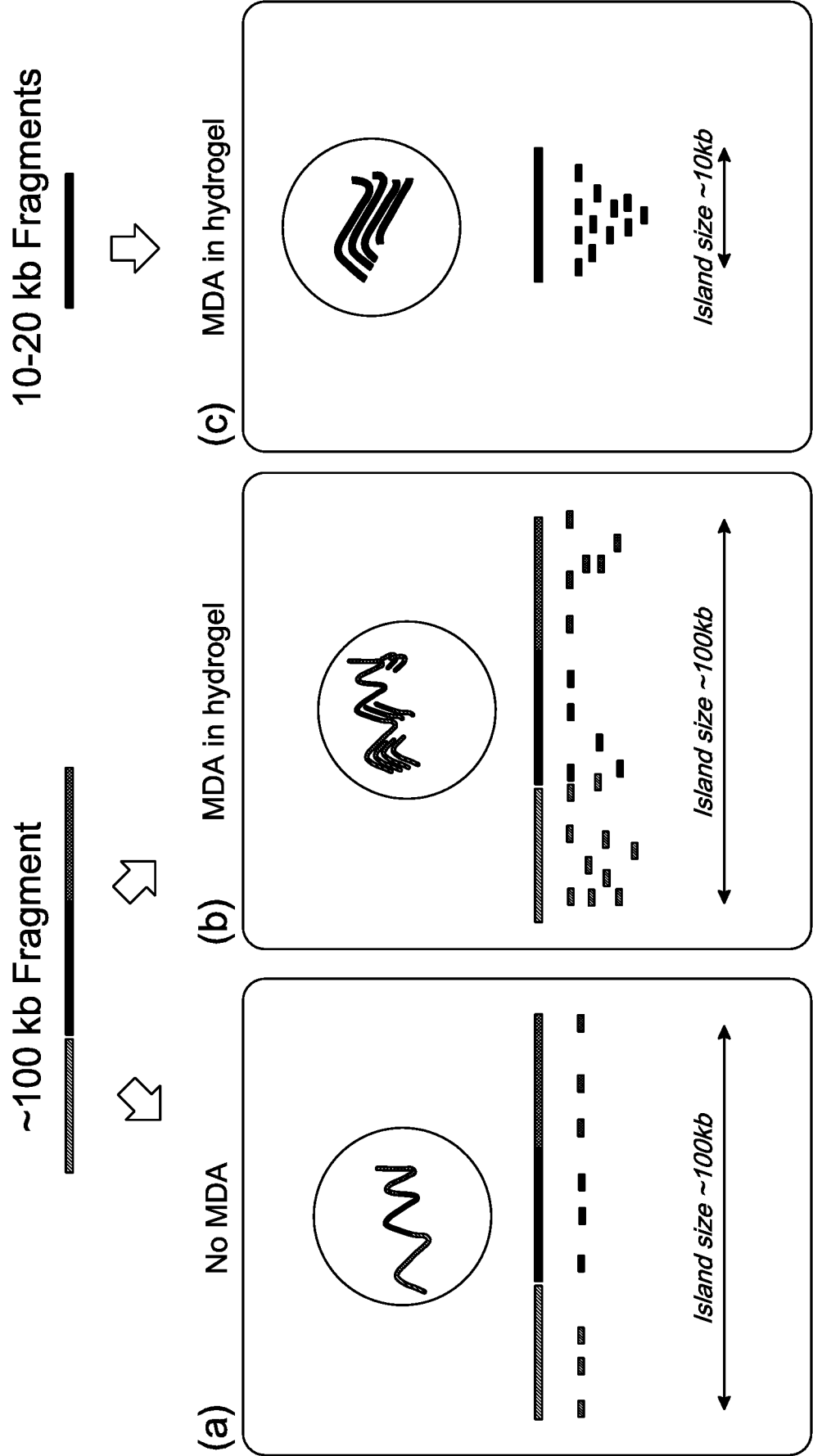
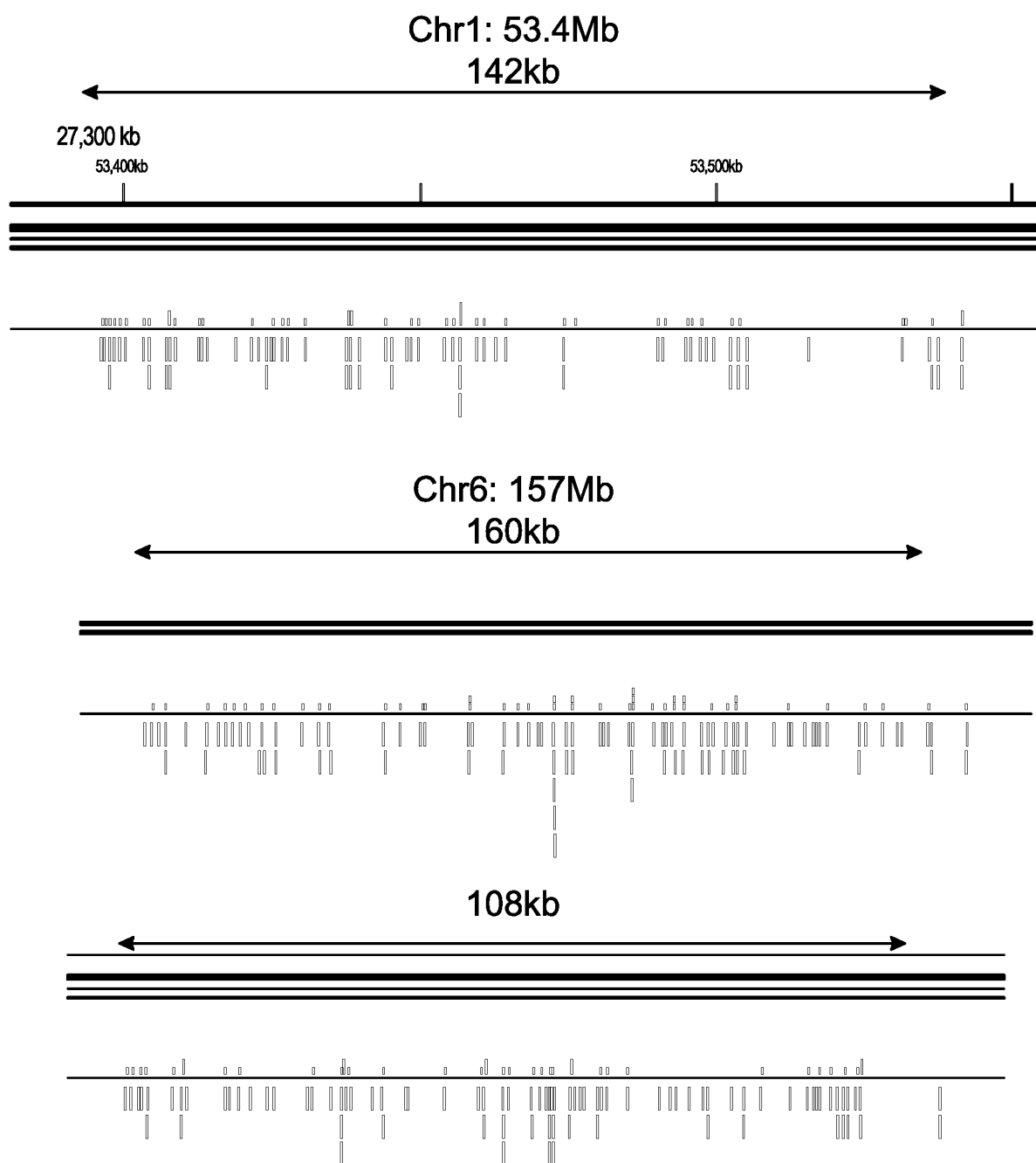


