(19) **DANMARK** (10) **DK/EP 2943589 T3**

(12) Oversættelse af
europæisk patentskrift

Patent- og
Varemærkestyrelsen

(73) Patenthaver: **Illumina Cambridge Limited, Chesterford Research Park, Little Chesterford, Nr Saffron Walden,
Essex CB10 1XL, Storbritannien**

(72) Opfinder: **GORMLEY, Niall, Anthony, Little Chesterford, Nr Saffron Walden, Essex CB10 1XL, Storbritannien
SMITH, Geoffrey, Paul, Little Chesterford, Nr Saffron Walden, Essex CB10 1XL, Storbritannien**

(74) Fuldmægtig i Danmark: **HEGNER & PARTNERS A/S, Symbion Science Park, Fruebjergvej 3, 2100 København Ø,
Danmark**

(54) Benævnelse: **FREMGANGSMÅDER OG FASTE BÆRERE TIL PRØVEPRÆPARERING UNDER ANVENDELSE AF
TRANSPOSOMER**

(56) Fremdragne publikationer:
EP-A1- 2 712 931
WO-A1-98/44151
WO-A1-2010/048605
WO-A1-2012/103545
WO-A2-2012/106546
US-A1- 2012 053 063
A. ADEY ET AL: "Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing", GENOME
RESEARCH, vol. 22, no. 6, 30 March 2012 (2012-03-30) , pages 1139-1143, XP055136909, ISSN: 1088-9051, DOI:
10.1101/gr.136242.111
CARUCCIO NICHOLAS: "Preparation of next-generation sequencing libraries using Nextera(TM) technology:
simultaneous DNA fragmentation and adaptor tagging by in vitro transposition", METHODS IN MOLECULAR
BIOLOGY, HUMANA PRESS, INC, US, vol. 733, 1 January 2011 (2011-01-01), pages 241-255, XP009157814,
ISSN: 1940-6029, DOI: 10.1007/978-1-61779-089-8_17 [retrieved on 2011-01-01]
SYED NAZAKAT ET AL: "Next-generation sequencing library preparation: simultaneous fragmentation and
tagging using in vitro transposition", APPLICATION NOTES. NATURE METHODS. EPICENTRE BIOTECH., 1

January 2009 (2009-01-01), pages I-II, XP055127060, Sage India: New Delhi, India

DESCRIPTION

BACKGROUND

[0001] There are a variety of methods and applications for which it is desirable to generate a library of fragmented and tagged DNA molecules from double-stranded DNA (dsDNA) target molecules. Often, the purpose is to generate smaller DNA molecules (e.g., DNA fragments) from larger dsDNA molecules for use as templates in DNA sequencing reactions.

[0002] Many of the methods currently used for fragmentation and tagging of double-stranded DNA for use in next-generation sequencing are wasteful of the DNA, require expensive instruments for fragmentation, and the procedures for fragmentation, tagging and recovering tagged DNA fragments are difficult, tedious, laborious, time-consuming, inefficient, costly, require relatively large amounts of sample nucleic acids. In addition, many of these methods generate tagged DNA fragments that are not fully representative of the sequences contained in the sample nucleic acids from which they were generated. Thus, what is needed in the art are methods that provide speed and ease of use when generating libraries of tagged DNA fragments from target DNA and which can be easily applied to nucleic acid analysis methods such as next-generation sequencing and amplification methods.

[0003] WO 2012/1106546 A2 discloses methods that include treating a target DNA sequence with a transposase resulting in one or more fragmentation or insertion events; adding or inserting one or more recognition sequences to the target DNA sequence (i) during the transposase treatment of (ii) during a subsequent amplification; sequencing the treated DNA; and capturing contiguity information by identifying target DNA sequences or recognition sequences having a shared property.

[0004] WO 2010/048605 A1 discloses methods, compositions and kits for using a transposase and a transposon end for generating extensive fragmentation and 5'-tagging of double-stranded target DNA in vitro, then using a DNA polymerase for generating 5'- and 3'-tagged single-stranded DNA fragments without performing a PCR amplification reaction, wherein the first tag on the 5'-ends exhibits the sequence of the transferred transposon end and optionally, an additional arbitrary sequence, and the second tag on the 3'-ends exhibits a different sequence from the sequence exhibited by the first tag.

[0005] WO 2012/103545 A1 discloses method for adding one or more tags to the double-stranded product of a tagmentation reaction. The method involves providing a double-stranded target nucleic acid and a transposome having a transposase with two transposon end sequences: a "transferred strand" and a "non-transferred strand". The transposome breaks the target nucleic acid into fragments while covalently transferring the transferred strand to a first strand of the fragment; the non-transferred strand of the transposome remains hybridized to the transferred strand.

[0006] Adey et al. disclose an ultra-low-input, tagmentation-based whole-genome bisulfite sequencing method in Genome Research, 22:1139-1143, 2012.

[0007] US 2012/0530630 A1 discloses a method for sequencing polynucleotide molecules. The method includes the steps of providing a plurality of polynucleotide molecules attached to a surface, wherein a first portion of each polynucleotide molecule is attached to a first location of the surface and a second portion of each polynucleotide molecule is attached to a second location of the surface, the relative proximity of the first and second locations being correlated with the probability that the first and second portions are paired, separating the first and second portions of the polynucleotide molecules on the surface, determining the sequences of the first and second portions of the polynucleotide molecules and comparing the relative proximities and the sequences to determine which first and second portions are paired and to determine the sequence of the target polynucleotide molecules.

BRIEF SUMMARY

[0008] Presented herein are methods and compositions for nucleic acid sample preparation on a solid support. The methods and compositions especially relate to methods and compositions for fragmenting and tagging DNA using transposon compositions immobilized to a solid support. The methods and compositions presented herein are useful, for example, for generating libraries of tagged DNA fragments for use, e.g., in next generation sequencing methods, and the like. In some preferred embodiments, the present invention relates to preparation of linear ssDNA fragments on a solid support from target DNA comprising any dsDNA of interest (including double-stranded cDNA prepared from RNA), from any source, for genomic, subgenomic, transcriptomic, or metagenomic analysis, or analysis of RNA expression.

[0009] The invention is defined by the appended claims. Accordingly, presented herein are methods as defined in the claims of preparing an immobilized library of tagged DNA fragments comprising: (a) providing a solid support having transposome complexes immobilized thereon, wherein the transposome complexes comprise a transposase bound to a first polynucleotide, the first polynucleotide comprising (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain; and (b) applying a target DNA to the solid support under conditions whereby the target DNA is fragmented by the transposome complexes, and the 3' transposon end sequence of the first polynucleotide is transferred to a 5' end of at least one strand of the fragments; thereby producing an immobilized library of double-stranded fragments wherein at least one strand is 5'- tagged with the first tag. The transposome complexes comprise a second polynucleotide comprising a region complementary to said transposon end sequence. The methods can further comprise (c) providing transposome complexes in solution and contacting the transposome complexes with the immobilized fragments under conditions whereby the target DNA is fragmented by the transposome complexes in solution; thereby obtaining immobilized nucleic acid fragments having one end in

solution. In some embodiments, the transposome complexes in solution can comprise a second tag, such that the method generates immobilized nucleic acid fragments having a second tag, the second tag in solution. The first and second tags can be different or the same.

[0010] Also presented herein are solid supports as defined in the claims having a library of tagged DNA fragments immobilized thereon prepared according to the above methods or other methods. For example, presented herein are solid supports having transposome complexes immobilized thereon, wherein the transposome complexes comprise a transposase bound to a first polynucleotide, the polynucleotide comprising (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain.

[0011] Also disclosed herein are methods of generating a flowcell, comprising immobilizing a plurality of transposome complexes to a solid support, the transposome complexes comprising a transposase bound to a first polynucleotide, the first polynucleotide comprising (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain.

[0012] The methods can further comprise providing a solid support having a plurality the first polynucleotides immobilized thereon, and contacting the solid support with transposase holoenzyme and a second polynucleotide, the second polynucleotide comprising a region complementary to the transposon end sequence. In some methods, immobilizing comprises (a) providing a solid support having amplification primers coupled thereto; (b) hybridizing a second polynucleotide to one of the amplification primers, the second oligonucleotide comprising a region complementary to a transposon end sequence and a region complementary to the first tag; (c) extending the amplification primer using a polymerase to generate a duplex comprising the first polynucleotide hybridized to the second polynucleotide, the first polynucleotide immobilized directly to the solid support; and (d) contacting the solid support with transposase holoenzyme, thereby assembling a transposome complex on the solid support.

[0013] Also disclosed herein is a population of microparticles having transposome complexes immobilized thereto, the transposome complexes comprising a transposase bound to a first polynucleotide and a second polynucleotide; wherein the first polynucleotide is immobilized at its 5' end to the surface of the microparticle and the second polynucleotide is hybridized to the 3' end of the first polynucleotide; and wherein the first polynucleotide comprises: (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain. Also presented herein are methods of producing an immobilized library of tagged DNA fragments comprising contacting a target DNA with the above population of microparticles to generate immobilized tagged DNA fragments.

[0014] Also disclosed herein are methods of generating a library tagged DNA fragments for index-directed assembly into a longer sequence read, the method comprising: providing a population of microparticles having transposome complexes immobilized thereto, the transposome complexes comprising a transposase bound to a first polynucleotide comprising an index domain associated with the microparticle and a second polynucleotide; applying a

target DNA to the population of microparticles, thereby generating immobilized DNA fragments that are tagged with the index domain. In the above methods, the first polynucleotide can be immobilized at its 5' end to the surface of the microparticle and the second polynucleotide can be hybridized to the 3' end of the first polynucleotide; and wherein the first polynucleotide comprises: (i) a 3' portion comprising a transposon end sequence, and (ii) the index domain; and wherein the population of microparticles comprises at least a plurality of index domains, and wherein the first polynucleotides on an individual microparticle share the same index domain.

Also disclosed herein is a method for sequencing a plurality of target DNA molecules, comprising: applying a plurality of target DNA to a solid support having transposome complexes immobilized thereon under conditions whereby the target DNA is fragmented by the transposome complexes; thereby producing an immobilized library of double-stranded fragments, wherein a first portion of each target DNA is attached to said solid support at a first location on said solid support and a second portion of each target DNA is attached to said solid support at a second location on said solid support; and mapping said immobilized library of double-stranded fragments to generate a set of locations that are linked by each target DNA; determining the sequences of said first and second portions of the target DNA; and correlating said set of locations to determine which first and second portions are linked by said target DNA and to determine the sequence of the target DNA molecules.

[0015] In some embodiments of the methods and compositions presented herein, the transposome complexes are present on the solid support at a density of at least 10^3, 10^4, 10^5, 10^6 complexes per mm^2. In some embodiments, the transposome complex comprises a hyperactive transposase, such as Tn5 transposase.

[0016] In some embodiments of the methods and compositions presented herein, the tag domain can comprise, for example, a region for cluster amplification. In some embodiments, the tag domain can comprise a region for priming a sequencing reaction.

[0017] In some embodiments of the methods and compositions presented herein, the solid support can comprise, for example, microparticles, a patterned surface, wells and the like. In some embodiments, the transposome complexes are randomly distributed upon the solid support. In some embodiments, the transposome complexes are distributed on a patterned surface.

[0018] The details of one or more embodiments are set forth in the accompanying drawings and the description below. Other features, objects, and advantages will be apparent from the description and drawings, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019]

Figure 1a shows the general concept according to one embodiment. Figure 1b shows the general concept according to one embodiment.

Figure 2a is a schematic illustrating a tagmentation reaction.

Figure 2b is a schematic with the resulting bridge redrawn to clarify the nature of the resulting bridge.

Figure 3a illustrates an embodiment where two forms of transposome are assembled on the surface of a flowcell. Addition of DNA to the flowcell results in tagmentation and coupling of the DNA to the transposomes. Also shown in Figure 3a are different types of resulting clusters: P7:P7, P5:P5, and P5:P7 clusters. 3b illustrates an embodiment where two forms of transposome are assembled on the surface of a flowcell.

Figure 4a illustrates another embodiment. In Figure 4a, only one form of surface bound transposome (e.g. P5 transposome) is present, resulting in bridges having the same tag sequence at each end. Additional transposomes are added to further fragment the bridge structures and incorporate an additional tag sequence (P7).

Figure 4b shows amplification to generate pairs of clusters resulting from each transposome stump which represent two adjacent fragments in the original intact DNA sample (Figure 4b).

Figures 5a, 5b, 5c and 5d illustrate different methods for assembling surface bound transposome complexes. Figure 5a shows one embodiment of this method for assembling the surface bound transposome complexes. Figure 5b shows transposome complexes are assembled in solution and immobilizing comprises a further step of ligating the first polynucleotide to a splint oligonucleotide coupled to the solid support. Figure 5c shows a transposome dimer is assembled by hybridizing a looped oligonucleotide to an immobilized first polynucleotide. Figure 5d shows transposome complexes can be assembled on a standard paired end flow cell with amplification primers immobilized thereto.

Figure 6a sets forth the design for an experiment set forth as Example 1. Figure 6b sets forth the design for an experiment set forth as Example 1.

Figure 7 sets forth representative data obtained in an experiment conducted according to the design set forth in Example 1.

Figure 8 sets forth representative data obtained in an experiment conducted according to the design set forth in Example 1.

Figure 9 is an illustration of assembly of bead-bound transposomes and subsequent tagmentation according to one embodiment.

Figure 10 is an illustration of assembly of bead-bound transposomes and subsequent tagmentation according to one embodiment.

Figure 11 is an illustration of assembly of bead-bound transposomes and subsequent tagmentation according to one embodiment.

Figure 12 is an illustration of assembly of bead-bound transposomes and subsequent tagmentation according to one embodiment.

Figure 13 is an illustration of surface bound transposomes according to one embodiment.

Figure 14 is an illustration of tagmentation performed on bead-bound transposomes according to one embodiment.

Figure 15 is an illustration of tagmentation and barcoding of sub-fragments performed on bead-bound transposomes according to one embodiment.

Figure 16 is an illustration of tagmentation and barcoding of sub-fragments performed on bead-bound transposomes according to one embodiment.

Figure 17 is an illustration of transposome assembly as described in Example 3.

Figure 18a is an illustration of transposome assembly as described in Example 3.

Figure 18b sets forth results using transposome assembly as described in Example 3.

DETAILED DESCRIPTION

[0020] Current protocols for sequencing nucleic acid samples routinely employ a sample preparation process that converts DNA or RNA into a library of templates. These methods can result in loss of DNA sample and often require expensive instruments for fragmentation. In addition, the sample preparation methods are often difficult, tedious, and inefficient.

[0021] In standard sample preparation methods, each template contains an adaptor at either end of the insert and often a number of steps are required to both modify the DNA or RNA and to purify the desired products of the modification reactions. These steps are performed in solution prior to the addition of the adapted fragments to a flowcell where they are coupled to the surface by a primer extension reaction that copies the hybridized fragment onto the end of a primer covalently attached to the surface. These 'seeding' templates then give rise to monoclonal clusters of copied templates through several cycles of amplification.

[0022] The number of steps required to transform DNA into adaptor-modified templates in solution ready for cluster formation and sequencing can be minimized by the use of transposase mediated fragmentation and tagging. This process, referred to herein as 'tagmentation,' often involves 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 are added to the ends of the adapted

fragments by PCR.

[0023] Solution-based tagmentation has drawbacks and requires several labor-intensive steps. Additionally, bias can be introduced during PCR amplification steps. The methods and compositions presented herein overcome those drawbacks and allow unbiased sample preparation, cluster formation and sequencing to occur on a single solid support with minimal requirements for sample manipulation or transfer.

[0024] The present disclosure relates to the surprising discovery that transposome complexes pre-coupled to the surface of a flowcell can effectively fragment, tag and immobilize intact DNA within the flowcell. In specific embodiments, one or more of the strands that comprise the transposome adaptors are attached to the surface of the flowcell via their 5' end. When intact DNA is pumped onto the flowcell, the tagmentation reaction occurs in the same manner as occurs in solution-based tagmentation reactions, but the resulting product fragments are physically attached to the surface of the flowcell by their ends. The transposome adaptor sequences can contain sequences that enable subsequent cluster generation and sequencing.

[0025] The methods and compositions presented herein provide several advantages over solution-based tagmentation methods. For example, purified, partially purified or even unpurified intact DNA template can be loaded directly onto a flowcell for generation of clusters, without prior sample preparation. In addition, the contiguity of sequence information in the original intact DNA can be physically preserved by the juxtaposition of tagmented fragments on the surface of the flowcell. As a further advantage, DNA is physically linked to the surface of the flowcell so purification of reagents following further manipulation of the DNA can be achieved by flow-through buffer exchange in the flowcell channel.

Tagmentation on a solid support

[0026] In accordance with the above, presented herein are methods as defined in the claims of preparing an immobilized library of tagged DNA fragments. The methods comprise: (a) providing a solid support having transposome complexes immobilized thereon, wherein the transposome complexes comprise a transposase bound to a first polynucleotide, the first polynucleotide comprising (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain; and (b) applying a target DNA to the solid support under conditions whereby the target DNA is fragmented by the transposome complexes, and the 3' transposon end sequence of the first polynucleotide is transferred to a 5' end of at least one strand of the fragments; thereby producing an immobilized library of double-stranded fragments wherein at least one strand is 5'- tagged with the first tag.

[0027] As used herein, the term "transposome complex" refers generally to a transposase enzyme non-covalently bound to a double stranded nucleic acid. For example, the complex can be a transposase enzyme preincubated with double-stranded transposon DNA under conditions that support non-covalent complex formation. Double-stranded transposon DNA can

include, without limitation, Tn5 DNA, a portion of Tn5 DNA, a transposon end composition, a mixture of transposon end compositions or other double-stranded DNAs capable of interacting with a transposase such as the hyperactive Tn5 transposase.

[0028] A "transposase" means an enzyme that is capable of forming a functional complex with a transposon end-containing composition (e.g., transposons, transposon ends, transposon end compositions) and catalyzing insertion or transposition of the transposon end-containing composition into the double-stranded target DNA with which it is incubated, for example, in an *in vitro* transposition reaction. A transposase as presented herein can also include integrases from retrotransposons and retroviruses. Transposases, transposomes and transposome complexes are generally known to those of skill in the art, as exemplified by the disclosure of US 2010/0120098. Although many embodiments described herein refer to Tn5 transposase and/or hyperactive Tn5 transposase, it will be appreciated that any transposition system that is capable of inserting a transposon end with sufficient efficiency to 5'-tag and fragment a target DNA for its intended purpose can be used in the present invention. In particular embodiments, a preferred transposition system is capable of inserting the transposon end in a random or in an almost random manner to 5'-tag and fragment the target DNA.

[0029] The term "transposon end" refers to a double-stranded nucleic acid DNA that exhibits only the nucleotide sequences (the "transposon end sequences") that are necessary to form the complex with the transposase or integrase enzyme that is functional in an *in vitro* transposition reaction. In some embodiments, a transposon end is capable of forming a functional complex with the transposase in a transposition reaction. As non-limiting examples, transposon ends can include the 19-bp outer end ("OE") transposon end, inner end ("IE") transposon end, or "mosaic end" ("ME") transposon end recognized by a wild-type or mutant Tn5 transposase, or the R1 and R2 transposon end as set forth in the disclosure of US 2010/0120098. Transposon ends can comprise any nucleic acid or nucleic acid analogue suitable for forming a functional complex with the transposase or integrase enzyme in an *in vitro* transposition reaction. For example, the transposon end can comprise DNA, RNA, modified bases, non-natural bases, modified backbone, and can comprise nicks in one or both strands. Although the term "DNA" is used throughout the present disclosure in connection with the composition of transposon ends, it should be understood that any suitable nucleic acid or nucleic acid analogue can be utilized in a transposon end.

[0030] The term "transferred strand" refers to the transferred portion of both transposon ends. Similarly, the term "non-transferred strand" refers to the non-transferred portion of both "transposon ends." The 3'-end of a transferred strand is joined or transferred to target DNA in an *in vitro* transposition reaction. The non-transferred strand, which exhibits a transposon end sequence that is complementary to the transferred transposon end sequence, is not joined or transferred to the target DNA in an *in vitro* transposition reaction.

[0031] In some embodiments, the transferred strand and non-transferred strand are covalently joined. For example, in some embodiments, the transferred and non-transferred strand sequences are provided on a single oligonucleotide, e.g., in a hairpin configuration. As such,

although the free end of the non-transferred strand is not joined to the target DNA directly by the transposition reaction, the non-transferred strand becomes attached to the DNA fragment indirectly, because the non-transferred strand is linked to the transferred strand by the loop of the hairpin structure. Additional examples of transposome structure and methods of preparing and using transposomes can be found in the disclosure of US 2010/0120098.

[0032] The terms "tag" and "tag domain" as used herein refer to a portion or domain of a polynucleotide that exhibits a sequence for a desired intended purpose or application. Some embodiments presented herein include a transposome complex comprising a polynucleotide having a 3' portion comprising a transposon end sequence, and tag comprising a tag domain. Tag domains can comprise any sequence provided for any desired purpose. For example, in some embodiments, a tag domain comprises one or more restriction endonuclease recognition sites. In some embodiments, a tag domain comprises one or more regions suitable for hybridization with a primer for a cluster amplification reaction. In some embodiments, a tag domain comprises one or more regions suitable for hybridization with a primer for a sequencing reaction. It will be appreciated that any other suitable feature can be incorporated into a tag domain. In some embodiments, the tag domain comprises a sequence having a length between 5 and 200 bp. In some embodiments, the tag domain comprises a sequence having a length between 10 and 100 bp. In some embodiments, the tag domain comprises a sequence having a length between 20 and 50 bp. In some embodiments, the tag domain comprises a sequence having a length between 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150 and 200 bp.

[0033] In the methods and compositions presented herein, transposome complexes arc immobilized to the solid support. In some embodiments, the transposome complexes are immobilized to the support via one or more polynucleotides, such as a polynucleotide comprising a transposon end sequence. In some embodiments, the transposome complex may be immobilized via a linker molecule coupling the transposase enzyme to the solid support. In some embodiments, both the transposase enzyme and the polynucleotide are immobilized to the solid support. When referring to immobilization of molecules (e.g. nucleic acids) to a solid support, the terms "immobilized" and "attached" are used interchangeably herein and both terms are intended to encompass direct or indirect, covalent or non-covalent attachment, unless indicated otherwise, either explicitly or by context. In certain embodiments of the invention covalent attachment may be preferred, but generally all that is required is that the molecules (e.g. nucleic acids) remain immobilized or attached to the support under the conditions in which it is intended to use the support, for example in applications requiring nucleic acid amplification and/or sequencing.

[0034] Certain embodiments of the invention may make use of solid supports comprised of an inert substrate or matrix (e.g. glass slides, polymer beads etc.) which has been functionalized, for example by application of a layer or coating of an intermediate material comprising reactive groups which permit covalent attachment to biomolecules, such as polynucleotides. Examples of such supports include, but are not limited to, polyacrylamide hydrogels supported on an inert substrate such as glass, particularly polyacrylamide hydrogels as described in WO

2005/065814 and US 2008/0280773. In such embodiments, the biomolecules (e.g. polynucleotides) may be directly covalently attached to the intermediate material (e.g. the hydrogel) but the intermediate material may itself be non-covalently attached to the substrate or matrix (e.g. the glass substrate). The term "covalent attachment to a solid support" is to be interpreted accordingly as encompassing this type of arrangement.

[0035] The terms "solid surface," "solid support" and other grammatical equivalents herein refer to any material that is appropriate for or can be modified to be appropriate for the attachment of the transposome complexes. As will be appreciated by those in the art, the number of possible substrates is very large. Possible substrates include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon™, etc.), polysaccharides, nylon or nitrocellulose, ceramics, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, optical fiber bundles, and a variety of other polymers. Particularly useful solid supports and solid surfaces for some embodiments are located within a flow cell apparatus. Exemplary flow cells are set forth in further detail below.

[0036] In some embodiments, the solid support comprises a patterned surface suitable for immobilization of transposome complexes in an ordered pattern. A "patterned surface" refers to an arrangement of different regions in or on an exposed layer of a solid support. For example, one or more of the regions can be features where one or more transposome complexes are present. The features can be separated by interstitial regions where transposome complexes are not present. In some embodiments, the pattern can be an x-y format of features that are in rows and columns. In some embodiments, the pattern can be a repeating arrangement of features and/or interstitial regions. In some embodiments, the pattern can be a random arrangement of features and/or interstitial regions. In some embodiments, the transposome complexes are randomly distributed upon the solid support. In some embodiments, the transposome complexes are distributed on a patterned surface. Exemplary patterned surfaces that can be used in the methods and compositions set forth herein are described in US Ser. No. 13/661,524 or US Pat. App. Publ. No. 2012/0316086 A1.