FIG. 3

*FIG. 4*

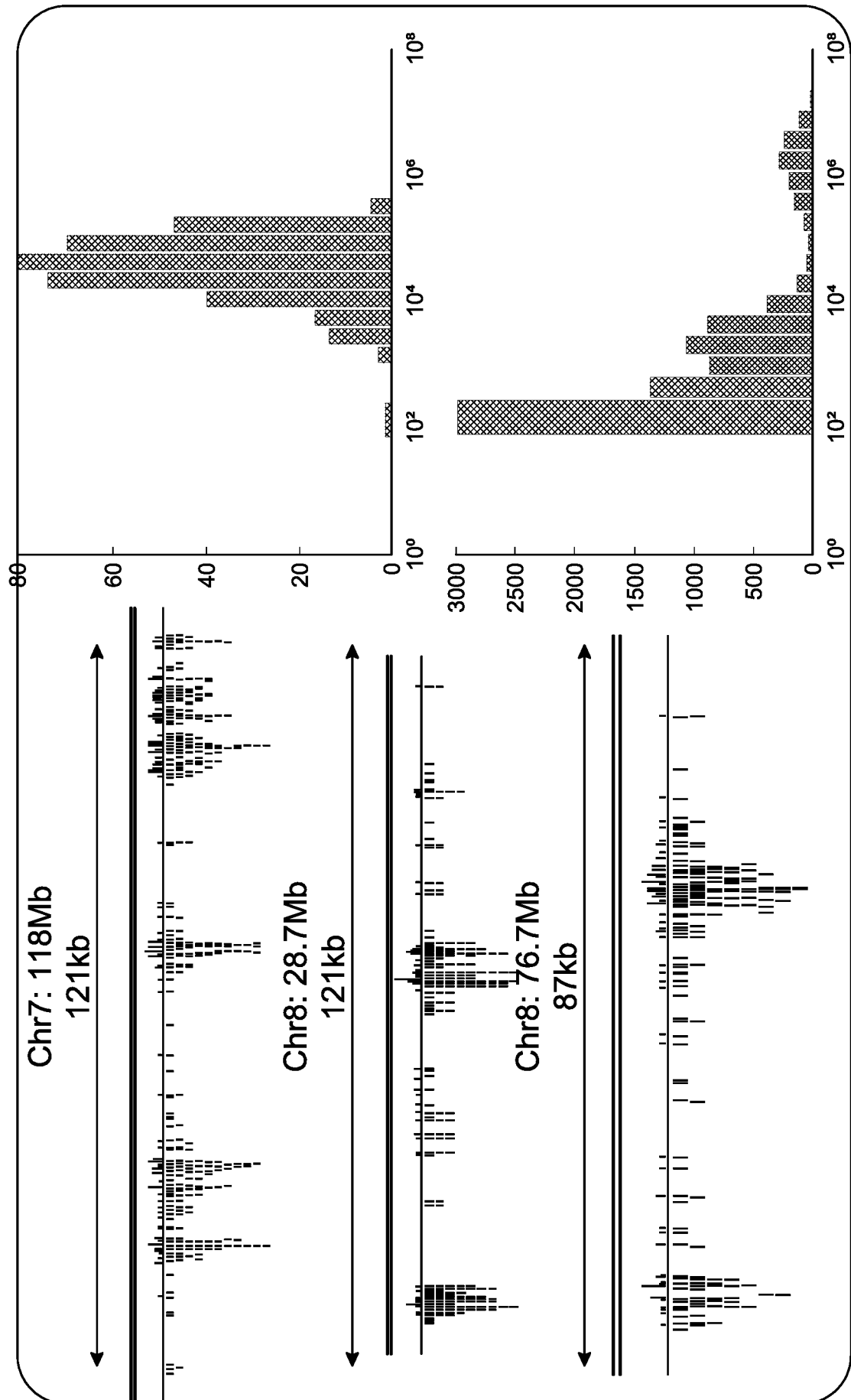


FIG. 5

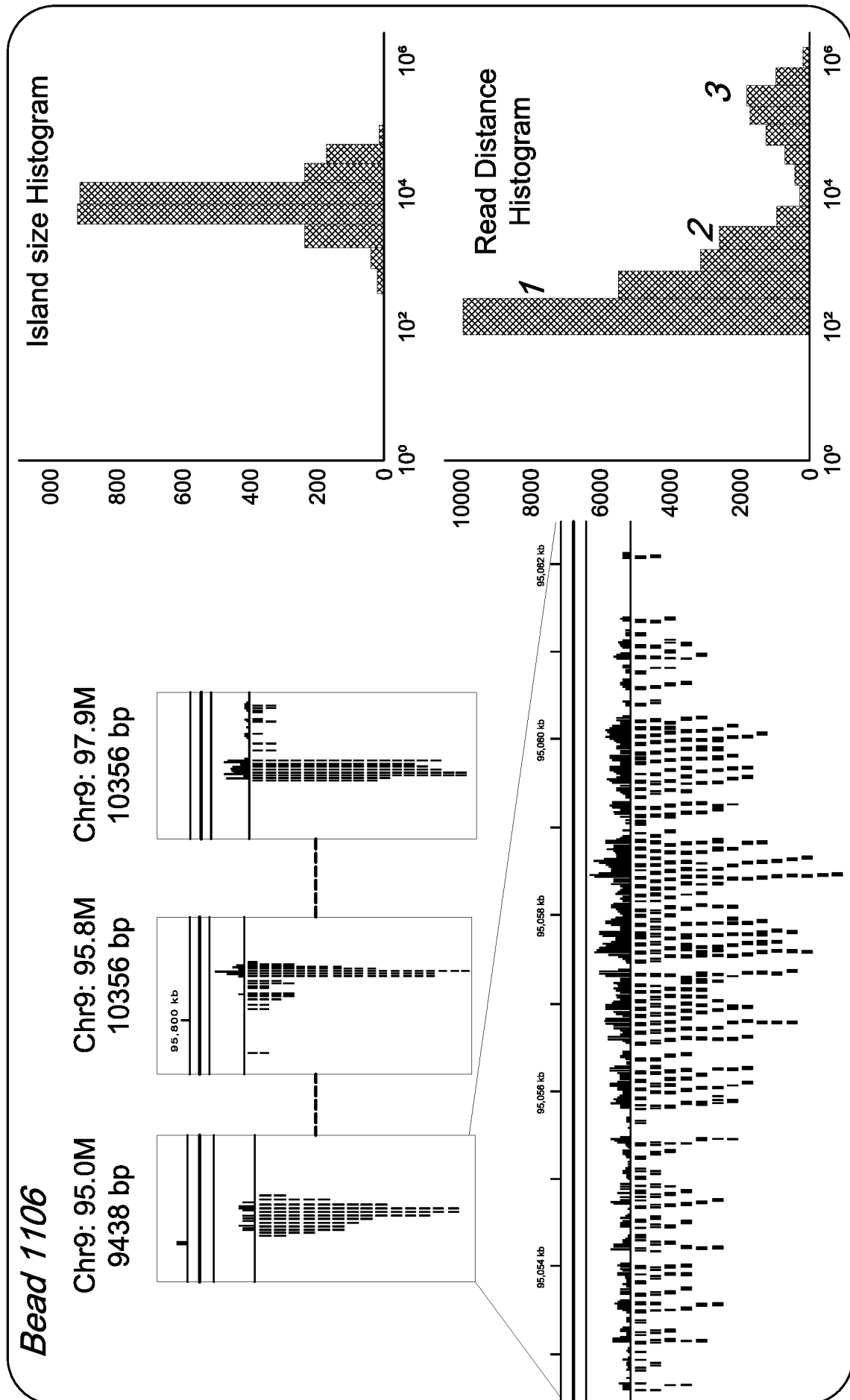