[0037] In some embodiments, the solid support comprises an array of wells or depressions in a surface. This may be fabricated as is generally known in the art using a variety of techniques, including, but not limited to, photolithography, stamping techniques, molding techniques and microetching techniques. As will be appreciated by those in the art, the technique used will depend on the composition and shape of the array substrate.

[0038] The composition and geometry of the solid support can vary with its use. In some embodiments, the solid support is a planar structure such as a slide, chip, microchip and/or array. As such, the surface of a substrate can be in the form of a planar layer. In some embodiments, the solid support comprises one or more surfaces of a flowcell. The term "flowcell" as used herein refers to a chamber comprising a solid surface across which one or more fluid reagents can be flowed. Examples of flowcells and related fluidic systems and

detection platforms that can be readily used in the methods of the present disclosure are described, for example, in Bentley et al., Nature 456:53-59 (2008), WO 04/018497; US 7,057,026; WO 91/06678; WO 07/123744; US 7,329,492; US 7,211,414; US 7,315,019; US 7,405,281, and US 2008/0108082.

[0039] In some embodiments, the solid support or its surface is non-planar, such as the inner or outer surface of a tube or vessel. In some embodiments, the solid support comprises microspheres or beads. By "microspheres" or "beads" or "particles" or grammatical equivalents herein is meant small discrete particles. Suitable bead compositions include, but are not limited to, plastics, ceramics, glass, polystyrene, methylstyrene, acrylic polymers, paramagnetic materials, thoria sol, carbon graphite, titanium dioxide, latex or cross-linked dextrans such as Sepharose, cellulose, nylon, cross-linked micelles and teflon, as well as any other materials outlined herein for solid supports may all be used. "Microsphere Detection Guide" from Bangs Laboratories, Fishers Ind. is a helpful guide. In certain embodiments, the microspheres are magnetic microspheres or beads.

[0040] The beads need not be spherical; irregular particles may be used. Alternatively or additionally, the beads may be porous. The bead sizes range from nanometers, i.e. 100 nm, to millimeters, i.e. 1 mm, with beads from about 0.2 micron to about 200 microns being preferred, and from about 0.5 to about 5 micron being particularly preferred, although in some embodiments smaller or larger beads may be used.

[0041] Figures 1a and 1b generally illustrate the method according to one embodiment. A solid support coated with grafted oligonucleotides is shown, some of which contain the ME sequences, will form active transposome complexes that are physically coupled to the solid support in the presence of Tn5. The density of these surface bound transposomes can be modulated by varying the density of the grafted oligonucleotides containing the ME sequence or by the amount of transposase added to the solid support. For example, in some embodiments, the transposome complexes are present on the solid support at a density of at least 10^3, 10^4, 10^5, or at least 10^6 complexes per mm^2.

[0042] When double stranded DNA is added to the solid support, the transposome complexes will tagment added DNA, thus generating ds fragments coupled at both ends to the surface. In some embodiments, the length of bridged fragments can be varied by changing the density of the transposome complexes on the surface. In certain embodiments, the length of the resulting bridged fragments is less than 100 bp, 200 bp, 300 bp, 400 bp, 500 bp, 600 bp, 700 bp, 800 bp, 900 bp, 1000 bp, 1100 bp, 1200 bp, 1300 bp, 1400 bp, 1500 bp, 1600 bp, 1700 bp, 1800 bp, 1900 bp, 2000 bp, 2100 bp, 2200 bp, 2300 bp, 2400 bp, 2500 bp, 2600 bp, 2700 bp, 2800 bp, 2900 bp, 3000 bp, 3100 bp, 3200 bp, 3300 bp, 3400 bp, 3500 bp, 3600 bp, 3700 bp, 3800 bp, 3900 bp, 4000 bp, 4100 bp, 4200 bp, 4300 bp, 4400 bp, 4500 bp, 4600 bp, 4700 bp, 4800 bp, 4900 bp, 5000 bp, 10000 bp, 30000 bp or less than 100,000 bp. In such embodiments, the bridged fragments can then be amplified into clusters using standard cluster chemistry, as exemplified by the disclosure of US Patent Nos. 7,985,565 and 7,115,400.

[0043] In some embodiments, the length of the templates is longer than what can be suitably amplified using standard cluster chemistry. For example, in some embodiments, the length of templates is longer than 100 bp, 200 bp, 300 bp, 400 bp, 500 bp, 600 bp, 700 bp, 800 bp, 900 bp, 1000 bp, 1100 bp, 1200 bp, 1300 bp, 1400 bp, 1500 bp, 1600 bp, 1700 bp, 1800 bp, 1900 bp, 2000 bp, 2100 bp, 2200 bp, 2300 bp, 2400 bp, 2500 bp, 2600 bp, 2700 bp, 2800 bp, 2900 bp, 3000 bp, 3100 bp, 3200 bp, 3300 bp, 3400 bp, 3500 bp, 3600 bp, 3700 bp, 3800 bp, 3900 bp, 4000 bp, 4100 bp, 4200 bp, 4300 bp, 4400 bp, 4500 bp, 4600 bp, 4700 bp, 4800 bp, 4900 bp, 5000 bp, 10000 bp, 30000 bp or longer than 100,000 bp. In such embodiments, then a second tagmentation reaction can be performed by adding transposomes from solution that further fragment the bridges, as illustrated, for example, in Figure 4a. The second tagmentation reaction can thus remove the internal span of the bridges, leaving short stumps anchored to the surface that can converted into clusters ready for further sequencing steps. In particular embodiments, the length of the template can be within a range defined by an upper and lower limit selected from those exemplified above.

[0044] Prior to cluster generation, the DNA immobilized by surface tagmentation can be imaged. For example, the immobilized DNA can be stained with an interchelating dye and imaged to preserve a record of the position of the backbone of the DNA molecule on the surface. Following cluster generation and sequencing, the coordinates of clusters can be associated with their position on the original backbone, thus assisting in alignment of reads along a molecule and genome assembly.

[0045] The step of applying a target DNA can comprise adding a biological sample to said solid support. The biological sample can be any type that comprises DNA and which can be deposited onto the solid surface for tagmentation. For example, the sample can comprise DNA in a variety of states of purification, including purified DNA. However, the sample need not be completely purified, and can comprise, for example, DNA mixed with protein, other nucleic acid species, other cellular components and/or any other contaminant. As demonstrated in Example 2 below, the biological sample comprises a mixture of DNA, protein, other nucleic acid species, other cellular components and/or any other contaminant present in approximately the same proportion as found *in vivo*. For example the components can be found in the same proportion as found in an intact cell. The biological sample can hava a 260/280 ratio of less than 2.0, 1.9, 1.8, 1.7, 1.6, 1.5, 1.4, 1.3, 1.2, 1.1, 1.0, 0.9, 0.8, 0.7, or less than 0.60. The biological sample can have a 260/280 ratio of at least 2.0, 1.9, 1.8, 1.7, 1.6, 1.5, 1.4, 1.3, 1.2, 1.1, 1.0, 0.9, 0.8, 0.7, or at least 0.60. Because the methods provided herein allow DNA to be bound to a solid support, other contaminants can be removed merely by washing the solid support after surface bound tagmentation occurs. The biological sample can comprise, for example, a crude cell lysate or whole cells. For example, a crude cell lysate that is applied to a solid support in a method set forth herein, need not have been subjected to one or more of the separation steps that are traditionally used to isolate nucleic acids from other cellular components. Exemplary separation steps are set forth in Maniatis et al., Molecular Cloning: A Laboratory Manual, 2d Edition, 1989, and Short Protocols in Molecular Biology, ed. Ausubel, et al.

[0046] Thus, the biological sample can comprise, for example, blood, plasma, serum, lymph,

mucus, sputum, urine, semen, cerebrospinal fluid, bronchial aspirate, feces, and macerated tissue, or a lysate thereof, or any other biological specimen comprising DNA. One advantage of the methods and compositions presented herein that a biological sample can be added to the flowcell and subsequent lysis and purification steps can all occur in the flowcell without further transfer or handling steps, simply by flowing the necessary reagents into the flowcell. Examples 1 and 2 below demonstrate successful application of crude cell lysates to the methods and compositions provided herein.

[0047] Figures 2a and 2b further illustrate the tagmentation reaction according to one embodiment. As shown in Figure 2a, transposomes comprise a dimer of Tn5 with each monomer binding a double stranded molecule, i.e. the ME adaptor. One strand of the ME adaptor is covalently attached to the surface of a flowcell. Transposomes bind target DNA and generate two nicks in the DNA backbone, 9 bases apart on either strand. Although Figure 2a shows a 9 bp gap between nicks, in other embodiments, the transposome can generate a gap of 7, 8, 9, 10, 11, or 12 bp between nicks. Only one of the two strands of each ME adaptor is ligated to the 5' strand at each nick position. This strand 'the transferred strand' is grafted to the surface of the flowcell via its 5' end. The resulting bridge of the surface tagmentation is redrawn in Figure 2b to clarify the nature of the resulting bridge.

[0048] Figures 3a and 3b illustrate an example of the invention put into practice. Two forms of transposome are assembled on the surface of a flowcell. The first form comprises a P7 transposome in which the 'transfer strand' of the ME adaptor comprises an extended oligonucleotide sequence linking the transposome to the surface including an amplification domain (P7) and a sequencing primer (S2). The second form comprises a P5 transposome in which the 'transfer strand' of the ME adaptor comprises an extended oligonucleotide sequence linking the transposome to the surface including an amplification domain (P5) and a sequencing primer (S1). Addition of DNA to the flowcell results in tagmentation and coupling of the DNA to the transposomes. Three types of bridges result: P5-P5, P7-P7 and P5-P7. Following cluster formation and linearization (Figure 3b), either P5-P5 or P7-P7 clusters are removed (depending on linearization). As shown in Figure 3b, if linearization happens via P5, then only the P5-P7 and P7-P7 bridges remain. P7-P7 clusters are not linearized by this reaction and therefore will not sequence. P7-P7 clusters are subsequently removed during Read 2 linearization for sequencing.

[0049] The method presented herein can further comprise an additional step of providing transposome complexes in solution and contacting the solution-phase transposome complexes with the immobilized fragments under conditions whereby the target DNA is fragmented by the transposome complexes solution; thereby obtaining immobilized nucleic acid fragments having one end in solution. In some embodiments, the transposome complexes in solution can comprise a second tag, such that the method generates immobilized nucleic acid fragments having a second tag, the second tag in solution. The first and second tags can be different or the same.

[0050] In some embodiments, one form of surface bound transposome is predominantly

present on the solid support. For example, in some embodiments, at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or at least 99% of the tags present on said solid support comprise the same tag domain. In such embodiments, after an initial tagmentation reaction with surface bound transposomes, at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or at least 99% of the bridge structures comprise the same tag domain at each end of the bridge. A second tagmentation reaction can be performed by adding transposomes from solution that further fragment the bridges. In some embodiments, most or all of the solution phase transposomes comprise a tag domain that differs from the tag domain present on the bridge structures generated in the first tagmentation reaction. For example, in some embodiments, at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or at least 99% of The tags present in the solution phase transposomes comprise the a tag domain that differs from the tag domain present on The bridge structures generated in the first tagmentation reaction.

[0051] Figures 4a and 4b illustrate this embodiment. The solid support shown in Figure 4a comprises only surface bound transposome comprising a single type of tag sequence (e.g. P5 transposome). In this instance, all tagmented bridges are P5-P5 bridges. A P7 transposome is added from solution and tagments into the bridges making all surface bound molecules P5-P7 templates. The surface bound fragments can be subsequently converted into clusters (see, for example, Figure 4b). Furthermore, a pair of clusters will result from each transposome stump which represent the two adjacent fragments in the original intact DNA sample (Figure 4b).

[0052] Also presented herein are solid supports having a library of tagged DNA fragments immobilized thereon prepared according to the above methods.

Physical Maps of Immobilized Polynucleotide Molecules (for reference)

[0053] Also disclosed herein are methods of generating a physical map of immobilized polynucleotides. The methods can advantageously be exploited to identify clusters likely to contain linked sequences (i.e., the first and second portions from the same target polynucleotide molecule). The relative proximity of any two clusters resulting from an immobilized polynucleotide thus provides information useful for alignment of sequence information obtained from the two clusters. Specifically, the distance between any two given clusters on a solid surface is positively correlated with the probability that the two clusters are from the same target polynucleotide molecule, as described in greater detail in WO 2012/025250.

[0054] As an example long dsDNA molecules stretching out over the surface of a flowcell can be tagmented *in situ,* resulting in a line of connected dsDNA bridges across the surface of the flowcell. Further, a physical map of the immobilized DNA can then be generated. The physical map thus correlates the physical relationship of clusters after immobilized DNA is amplified. Specifically, the physical map is used to calculate the probability that sequence data obtained from any two clusters are linked, as described in WO 2012/025250.

[0055] The physical map can be generated by imaging the DNA to establish the location of the immobilized DNA molecules across a solid surface. The immobilized DNA can be imaged by adding an imaging agent to the solid support and detecting a signal from the imaging agent. The imaging agent can be a detectable label. Suitable detectable labels, include, but arc not limited to, protons, haptens, radionuclides, enzymes, fluorescent labels, chemiluminescent labels, and/or chromogenic agents. For example, the imaging agent can be an intercalating dye or non-intercalating DNA binding agent. Any suitable intercalating dye or non-intercalating DNA binding agent as are known in the art can be used, including, but not limited to those set forth in U.S. 2012/0282617.

[0056] In some embodiments, the immobilized double stranded fragments are further fragmented to liberate a free end (see Figure 4a) prior to cluster generation (Figure 4b). Cleaving bridged structures can be performed using any suitable methodology as is known in the art, as exemplified by WO 2012/025250. For example, cleavage can occur by incorporation of a modified nucleotide, such as uracil as described in WO 2012/025250, by incorporation of a restriction endonuclease site, or by applying solution-phase transposome complexes to the bridged DNA structures, as described elsewhere herein.

[0057] A plurality of target DNA molecules can be flowed onto a flowcell comprising a plurality of nano-channels, the nano-channel having a plurality of transposome complexes immobilized thereto. As used herein, the term nano-channel refers to a narrow channel into which a long linear DNA molecule is flown. In some embodiments, no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 or no more than 1000 individual long strands of target DNA are flowed into each nano-channel. The individual nano-channels can be separated by a physical barrier which prevents individual long strands of target DNA from interacting with multiple nano-channels. The solid support can comprise at least 10, 50, 100, 200, 500, 1000, 3000, 5000, 10000, 30000, 50000, 80000 or at least 100000 nano-channels. Transposomes bound to the surface of a nano-channel tagment the DNA. Contiguity mapping can then be performed, for example, by following the clusters down the length of one of these channels. In some embodiments, the long strand of target DNA can be at least 0.1kb, 1kb, 2kb, 3kb, 4kb, 5kb, 6kb, 7kb, 8kb, 9kb, 10kb, 15kb, 20kb, 25kb, 30kb, 35kb, 40kb, 45kb, 50kb, 55kb, 60kb, 65kb, 70kb, 75kb, 80kb, 85kb, 90kb, 95kb, 100kb, 150kb, 200kb, 250kb, 300kb, 350kb, 400kb, 450kb, 500kb, 550kb, 600kb, 650kb, 700kb, 750kb, 800kb, 850kb, 900kb, 950kb, 1000kb, 5000kb, 10000kb, 20000kb, 30000kb, or at least 50000kb in length. The long strand of target DNA can be no more than 0.1kb, 1kb, 2kb, 3kb, 4kb, 5kb, 6kb, 7kb, 8kb, 9kb, 10kb, 15kb, 20kb, 25kb, 30kb, 35kb, 40kb, 45kb, 50kb, 55kb, 60kb, 65kb, 70kb, 75kb, 80kb, 85kb, 90kb, 95kb, 100kb, 150kb, 200kb, 250kb, 300kb, 350kb, 400kb, 450kb, 500kb, 550kb, 600kb, 650kb, 700kb, 750kb, 800kb, 850kb, 900kb, 950kb, or no more than 1000kb in length. As an example, a flowcell having 1000 or more nano-channels with mapped immobilized tagmentation products in the nano-channels can be used to sequence the genome of an organism with short 'positioned' reads. Mapped immobilized tagmentation products in the nano-channels can be used resolve haplotypes. Mapped immobilized tagmentation products in the nano-channels can be used to

resolve phasing issues.

Amplification and Sequencing Immobilized DNA Fragments

[0058] Amplification. The present disclosure further relates to amplification of the immobilized DNA fragments produced according to the methods provided herein. The immobilized DNA fragments produced by surface bound transposome mediated tagmentation can be amplified according to any suitable amplification methodology known in the art. In some embodiments, the immobilized DNA fragments are amplified on a solid support. In some embodiments, the solid support is the same solid support upon which the surface bound tagmentation occurs. In such embodiments, the methods and compositions provided herein allow sample preparation to proceed on the same solid support from the initial sample introduction step through amplification and optionally through a sequencing step.

[0059] For example, in some embodiments, the immobilized DNA fragments are amplified using cluster amplification methodologies as exemplified by the disclosures of US Patent Nos. 7,985,565 and 7,115,400. US Patent Nos. 7,985,565 and 7,115,400 describe methods of solid-phase nucleic acid amplification which allow amplification products to be immobilized on a solid support in order to form arrays comprised of clusters or "colonies" of immobilized nucleic acid molecules. Each cluster or colony on such an array is formed from a plurality of identical immobilized polynucleotide strands and a plurality of identical immobilized complementary polynucleotide strands. The arrays so-formed are generally referred to herein as "clustered arrays". The products of solid-phase amplification reactions such as those described in US Patent Nos. 7,985,565 and 7,115,400 are so-called "bridged" structures formed by annealing of pairs of immobilized polynucleotide strands and immobilized complementary strands, both strands being immobilized on the solid support at the 5' end, preferably via a covalent attachment. Cluster amplification methodologies are examples of methods wherein an immobilized nucleic acid template is used to produce immobilized amplicons. Other suitable methodologies can also be used to produce immobilized amplicons from immobilized DNA fragments produced according to the methods provided herein. For example one or more clusters or colonies can be formed via solid-phase PCR whether one or both primers of each pair of amplification primers are immobilized.

[0060] The immobilized DNA fragments can be amplified in solution. For example, the immobilized DNA fragments can be cleaved or otherwise liberated from the solid support and amplification primers are then hybridized in solution to the liberated molecules. Amplification primers can be hybridized to the immobilized DNA fragments for one or more initial amplification steps, followed by subsequent amplification steps in solution. Thus, an immobilized nucleic acid template can be used to produce solution-phase amplicons.

[0061] It will be appreciated that any of the amplification methodologies described herein or generally known in the art can be utilized with universal or target-specific primers to amplify immobilized DNA fragments. Suitable methods for amplification include, but are not limited to,

the polymerase chain reaction (PCR), strand displacement amplification (SDA), transcription mediated amplification (TMA) and nucleic acid sequence based amplification (NASBA), as described in U.S. Patent No. 8,003,354. The above amplification methods can be employed to amplify one or more nucleic acids of interest. For example, PCR, including multiplex PCR, SDA, TMA, NASBA and the like can be utilized to amplify immobilized DNA fragments. In some embodiments, primers directed specifically to the nucleic acid of interest are included in the amplification reaction.

[0062] Other suitable methods for amplification of nucleic acids can include oligonucleotide extension and ligation, rolling circle amplification (RCA) (Lizardi et al., Nat. Genet. 19:225-232 (1998)) and oligonucleotide ligation assay (OLA) (See generally U.S. Pat. Nos. 7,582,420, 5,185,243, 5,679,524 and 5,573,907; EP 0 320 308 B1; EP 0 336 731 B1; EP 0 439 182 B1; WO 90/01069; WO 89/12696; and WO 89/0983) technologies. It will be appreciated that these amplification methodologies can be designed to amplify immobilized DNA fragments. For example, the amplification method can include ligation probe amplification or oligonucleotide ligation assay (OLA) reactions that contain primers directed specifically to the nucleic acid of interest. The amplification method can include a primer extension-ligation reaction that contains primers directed specifically to the nucleic acid of interest. As a non-limiting example of primer extension and ligation primers that can be specifically designed to amplify a nucleic acid of interest, the amplification can include primers used for the GoldenGate assay (Illumina, Inc., San Diego, CA) as exemplified by U.S. Pat. No. 7,582,420 and 7,611,869.

[0063] Exemplary isothermal amplification methods that can be used in a method of the present disclosure include, but are not limited to, Multiple Displacement Amplification (MDA) as exemplified by, for example Dean et al., Proc. Natl. Acad. Sci. USA 99:5261-66 (2002) or isothermal strand displacement nucleic acid amplification exemplified by, for example U.S. Pat. No. 6,214,587. Other non-PCR-based methods that can be used in the present disclosure include, for example, strand displacement amplification (SDA) which is described in, for example Walker et al., Molecular Methods for Virus Detection, Academic Press, Inc., 1995; U.S. Pat. Nos. 5,455,166, and 5,130,238, and Walker et al., Nucl. Acids Res. 20:1691-96 (1992) or hyperbranched strand displacement amplification which is described in, for example Lage et al., Genome Research 13:294-307 (2003). Isothermal amplification methods can be used with the strand-displacing Phi 29 polymerase or Bst DNA polymerase large fragment, 5'->3' exo$^-$ for random primer amplification of genomic DNA. The use of these polymerases takes advantage of their high processivity and strand displacing activity. High processivity allows the polymerases to produce fragments that are 10-20 kb in length. As set forth above, smaller fragments can be produced under isothermal conditions using polymerases having low processivity and strand-displacing activity such as Klenow polymerase. Additional description of amplification reactions, conditions and components are set forth in detail in the disclosure of U.S. Patent No. 7,670,810.

[0064] Another nucleic acid amplification method that is useful in the present disclosure is Tagged PCR which uses a population of two-domain primers having a constant 5' region followed by a random 3' region as described, for example, in Grothues ct al. Nucleic Acids Res.

21(5):1321-2 (1993). The first rounds of amplification arc carried out to allow a multitude of initiations on heat denatured DNA based on individual hybridization from the randomly-synthesized 3' region. Due to the nature of the 3' region, the sites of initiation are contemplated to be random throughout The genome. Thereafter, the unbound primers can be removed and further replication can take place using primers complementary to the constant 5' region.

[0065] Sequencing. (for reference) The present disclosure further relates to sequencing of the immobilized DNA fragments produced according to the methods provided herein. The immobilized DNA fragments produced by surface bound transposome mediated tagmentation 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. The immobilized DNA fragments can be sequenced on a solid support. The solid support for sequencing can be the same solid support upon which the surface bound tagmentation occurs. The solid support for sequencing can be the same solid support upon which the amplification occurs.

[0066] One preferred sequencing methodology is sequencing-by-synthesis (SBS). 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.

[0067] Flow cells provide a convenient solid support for housing amplified DNA fragments produced by the methods of the present disclosure. One or more amplified DNA fragments in such a format 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 flow cell 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; US 7,057,026; WO 91/06678; WO 07/123744; US 7,329,492; US 7,211,414; US 7,315,019; US 7,405,281, and US 2008/0108082.

[0068] Other sequencing procedures that use cyclic reactions can be used, such as pyroscquencing. 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); US 6,210,891; US 6,258,568 and US. 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, US 7,595,883, and US 7,244,559.

[0069] Methods involving the real-time monitoring of DNA polymerase activity can be used. 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 zeromode 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).

[0070] Some SBS methods 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 from Ion Torrent (Guilford, CT, a Life Technologies subsidiary) 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.

[0071] Another useful 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 methods, 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. Patent 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)).

[0072] Exemplary methods for array-based expression and genotyping analysis that can be applied to detection according to the present disclosure are described in US Pat.

Nos.7,582,420; 6,890,741; 6,913,884 or 6,355,431 or US Pat. Pub. Nos. 2005/0053980 A1; 2009/0186349 A1 or US 2005/0181440 A1.

[0073] An advantage of the methods set forth herein is that they provide for rapid and efficient detection of a plurality of target nucleic acid in parallel. Accordingly the present disclosure provides integrated systems capable of preparing and detecting nucleic acids using techniques known in the art such as those exemplified above. Thus, an integrated system of the present disclosure can include fluidic components capable of delivering amplification reagents and/or sequencing reagents to one or more immobilized DNA fragments, the system comprising components such as pumps, valves, reservoirs, fluidic lines and the like. A flow cell can be configured and/or used in an integrated system for detection of target nucleic acids. Exemplary flow cells are described, for example, in US 2010/0111768 A1 and US Ser. No. 13/273,666. As exemplified for flow cells, one or more of the fluidic components of an integrated system can be used for an amplification method and for a detection method. Taking a nucleic acid sequencing method as an example, one or more of the fluidic components of an integrated system can be used for an amplification method set forth herein and for the delivery of sequencing reagents in a sequencing method such as those exemplified above. Alternatively, an integrated system can include separate fluidic systems to carry out amplification methods and to carry out detection methods. Examples of integrated sequencing systems that are capable of creating amplified nucleic acids and also determining the sequence of the nucleic acids include, without limitation, the MiSeq™ platform (Illumina, Inc., San Diego, CA) and devices described in US Ser. No. 13/273,666.