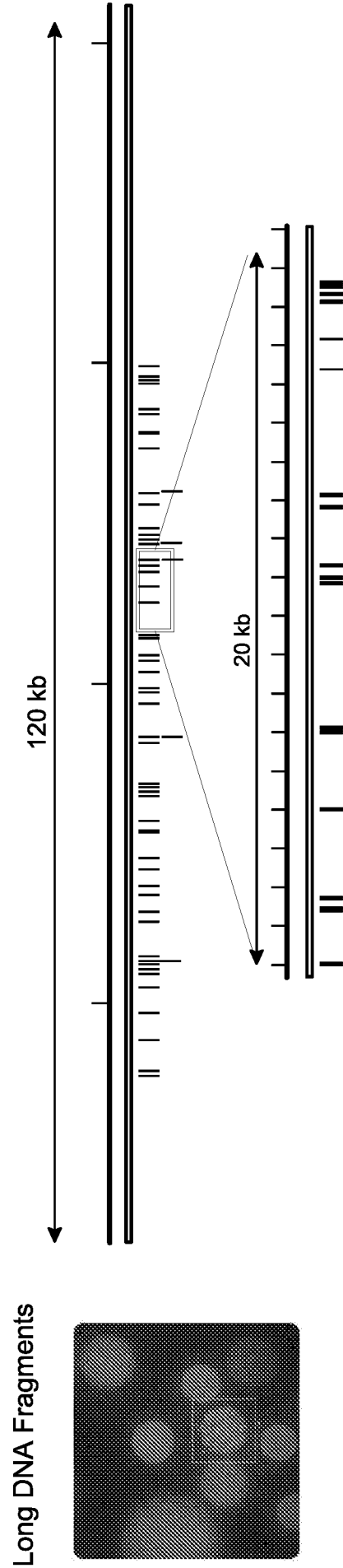