Solid supports with immobilized transposomes and methods of preparation

[0074] Other embodiments presented herein include solid supports, such as flowcells, having transposome complexes immobilized thereon. In certain embodiments, the transposome complexes comprise a transposase bound to a first polynucleotide, the polynucleotide comprising (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain. The density of these surface bound transposomes can vary. For example, in some embodiments, the transposome complexes are present on the solid support at a density of at least 10^3, 10^4, 10^5, or at least 10^6 complexes per mm^2.

[0075] Also presented herein are methods of generating a flowcell for tagmentation. The methods can comprise, for example, immobilizing a plurality of transposome complexes to a solid support, the transposome complexes comprising a transposase bound to a first polynucleotide, the first polynucleotide comprising (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain.

[0076] Transposome complexes can be immobilized to a solid support in a variety of methods, which will be appreciated by one of skill in the art. In one embodiment, the method comprises providing a solid support having a plurality of first polynucleotides immobilized thereon, and

contacting the solid support with transposase holoenzyme and a second polynucleotide, the second polynucleotide comprising a region complementary to the transposon end sequence. In some embodiments, The second polynucleotide is hybridized to the immobilized first polynucleotide before the transposase holoenzyme is added. In some embodiments, the second polynucleotide and the transposase holoenzyme are provided together. Figure 5a shows one embodiment of this method for assembling the surface bound transposome complexes. In one method, transposase holoenzyme is added along with the non-transferred strand of the ME adaptor to a flowcell containing oligonucleotides comprising from the 5' surface grafted end: a cluster amplification primer (A.pr) and sequencing primer (S.pr) and the ME sequence (Figure 5a). Alternatively, the non-transferred ME' strand can hybridized to the surface grafted ME strand first and then the transposase added to the flowcell.

[0077] The transposome complexes can be assembled in solution and immobilizing can comprise a further step of ligating the first polynucleotide to a splint oligonucleotide coupled to the solid support. This is illustrated in Figure 5b. The splint oligonucleotide can be extended using a polymerase before a ligating step occurs.

[0078] Transposome dimer can be assembled by hybridizing a looped oligonucleotide to an immobilized first polynucleotide. For example, a looped oligonucleotide can comprise a first end and a second end, with the first end being complementary to the transposon end sequence of the first polynucleotide. The second end of the looped oligonucleotide can be complementary to a second transposon end sequence. The second transposon end sequence can be, for example, part of a solution-phase first polynucleotide. The immobilized first polynucleotide and the solution-phase first polynucleotide can comprise dissimilar transposon end sequences or complements thereof. The looped oligonucleotide can comprise sequences complementary to each of the dissimilar transposon end sequences at the first and second ends. An illustration of this is shown in Figure 5c. As shown in Figure 5c, a contiguous adaptor pair is generated by hybridizing two oligonucleotides to the surface bound oligonucleotide. This can be achieved by employing two dissimilar transposon end sequences (ME1 and ME2). Addition of transposase holoenzyme reconstitutes an active 'looped' transposome complex, where only one of the two adaptors of each transposome is coupled to the surface Figure 5c.

[0079] Transposome complexes can be assembled on a standard paired end flow cell with amplification primers immobilized thereto (e.g. a HiSeq flow cell or MiSeq flow cell sold by Illumina Inc, San Diego, CA). This can be accomplished by, for example, hybridization of a 'splint' oligonucleotide that anneals to one or both species of surface grafted amplification primer. The splint acts as a template to then extend the grafted surface primer with a polymerase and dNTPs to form an oligonucleotide duplex that contains a surface amplification primer, a sequencing primer and the transposon end sequences of the transposase. Addition of transposase assembles a transposome of the surface. This embodiment is illustrated in Figure 5d.

[0080] In any of the embodiments provided herein, transposome complexes may be homodimers, or heterodimers. For example, as illustrated in Figure 11, a homodimeric

transposome would comprise two P5-ME adaptors at both sites or alternatively, two P7-ME adaptors. Similarly, a heterodimeric transposome complex would comprise both P5-ME and P7-ME adaptors, as shown in Figure 11.

Tagmentation using transposome beads

[0081] One embodiment presented herein is a population of microparticles having transposome complexes immobilized thereto. The use of a solid support such as beads can provide several advantages over solution-based tagmentation. For example, in standard solution-based tagmentation, it is difficult to control the final fragment size of the tagmentation reaction. Fragment size is a function of the ratio of transposomes to the amount and size of DNA and to the duration of the reaction. Even if these parameters are controlled, size selection fractionation is commonly required as an additional step to remove excess small fragments shorter than the combined paired-read lengths. The methods provided herein avoid those disadvantages. Specifically, bead-immobilized transposomes allow for selection of final fragment size as a function of the spatial separation of the bound transposomes, independent of the quantity of transposome beads added to the tagmentation reaction. An additional limitation of solution-based tagmentation is that it is typically necessary to do some form of purification of the products of the tagmentation reaction both before and after PCR amplification. This typically necessitates some transfer of reactions from tube to tube. In contrast, tagmentation products on the bead based transposomes can be washed and later released for amplification or other downstream processing, thus avoiding the need for sample transfer. For example, where transposomes are assembled on paramagnetic beads, purification of the tagmentation reaction products can easily be achieved by immobilizing the beads with a magnets and washing. Thus, tagmentation and other downstream processing such as PCR amplification can all be performed in a single tube, vessel, droplet or other container. Tagmentation and downstream processing of samples can take place on a microfluidic droplet based device, as exemplified in the disclosure of U.S. Application Ser. No. 13/670,318, filed on November 6, 2012 entitled "INTEGRATED SEQUENCING APPARATUSES AND METHODS OF USE". For example, in a microfluidic droplet based device, a droplet containing target nucleic acid, wash buffer or other reagents may be passed over a surface comprising immobilized transposome complexes. Likewise, a droplet comprising beads having transposomes immobilized thereon may be contacted with target nucleic acid, wash buffer or other reagents in a microfluidic droplet based device.

[0082] In some embodiments, the immobilized transposome complexes comprise a transposase bound to a first polynucleotide and a second polynucleotide; wherein the first polynucleotide is immobilized at its 5' end to the surface of the microparticle and the second polynucleotide is hybridized to the 3' end of the first polynucleotide; and wherein the first polynucleotide comprises: (i) a 3' portion comprising a transposon end sequence, and (ii) a first tag comprising a first tag domain.

[0083] Figures 9-12 provide an illustration of transposomes assembled on the surface of

paramagnetic beads. Figure 9 shows a bead surface with two different first polynucleotides immobilized thereto. The first polynucleotides that are shown comprise a transposon end sequence (ME). One of the first polynucleotides shown in comprises a tag domain comprising an amplification primer sequence (P5) and a sequencing primer sequence (Read 1). The other polynucleotide shown in Figure 9 comprises a different amplification primer sequence (P7) and a sequencing primer sequence (Read 2). The first polynucleotides may also comprise an index tag. The index tag may be unique to the bead, or it may be shared with one or more other beads in the population. A single bead may have only a single index tag immobilized thereto, or it may have a plurality of index tags immobilized thereto.

[0084] Figure 10 shows a second polynucleotide hybridized to each first polynucleotide. The second polynucleotide comprises a region complementary to the transposon end sequence of the first polynucleotide. The second polynucleotide can be 15 bp in length. The second polynucleotide can be approximately 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 bp or more than 20 bp in length. The second polynucleotide can be phosphorylated at its 5' end.

[0085] Figure 11 illustrates assembly of transposomes on the surface of the bead, where transposase enzyme is contacted with the immobilized oligonucleotides. As shown in Figure 11, when the transposase is added to the beads, three types of transposome complexes are formed: a P5:P5, a P7:P7, and a P5:P7 complex (Figure 11).

[0086] When the transposome beads are added to a solution of target DNA in a tagmentation buffer, tagmentation takes place, linking the DNA to the surface of the beads. An immobilized library of tagged DNA fragments is generated.

[0087] Figure 12 illustrates an embodiment where successive tagmentation into the DNA results in bridged molecules between transposomes. Where three types of transposomes are present (as in Figure 11), three types of bridge complexes result: a P5:P5, P7:P7 and P5:P7 complex in a ratio of 25:25:50, respectively.

[0088] The length of the bridged fragments can be dictated by the density of the transposomes on the surface of the bead. This density is tunable via the amount of oligonucleotide on the surface, the amount of duplex transposon end complexes on the surface and the amount of transposase enzyme added during the transposome assembly. Once tagmentation is complete, the P5:P7 tagmentation products can be liberated from the surface of the bead using any suitable method. The tagmentation products can be liberated from the beads using an amplification method such as suppression PCR, step-out PCR and the like. The tagmentation products can be liberated from the beads by cleavage. The cleavage can be, for example, chemical, enzymatic, photochemical or a combination thereof. It will be appreciated that any suitable method for releasing one or more tagmentation products from a solid support can be utilized in the methods provided herein.

[0089] DNA can be efficiently contacted with surface bound transposomes using any suitable method for increasing the probability of contact. For example, precipitation of DNA onto the

solid surface can be utilized to increase contact between the target DNA and the transposome complexes on the solid surface. Any one of a number of methods that are known in the art for contacting DNA with a solid support can be utilized, as exemplified by the disclosure of WO 2010/115122. As will be appreciated by one of skill in the art, DNA can be precipitated onto a solid support by the addition of PEG, ethanol or any one of a variety of other agents known to precipitate DNA onto surfaces, including, for example, any one of a number of buffers used in solid phase reversible immobilization (SPRI) technology.

[0090] A population of beads bearing immobilized transposome complexes can be mixed with an excess of beads that bear no transposomes or oligonucleotides, thereby reducing the likelihood of tagmentation across two or more different beads. Another method to reduce the likelihood of tagmentation across two or more different beads includes immobilizing beads so contact between beads is minimized. Immobilization of beads can be accomplished by any of a number of techniques known in the art, including, for example, immobilizing the beads via magnetism to the sides of a solid surface such as a microcentrifuge tube, or any other immobilization technique as exemplified by WO 2010/115122.

[0091] Transposome beads can be used to isolate and identify nucleic acids from a single cell, such as a prokaryotic or eukaryotic cell. For example, particles such as beads can be coated with indexed transposomes which share the same index (all of the transposomes present on a particular bead carry the same index, which is different from the index present on another bead). The beads can then be placed inside cells through any one of variety of methodologies known in the art. For example, methods for delivering beads inside cells include, but are not limited to gene guns, photothermal nanoblades (Wu et al. Anal Chem. (2011) 4:1321-7), and peptides used in conjunction with cell permeabilizing substances (Nitin et al Ann Biomed Eng. (2009) 37:2018-2027) and the like. It will be appreciated that any suitable method for associating DNA from a single cell with a particle bearing indexed transposomes can be used in the methods presented herein.

[0092] Transposomes can be covalently attached to the beads as described in detail hereinabove. Additionally or alternatively, transposomes can be released from the beads upon the application of a chemical or physical stimulus. Some examples of stimuli which can trigger release of transposome from a solid support include light and/or temperature changes. The transposomes can be released from the solid support using the activity of an enzyme such as a restriction endonuclease. The transposomes can be detached from the beads and move freely inside the cell. Once the beads (or alternatively, the released transposomes) come into contact with chromatin or DNA, tagmentation can take place. It will be understood that in eukaryotic and prokaryotic systems, not all genomic DNA will always be accessible and/or available for tagmentation. Up to 0.001%, 0.01%, 0.1%, 1%, 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% or 99% or more that 99% of the total DNA in a cell can be tagmented by the transposomes. The use of uniquely tagged beads makes it possible to identify reads from the same cell by grouping together reads that share the same indices. These reads can be considered as derived from the same bead (and therefore from the same cell). Additionally or

alternatively, a cross-linking step can be performed to cross-link cellular DNA while it is still inside the cell's nucleus. Such a step ensures that the DNA from a single cell is held together. After extracting the DNA from cells, it can then mixed with the indexed beads and subjected to tagmentation as described in detail hereinabove. Additionally or alternatively, variations of the method include the tagmentation of flow sorted chromosomes. Flow sorted chromosomes can be mixed with the indexed beads described above. A particular chromosome would adhere to one bead and the fragments generated through tagmentation would contain the same index. This could enable phasing of SNPs from whole chromosomes (haplotyping).

[0093] The total number of indices can be greater than the total number of cells that are successfully tagmented. A dual index approach can be implemented, for example, as a way to increase the number of possible combinations of indices. As one example, the use of two 8 base indices would give a theoretical number of combinations of $4^8 \times 4^8 = 4.3 \times 10^9$.

[0094] An approach can be used to ensure that an individual cell is not tagmented by multiple beads. For example, one approach is to use beads of a size which is similar to that of the cell. This would ensure that a cell would not be able to accommodate multiple beads. Additionally or alternatively, another approach is to make use of a cell to bead ratio which favors single cell targeting. For example, if there are far more cells than beads, then the poisson distribution of beads inside the cells means that the cells that have taken up a single bead far outnumber the cells that have taken up two or more beads.

[0095] The single cell approach described above can be used to determine whether two SNPs (or structural rearrangements) arc present in the same cell. For example, in the case of heterogeneous populations of cancer cells, knowing whether two SNPs are present in the same cell or in different cells could aid in implementing the right cancer therapy.

[0096] The single cell approach described above can be used to study RNA. For example, by coating the beads with suitable enzymes (i.e. reverse transcriptase) and oligonucleotides for a reverse transcription step, gene expression at the single cell level can be analyzed. After introducing the beads coated with reverse transcriptase, oligonucleotides and transposomes, the cytoplasmic RNA can be converted into cDNA, tagged and prepared inside the cells.

Methods of assembling long reads using barcodes on immobilized transposomes (for reference)

[0097] Barcode-assisted assembly of DNA fragments enables isolation of individual long DNA molecules within a population of DNA molecules and conversion of each molecule into a uniquely barcoded sub-fragment library. When the entire population of sub-fragmented DNA molecules is sequenced, the subfragments can be assembled back into their original long molecule by reference to the barcodes they contain.

[0098] Various methods of barcoding individual DNA molecules are known. For example, the 'dilution method,' isolates individual long molecules by extreme dilution and aliquoting into separate compartments (e.g., wells of a plate), such that each well contains only one or just a few molecules of DNA. Because each well is physically separate, a library preparation can be done in each well with a unique barcode. Thereafter the contents of the wells are pooled and sequenced. Another method employs an emulsion wherein each droplet contains long DNA molecules, library preparation reagents and a barcode unique to each droplet. Another approach uses a large library of indexed looped transposome complexes to insert multiple twin barcodes along the length of the DNA while preserving the intactness of the molecule. Subsequent cleavage between the barcode 'twins' yields fragments that can be sequenced and reassembled by matching up the twin barcodes. Each of the above-mentioned barcoding methods carries with it disadvantages that are overcome by the barcoding methods presented herein.

[0099] Presented herein are alternative methods to the above described ways of isolating and barcoding individual long molecules. The methods presented herein achieve advantages of physical isolation similar to the emulsion method without using emulsions, and at the same time provide a complexity that is much greater than that provided by the large number of 'wells' used in the dilution method. The unique barcodes of the present methods are in some ways analogous to the 'wells' of the dilution method except that the number of beads in the bead-based method can often be much higher than the number of wells in the dilution method. An additional advantage over the emulsion methods is that in the bead-based method, the barcodes are deterministically distributed (i.e., one barcode per bead) and not random (i.e. Poisson distributed). The methods presented herein also achieve the same initial preservation of molecule intactness and contiguity but without the need for looped transposome complexes as used in some other methods. Additionally, the methods presented herein do not require as large a code space as used in some other methods.

[0100] Accordingly, barcoding method can comprise providing a population of microparticles having transposome complexes immobilized thereto, the transposome complexes comprising a transposase bound to a first polynucleotide and second polynucleotide. The first polynucleotide can comprise an index domain associated with the microparticle. The index domain can be unique to the microparticle. The population of microparticles can comprise at least a plurality of index domains. The index domain can be present on more than one microparticle in the population of microparticles. A microparticle in the population of microparticles can comprise more than one index domain.

[0101] The barcoding methods presented herein further comprise applying a target DNA to the population of microparticles, thereby generating immobilized DNA fragments that are tagged with the index domain. DNA can be efficiently contacted with surface bound transposomes using any suitable method for increasing the probability of contact as discussed hereinabove, as exemplified by WO 2010/115122.

[0102] The methods can be performed using any one of a variety of known formats, for

example, with a combination of tagmentation reagents and a bead array for the library preparation, followed by an indexed sequencing run and bespoke data analysis. Any other suitable method that maintains beads in static separation from one another can be used for surface tagmentation and indexing of samples. For example, physical configurations such as wells or small depressions in the substrate that can retain the beads, such that a microsphere can rest in the well, or the use of other forces (magnetic or compressive), or chemically altered or active sites, such as chemically functionalized sites, electrostatically altered sites, hydrophobically and/or hydrophilically functionalized sites, or spots of adhesive.

[0103] The microspheres can be non-covalently associated in the wells, although the wells may additionally be chemically functionalized as is generally described below, cross-linking agents may be used, or a physical barrier may be used, e.g., a film or membrane over the beads.

[0104] The surface of the substrate can be modified to contain chemically modified sites that can be used to attach, either covalently or non-covalently, the microspheres of the invention to the discrete sites or locations on the substrate. "Chemically modified sites" in this context includes, but is not limited to, the addition of a pattern of chemical functional groups including amino groups, carboxy groups, oxo groups and thiol groups, that can be used to covalently attach microspheres, which generally also contain corresponding reactive functional groups; the addition of a pattern of adhesive that can be used to bind the microspheres (either by prior chemical functionalization for the addition of the adhesive or direct addition of the adhesive); the addition of a pattern of charged groups (similar to the chemical functionalities) for the electrostatic attachment of the microspheres, e.g., when the microspheres comprise charged groups opposite to the sites; the addition of a pattern of chemical functional groups that renders the sites differentially hydrophobic or hydrophilic, such that the addition of similarly hydrophobic or hydrophilic microspheres under suitable experimental conditions will result in association of the microspheres to the sites on the basis of hydroaffinity. For example, the use of hydrophobic sites with hydrophobic beads, in an aqueous system, drives the association of the beads preferentially onto the sites. As outlined above, "pattern" in this sense includes the use of a uniform treatment of the surface to allow attachment of the beads at discrete sites, as well as treatment of the surface resulting in discrete sites. As will be appreciated by those in the art, this may be accomplished in a variety of ways.

[0105] A multitude of beads comprising surface bound transposomes can be generated, wherein each bead contains many transposomes but all transposomes on any given bead all contain the same barcode. The generation of a population of monoclonal barcoded transposome oligonucleotides on beads can be performed according to any one of a number of techniques as is known in the art, as exemplified by the disclosure of U.S. 5,604,097.

[0106] Figure 13 illustrates one embodiment, where a surface bound transposome contains long oligonucleotides attached to the surface via their 5' end. As set forth in Figure 13, the 3' end of the oligonucleotide comprises the Mosaic End (METS) sequences of the Tn5 transposase enzyme. Upstream of this sequence (closer to the 5' end) is a barcode sequence

of typically 6-8 bases in length. Further upstream is a primer sequence. Additionally, the oligonucleotide can further comprise a cleavage site to enable the oligonucleotide to be cleaved off the surface: its presence in the oligonucleotide is optional to the overall method. A second short oligonucleotide (the non-transferred strand, MENTS) is hybridized to the METS sequence at the 3' end of the long surface grafted oligonucleotide. Finally, a transposase dimer is assembled at the 3' end of the oligonucleotides to form a functional transposome.

[0107] Figure 14 illustrates barcoding on beads in an array of wells. As shown in Figure 14, when long molecules of dsDNA are added to an array of bead-immobilized transposomes, a given molecule encounters a bead and gets tagmented many times by the transposomes on the bead. Each fragment becomes immobilized to the bead and tagged with the barcode associated with that bead. In the particular embodiment shown in Figure 14, the physical separation of beads in the array chip prevents a DNA molecule from reaching between two beads. In other embodiments, the beads are in close contact and one or more DNA molecules may stretch between two or more beads. In some embodiments, more than one DNA molecule can be tagmented per bead. The probability of two alleles being tagmented onto the same bead is low and is a function of the concentration of the DNA added, the number of beads and the number of barcodes. For example, to avoid two alleles occurring in the same well, 0.1x haplome equivalents (50,000 genomes equivalents) would need to be loaded to 1 million beads each with a unique barcode.

[0108] Figure 15 illustrates transfer of barcoded tagged molecules to a sequencing reaction. As shown in Figure 15, once the tagmentation is complete the DNA is transferred from the bead surface to solution so that individual fragments can be pooled and prepared for sequencing. Sequencing and assembly by reference to the barcodes enables the sequence of the original long tagmented DNA molecules to be re-created, thus enabling long or pseudo-long reads and phasing of SNPs.

[0109] Release of the barcoded surface tagmented fragments to the solutions can be achieved using any suitable methodology as is known in the art. For example, in some embodiments, the tagmented molecules can be cleaved off the surface of the beads via a cleavage moiety present at the 5' end of the surface bound oligonucleotides (see Figure 13). The cleavage moiety can be any moiety suitable for cleavage of a nucleic acid strand from a solid support. Examples of methods that utilize a cleavage moiety include, but are not limited to, restriction endonuclease cleavage, chemical cleavage, RNase cleavage, photochemical cleavage, and the like, including those cleavage methods and cleavage moieties set forth in WO 2006/064199.

[0110] Cleavage using a cleavage moiety yields a molecule having the following format: 5' - Primer - Barcode - ME - Insert - ME - Barcode - Primer - 3'

[0111] The "Primer" regions can be used as hybridization points to hybridize PCR step-out primers that enable additional sequences to be added such as amplification and sequencing primers. For example, amplification primers P5 and P7 can be added. Once added,

suppression PCR can be used, for example, to enrich for molecules that have P5 adaptors on one end and P7 on the other.

[0112] Amplification can be performed directly off the beads with step-out primers that add P5 and P7 adaptor sequences by suppression PCR. In another embodiment, each bead can have two types of surface grafted oligonucleotides where the primer sequence (as in Figure 13) is either P5-Read1 sequencing primer or P7-Read 2 sequencing primer. This will result in mixed P5-P7 transposomes. These can either be cleaved off the beads and followed by suppression PCR to enrich the P5/P7 molecules or amplified directly off the beads, as described above.

[0113] A single transposome type (e.g. P5-Read 1-barcode) may be present on the surface of the bead. Once surface tagmentation is complete, a second transposome bearing a different amplification and/or sequencing primer can be added in solution to cleave the bridged molecules. This yields all molecules with the same adaptor format that can either be cleaved or amplified off the bead surface. An additional sample-specific barcode could be added to the solution transposome such that multiple samples can be pooled by the method. Figure 16 is an illustration of this embodiment. As shown in Figure 16, solution-phase transposomes bearing a P7 amplification primer sequence are added to a bead with immobilized bridged tagmentation products. The immobilized fragments arc tagged with a polynucleotide sequence comprising a P5 amplification primer sequence and an barcode index tag (i9) that is unique to the bead. As shown in Figure 16, after the second tagmentation reaction, each fragment has a free end bearing a P7 primer sequence and an immobilized end bearing a P5 primer sequence.

EXAMPLE 1

Surface tagmentation on a flowcell

[0114] This example describes an experiment confirming the embodiment illustrated in Figures 4a and 4b (surface tagmentation followed by solution tagmentation).

[0115] An experiment on an 8 lane flowcell was carried out using a list of conditions and controls as shown in Figures 6a and 6b. Lane 6 indicates the complete end to end exposition of the method: unfragmented *E. coli* genomic DNA was added to a flowcell where upon it was tagmented by the surface bound transposomes. Next heat was applied to the flowcell to selectively dehybridize the naked ME sequences thus abolishing them as a target for a second tagmentation reaction when transposome was next added from solution. Following the second tagmentation reaction, clusters were generated and 'paired-end' sequenced (2x36 base reads) was performed. Sequencing metrics are given in slide 7 which show that in lane 6, 73.14% of clusters passed filters and 74.69 of these aligned. This provides enough data to yield a gap size plot of the inserts and an estimation of library diversity (see Figure 8).

[0116] Other lanes included controls which are listed below:

Lane 1 comprises PhiX DNA as a positive control to ensure everything has been pumping correctly and cluster generation and sequencing works as expected.

Lane 2 is another control lane and illustrates tagmentation into surface bound transposons in the absence of target DNA. The flowcell (FC) comprises a standard paired end FC onto which an oligonucleotide containing the tagmentation primer sequence and ME' sequence (non-transferred ME' strand) was hybridized to the P7 surface oligonucleotide. This oligonucleotide was added at a saturating concentration. First extension resulted in a double stranded transposon with ME ends. A P5 transposome was assembled in solution and flowed onto the FC (6.25 nM/lane). The P5 transposome tagments the double stranded surface bound transposons. The tagmentation products are subsequently converted into clusters.