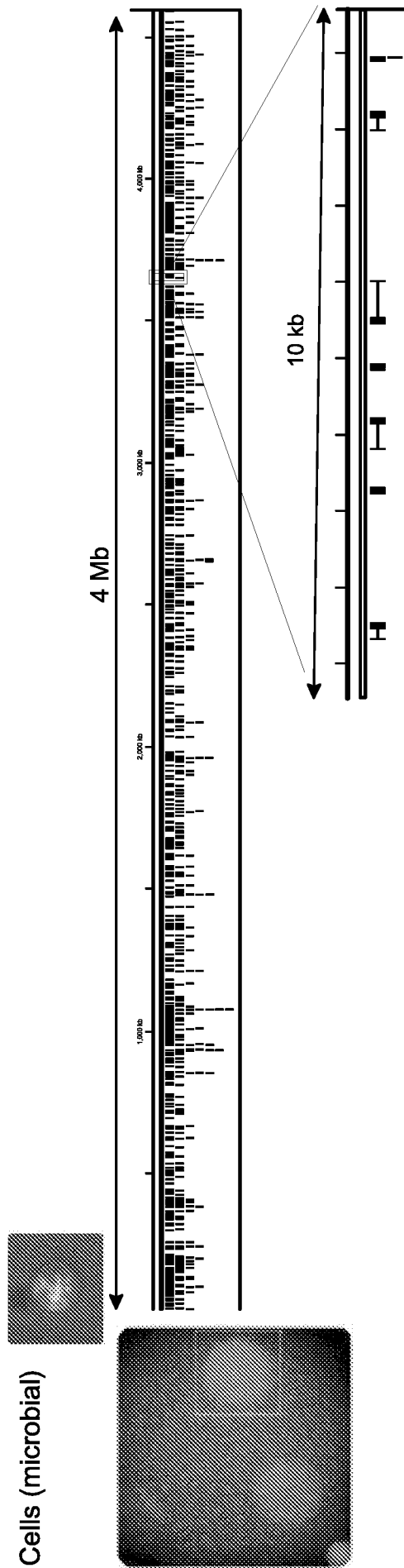


FIG. 6



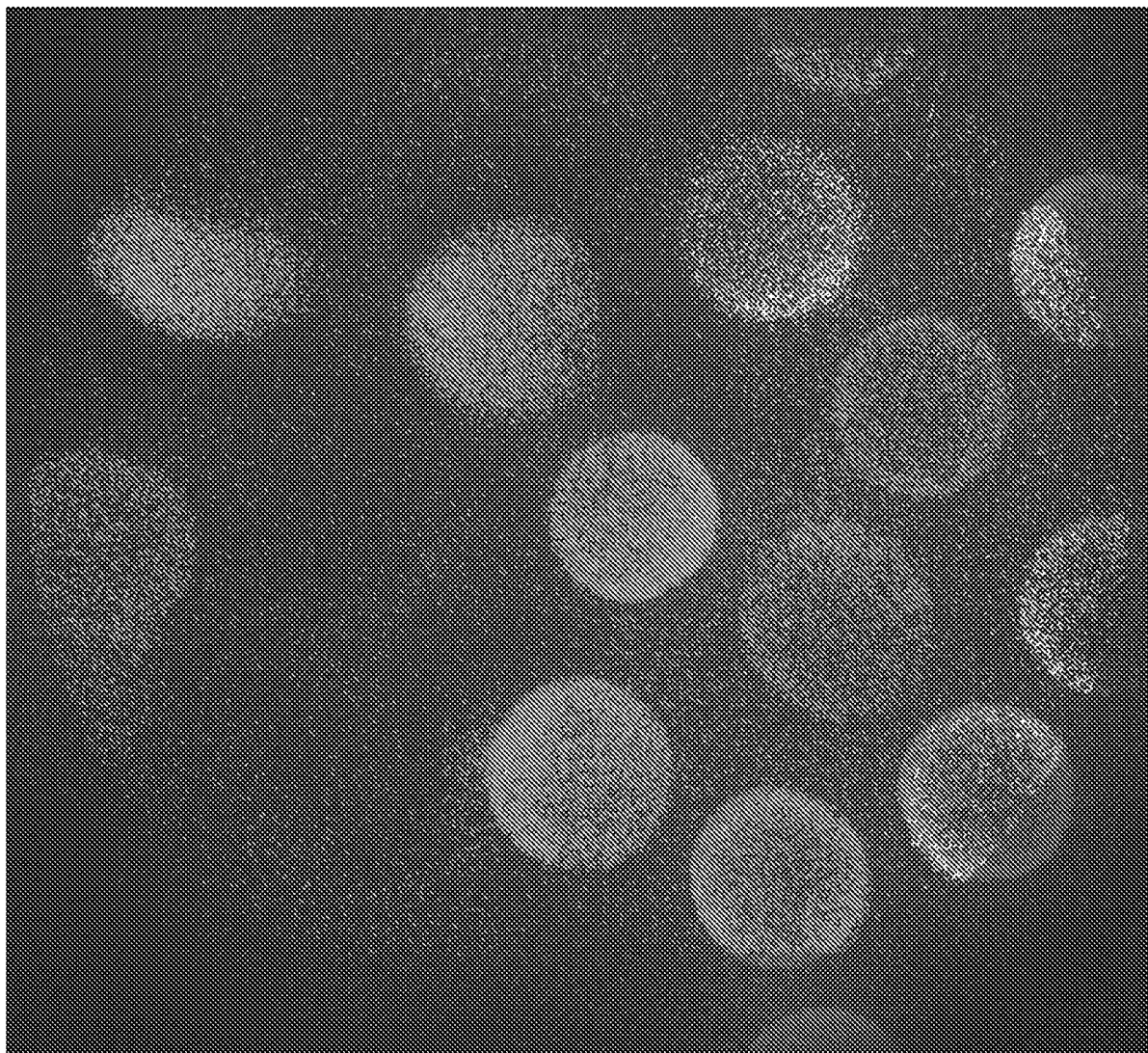