Lane 3 illustrates tagmentation into surface bound transposons in the absence of target DNA, only in this instance the FC has been heated to 75°C in order to convert the ds surface bound transposons into single stranded oligonucleotides prior to the addition of the P5 Transposome from solution in order to prevent tagmentation into these constructs, as this was interfering with the tagmentation of the target DNA.

Lane 4 illustrates the addition of a surface bound transposome to the lane 3 conditions. In this lane an oligonucleotide containing the tagmentation primer sequence and ME' sequence was hybridized to the P7 surface oligonucleotide. First extension results in a double stranded transposon with ME ends. Following this a P7 transposome was assembled on the surface of the FC by adding Tn5 enzyme to the lane at 50X concentration and incubating at 30°C for 30 minutes. The FC was then heated to 75°C prior to adding P5 Transposome from solution.

Lane 5 comprises the same conditions as lane 4, only in this case instead of adding P5 Transposome from solution, an *E. coli* 900 bp library (with P5/P7 ends) was added to determine whether the P7 surface bound transposome remains active after the heating step.

Lane 6 illustrates an example of the invention put into practice. In this instance the FC comprises of a standard paired end FC onto which an oligonucleotide containing the tagmentation primer sequence and ME' sequence was hybridized to the P7 surface oligonucleotide. First extension results in double stranded transposons with ME ends. P7 transposomes are assembled on the surface of the FC by adding Tn5 enzyme to the lane at 50X concentration and incubating at 30°C for 30 minutes. Target DNA (300ng of unfragmented *E. coli* genomic DNA) was added onto the FC lane and an incubation step of 15 min at 55°C was carried out in order for tagmentation to take place. The P7 surface bound transposome was washed off using PBI (Qiagen) and the FC was then heated to 75°C prior to adding P5 transposome from solution. Following addition of P5 transposome from solution and an incubation step of 15 min at 55°C, a stand displacement extension reaction was carried out in order to fill in the 9-bp gaps generated in the DNA backbone by the transposition reaction. The stand displacement extension reaction comprises of the addition of a Bst and dNTP mix and incubation at 65°C for 5 min. The P5 transposome was washed off in the final step. At this point all surface bound molecules should be P5-P7 templates and can therefore be converted

into clusters.

Lane 7 comprises of the same conditions as lane 6, only in this instance the heat step has been left out in order to highlight the effect of heat on the cluster number and % align.

Lane 8 comprises a negative control where the P7 transposome has been assembled on the surface at 50X concentration in the presence of a saturating concentration of ds surface bound transposons, however, no target DNA has been added. This allows an assessment of whether the surface bound P7 transposome tagments into its neighboring ds surface bound transposons.

EXAMPLE 2

Surface bound sample preparation from an _E. coli_ scrape

[0117] An _E. coli_ sample (5mm by 2mm scrape from a lawn on an agar plate) was scraped and resuspended in a tube containing water and glass beads. The suspended cells and beads were mixed using a vortex mixer to break open the cells and then centrifuged to pellet cellular debris. The supernatant (containing the cell lysate, including proteins and nucleic acids) was removed and added to a Genome Analyzer flowcell (Illumina, Inc., San Diego, CA) having immobilized transposomes according to the protocol described in Example 1.

[0118] Cluster generation was performed on the flowcell using a Cluster Station sample preparation device (Illumina, Inc., San Diego, CA). After cluster generation, a paired-end sequencing run was performed with reads of 36 bases in each direction.

[0119] For Read 1, 58.99% of clusters passed filters and 92.16 of these aligned. For Read 2, 58.99% of clusters passed filters and 55.08 of these aligned. These data confirm that unpurified cell lysates can be added directly to a flowcell having immobilized transposomes with surprisingly robust sequencing results.

EXAMPLE 3

[0120] This example describes methods to avoid tagmentation of surface-bound oligonucleotide duplexes when solution phase transposomes are added to a flowcell.

[0121] One method for assembling transposomes on the surface of a flow cell is to take a standard paired end flow cell, hybridize a 'splint' oligonucleotide against the P5 and/or P7 surface grafted oligonucleotides forming an extendable overhang that can be extended with a

polymerase to make a duplex containing a double stranded ME sequence. At this stage transposase enzyme can be added to form a functional surface bound transposome (figure 5d). If only a portion of the duplexes form a transposome, remaining 'naked' ME duplexes may subsequently become targets for tagmentation by either nearby surface-bound transposomes or transposomes added from solution. Furthermore, fully assembled surface transposomes contain dsDNA portions upstream of the ME sequences that can also become targets for tagmentation by nearby surface-bound transposomes or transposomes added from solution. This undesired tagmentation manifests itself in a reduction of the percentage of clusters passing purity filters that align to the target genome.

[0122] An example of this effect can be seen by comparing lanes 4 and 5 versus lanes 6 and 7 in figure 18b. In lanes 6 and 7 transposomes were assembled as described above with the long splint oligonucleotides that, after extension and transposome assembly, produced surface bound duplexes that have at least 50 double stranded bases. Upon use in a surface tagmentation reaction (Figure 18a) and subsequent sequencing, only 22.59 and 15.52% of clusters aligned to the target *E.coli* respectively (Figure 18b).

[0123] To avoid producing a population of clusters that contain sequences that do not align to the target genomic DNA, transposome assembly was done as indicated in Figure 17. A 'splint' oligonucleotide was first hybridized against the P5 and/or P7 surface grafted oligonucleotides forming an extendable overhang that could be extended with a polymerase to make a duplex containing a double stranded ME sequence. Next, the splint oligonucleotide was removed by de-annealing and washing away, and then was replaced by hybridizing a shorter oligonucleotide as little as 12 bases long but preferably 16 bases long to the 3' end of the surface attached ME oligonucleotide. Transposase was added to assemble a transposome that did not contain any exposed dsDNA. Any 'naked' duplexes that did not contain a bound transposase were removed by incubating the flow cell at 60°C and washing under flow conditions to selectively de-anneal and remove the short duplex DNA.

[0124] The beneficial effect of this approach to assembling transposomes can be seen in Lanes 4 and 5 where the surface bound transposomes were assembled by this method and used in a surface tagmentation reaction (Figure 18a). The percentage of clusters that aligned to *E.coli* increases from 22.59 and 15.52% for lane 6 and 7 respectively, to 96.47 and 96.41 for lanes 4 and 5 respectively (Figure 18b).

[0125] The term comprising is intended herein to be open-ended, including not only the recited elements, but further encompassing any additional elements.

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in

compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO20121106546A2 [0003]
- WO2010048605A1 [0004]
- WO2012103545A1 [0005]
- US20120530630A1 [0007]
- US20100120098A [0028] [0029] [0031]
- WO2005065814A [0034]
- US20080280773A [0034]
- US13661524B [0036]
- US20120316086A1 [0036]
- WO04018497A [0038] [0067]
- US7057026B [0038] [0067]
- WO9106678A [0038] [0067]
- WO07123744A [0038] [0067]
- US7329492B [0038] [0067]
- US7211414B [0038] [0067]
- US7315019B [0038] [0067]
- US7405281B [0038] [0067]
- US20080108082A [0038] [0067]
- US7985565B [0042] [0059] [0059] [0059]
- US7115400B [0042] [0059] [0059] [0059]
- WO2012025250A [0053] [0054] [0056] [0056]
- US20120282617A [0055]
- US8003354B [0061]
- US7582420B [0062] [0062] [0072]
- US5185243A [0062]
- US5679524A [0062]
- US5573907A [0062]
- EP0320308B1 [0062]
- EP0336731B1 [0062]
- EP0439182B1 [0062]
- WO9001069A [0062]
- WO8912696A [0062]
- WO890983A [0062]
- US7611869B [0062]
- US6214587B [0063]
- US5455166A [0063]
- US5130238A [0063]

- US7670810B [0063]
- US6210891B [0068]
- US6258568B [0068]
- US6274320B [0068]
- US1157111W [0068]
- US20050191698A1 [0068]
- US7595883B [0068]
- US7244559B [0068]
- US20090026082A1 [0070]
- US20090127589A1 [0070]
- US20100137143A1 [0070]
- US20100282617A1 [0070]
- US7001792B [0071]
- US6890741B [0072]
- US6913884B [0072]
- US6355431B [0072]
- US20050053980A1 [0072]
- US20090186349A1 [0072]
- US20050181440A1 [0072]
- US20100111768A1 [0073]
- US13273666B [0073] [0073]
- US67031812A [0081]
- WO2010115122A [0089] [0090] [0101]
- US5604097A [0105]
- WO2006064199A [0109]

Non-patent literature cited in the description

- **ADEY et al.**disclose an ultra-low-input, tagmentation-based whole-genome bisulfite sequencing methodGenome Research, 2012, vol. 22, 1139-1143 [0006]
- **BENTLEY et al.**Nature, 2008, vol. 456, 53-59 [0038] [0067]
- **MANIATIS et al.**Molecular Cloning: A Laboratory Manual19890000 [0045]
- Short Protocols in Molecular Biology [0045]
- **LIZARDI et al.**Nat. Genet., 1998, vol. 19, 225-232 [0062]
- **DEAN et al.**Proc. Natl. Acad. Sci. USA, 2002, vol. 99, 5261-66 [0063]
- **WALKER et al.**Molecular Methods for Virus DetectionAcademic Press, Inc.19950000 [0063]
- **WALKER et al.**Nucl. Acids Res., 1992, vol. 20, 1691-96 [0063]
- **LAGE et al.**Genome Research, 2003, vol. 13, 294-307 [0063]
- **GROTHUES**Nucleic Acids Res., 1993, vol. 21, 51321-2 [0064]

- **RONAGHI et al.**Analytical Biochemistry, 1996, vol. 242, 184-9 [0068]
- **RONAGHI**Genome Res., 2001, vol. 11, 13-11 [0068]
- **RONAGHI et al.**Science, 1998, vol. 281, 5375363- [0068]
- **LEVENE et al.**Science, 2003, vol. 299, 682-686 [0069]
- **LUNDQUIST et al.**Opt. Lett., 2008, vol. 33, 1026-1028 [0069]
- **KORLACH et al.**Proc. Natl. 'Acad. Sci. USA, 2008, vol. 105, 1176-1181 [0069]
- **DEAMER et al.**Trends Biotechnol., 2000, vol. 18, 147-151 [0071]
- **DEAMER et al.**Acc. Chem. Res., 2002, vol. 35, 817-825 [0071]
- **LI et al.**Nat. Mater., 2003, vol. 2, 611-615 [0071]
- **SONI et al.**Clin. Chem., 2007, vol. 53, 1996-2001 [0071]
- **HEALY**Nanomed., 2007, vol. 2, 459-481 [0071]
- **COCKROFT et al.**J. Am. Chem. Soc., 2008, vol. 130, 818-820 [0071]
- **WU et al.**Anal Chem., 2011, vol. 4, 1321-7 [0091]
- **NITIN et al.**Ann Biomed Eng., 2009, vol. 37, 2018-2027 [0091]

PATENTKRAV

1.   Fremgangsmåde til præparering af et immobiliseret bibliotek af mærkede DNA-fragmenter, hvilken fremgangsmåde omfatter:

(a)   tilvejebringelse af en fast bærer med en flerhed af transposomkomplekser immobiliseret derpå, hvor transposomkomplekserne hver især omfatter:

en transposase, der er ikke-kovalent bundet til et første polynukleotid i en dobbeltstrenget nukleinsyre, der omfatter det første og et andet polynukleotid, hvor det første polynukleotid er immobiliseret til den faste bærer, og hvor det første polynukleotid omfatter