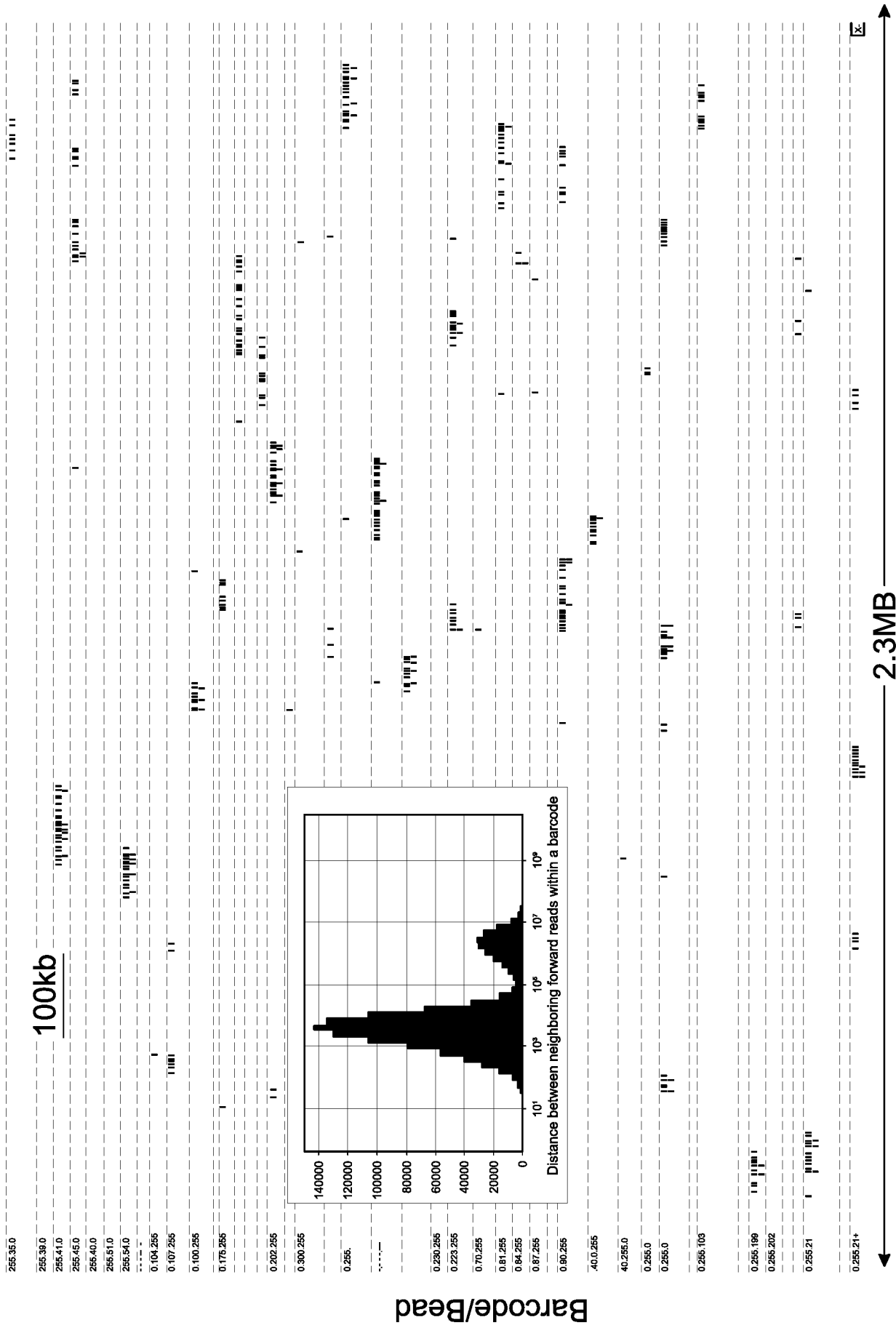


FIG. 8

FIG. 9



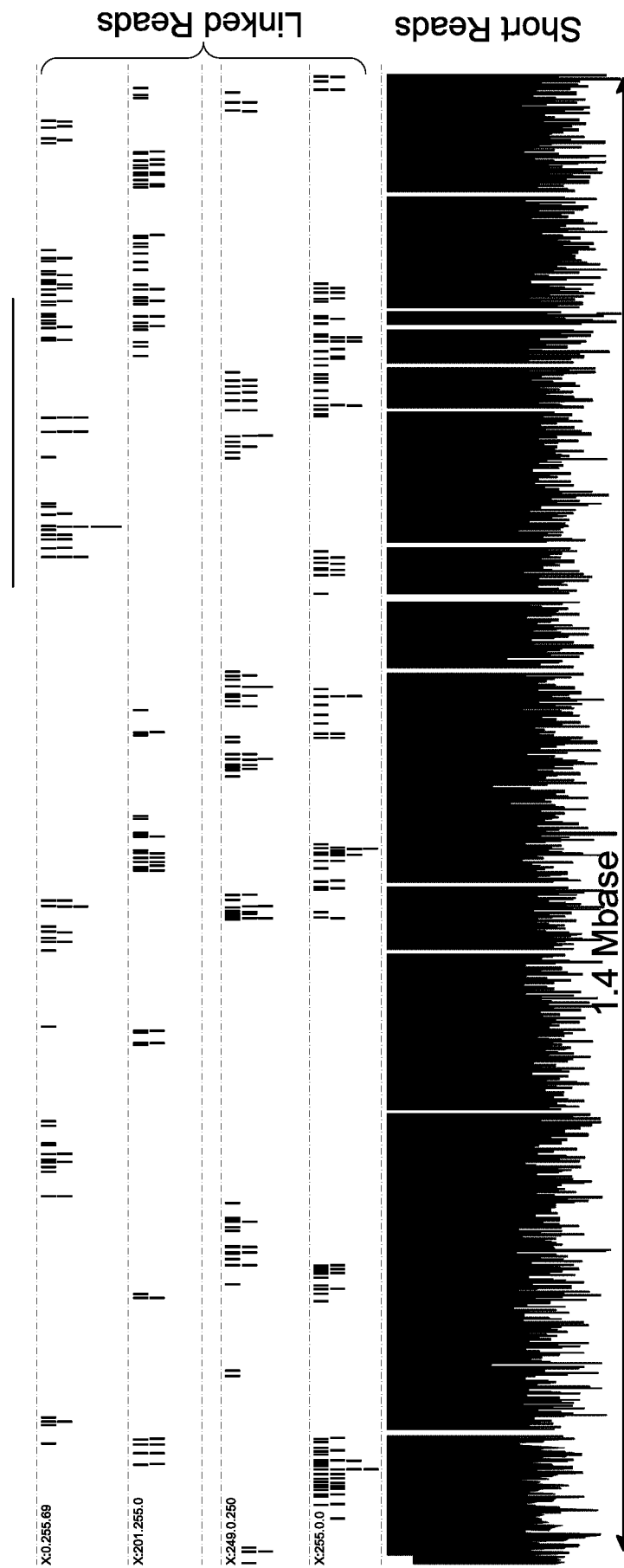


FIG. 10A

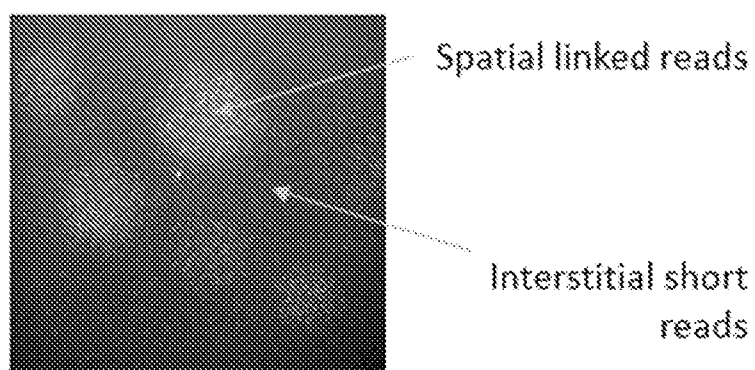


FIG. 10B