(i)   en 3'-del, der omfatter en transposonendesekvens, og
(ii)   et første mærke, der omfatter et første mærkedomæne;

og hvor det andet polynukleotid omfatter en region, der er komplementær til transposonendesekvensen;
og hvor transposomkomplekserne er homodimerer, idet homodimererne omfatter en første flerhed af homodimerer, der omfatter et første mærke for en første sekvens, og en anden flerhed af homodimerer, der omfatter et første mærke for en anden sekvens;
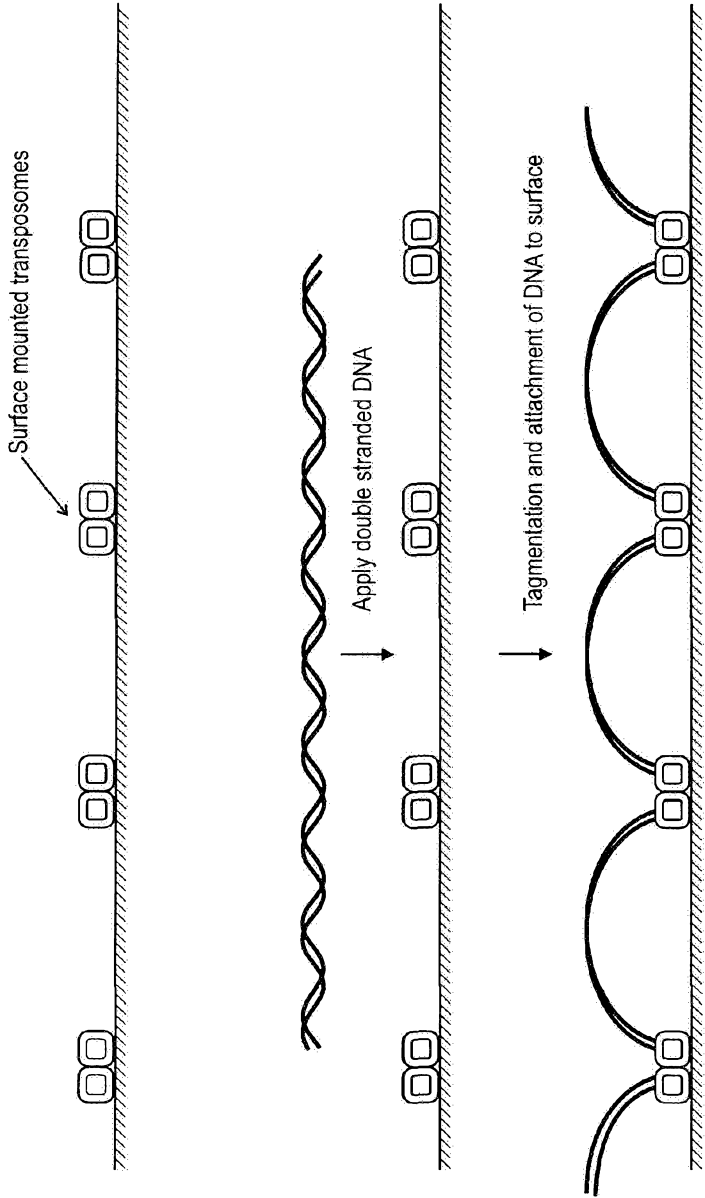
(b)   påføring af et dobbeltstrenget mål-DNA på den faste bærer under betingelser, hvorved mål-DNA'et fragmenteres af transposomkomplekserne, og 3'-transposonendesekvensen for det første polynukleotid overføres til en 5'-ende af mindst én streng af fragmenterne; hvorved der fremstilles et immobiliseret bibliotek af dobbeltstrengede fragmenter, hvor mindst én streng er 5'-mærket med det første mærke for den første sekvens, og mindst én streng er 5'-mærket med det første mærke for den anden sekvens.

2.   Fremgangsmåde ifølge krav 1, hvor en 3'-transposonendesekvens for et første polynukleotid overføres til 5'-enderne af begge strenge af det fragmenterede dobbeltstrengede mål-DNA.

3.   Fremgangsmåde ifølge krav 1 eller 2, der endvidere omfatter vask af den faste bærer med henblik på at fjerne eventuelle ubundne nukleinsyrer.

5   **4.**   Fremgangsmåde ifølge et hvilket som helst af kravene 1-3, hvor transposomkomplekserne forekommer på den faste bærer i en densitet på mindst 10^3, 10^4, 10^5, 10^6 komplekser pr. mm^2, eller hvor transposomkomplekset omfatter en hyperaktiv Tn5-transposase.

10   **5.**   Fremgangsmåde ifølge et hvilket som helst af kravene 1-4, hvor længderne af de dobbeltstrengede fragmenter i det immobiliserede bibliotek tilpasses ved at forøge eller formindske densiteten for transposomkomplekser, der forekommer på den faste bærer.

15   **6.**   Fremgangsmåde ifølge et hvilket som helst af kravene 1-5, hvor mindst 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % af de mærker, der forekommer på den faste bærer, omfatter det samme mærkedomæne.

20   **7.**   Fremgangsmåde ifølge et hvilket som helst af kravene 1-6, der endvidere omfatter, at:
(c)   der tilvejebringes transposomkomplekser i opløsning, og transposomkomplekserne i opløsning bringes i kontakt med de immobiliserede fragmenter under betingelser, hvorved mål-DNA'et fragmenteres yderligere med
25   transposomkomplekserne i opløsning; hvorved der opnås immobiliserede nukleinsyrefragmenter, der har den ene ende i opløsning.

8.   Fremgangsmåde ifølge krav 7, hvor transposomkomplekserne i opløsning omfatter et andet mærke, hvorved der genereres immobiliserede
30   nukleinsyrefragmenter, der har et andet mærke i opløsning.

9.   Fremgangsmåde ifølge krav 8, hvor det første og det andet mærke er forskelligt.

10. Fremgangsmåde ifølge et hvilket som helst af kravene 7-9, hvor mindst 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % af transposomkomplekserne i opløsning omfatter et andet mærke, der omfatter et andet mærkedomæne.

11. Fremgangsmåde ifølge krav 8, der endvidere omfatter amplificering af fragmenterne på den faste overflade ved tilvejebringelse af en polymerase og en amplificeringsprimer svarende til en del af det første polynukleotid.

12. Fremgangsmåde ifølge krav 1, hvor mærkedomænet omfatter en region til clusteramplificering eller en region til priming af en sekventeringsreaktion.

13. Fremgangsmåde ifølge krav 1, hvor den faste bærer omfatter mikropartikler eller en mønstret overflade eller brønde.

14. Fast bærer med et bibliotek af mærkede DNA-fragmenter, der er immobiliseret derpå, præpareret ifølge fremgangsmåden ifølge et hvilket som helst af kravene 1-13.

15. Fast bærer med transposomkomplekser immobiliseret derpå, idet hvert transposomkompleks omfatter en transposase, der ikke-kovalent er bundet til et første polynukleotid i en dobbeltstrenget nukleinsyre, der omfatter det første og et andet polynukleotid, hvor det første polynukleotid er immobiliseret til den faste bærer, og hvor det første polynukleotid omfatter

(i) en 3'-del, der omfatter en transposonendesekvens, og

(ii) et første mærke, der omfatter et første mærkedomæne;

og hvor det andet polynukleotid omfatter en region, der er komplementær til transposonendesekvensen;

og hvor transposomkomplekserne er homodimerer, idet homodimererne omfatter en første flerhed af homodimerer, der omfatter et første polynukleotid for en første sekvens, og en anden flerhed af homodimerer, der omfatter et første polynukleotid for en anden sekvens.

DRAWINGS

Fig. 1a)

Fig. 1b)

Fig. 2a)

Fig. 2b)

Apply double stranded DNA

Tagment

Make clusters

Fig. 3a)

Fig. 3b)
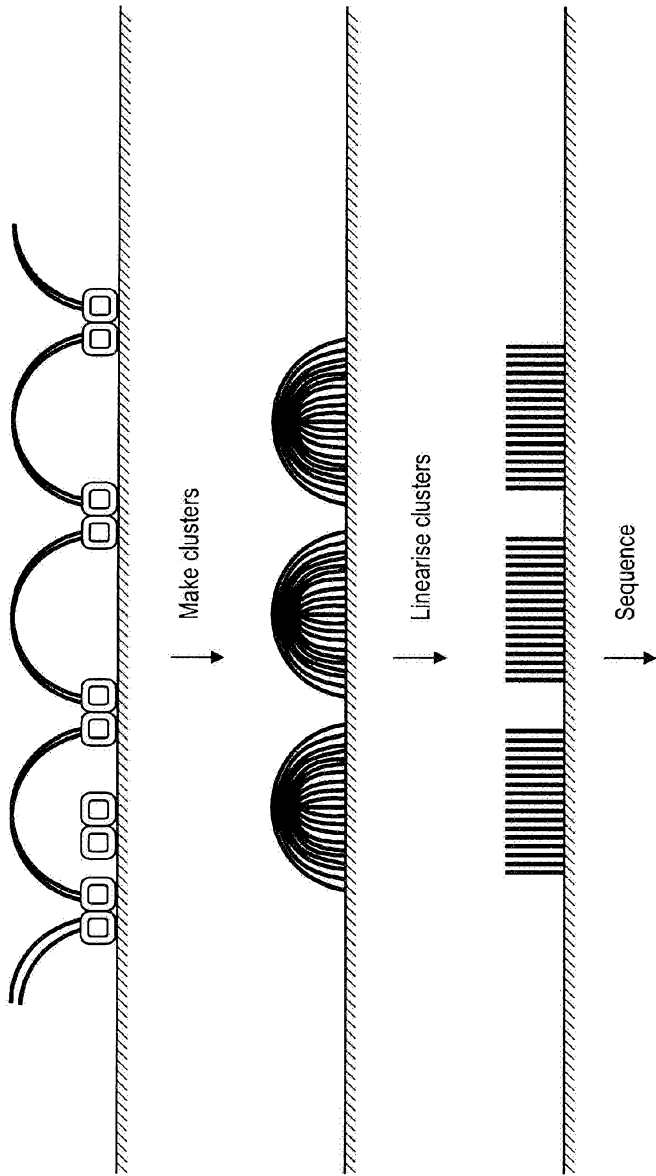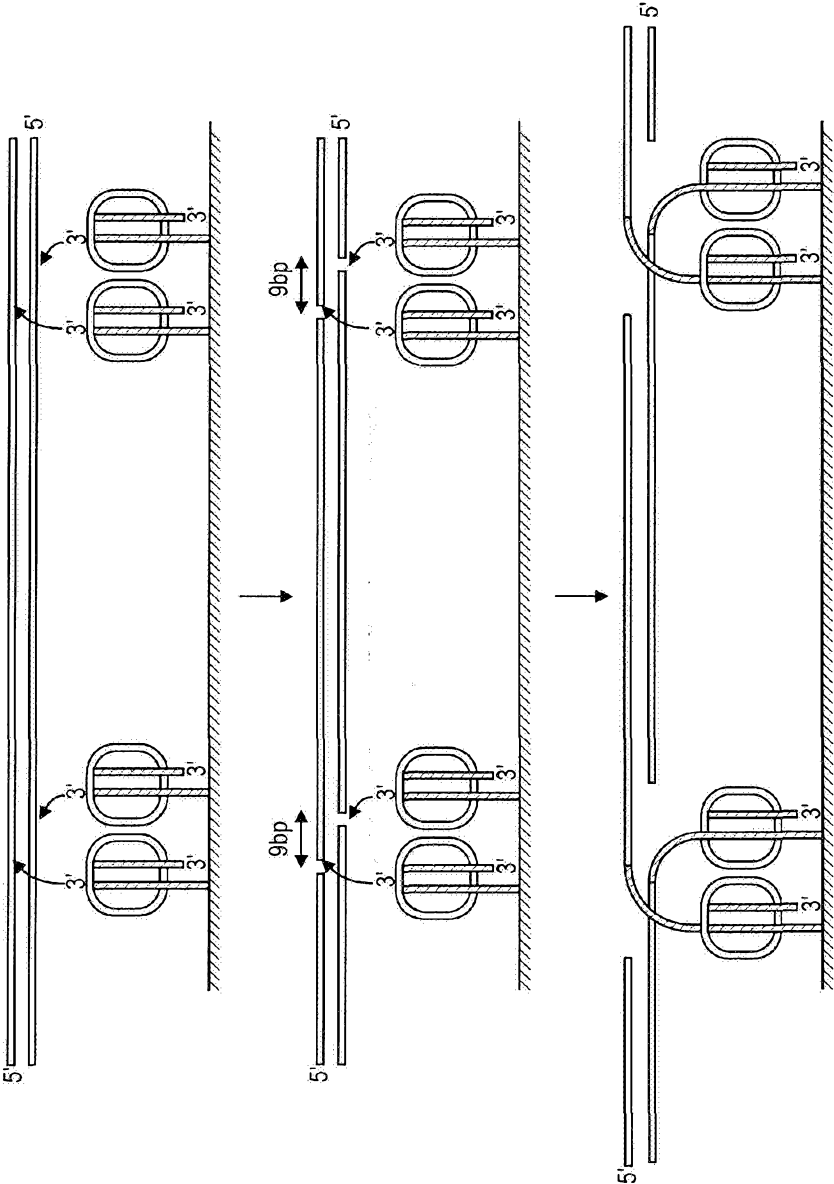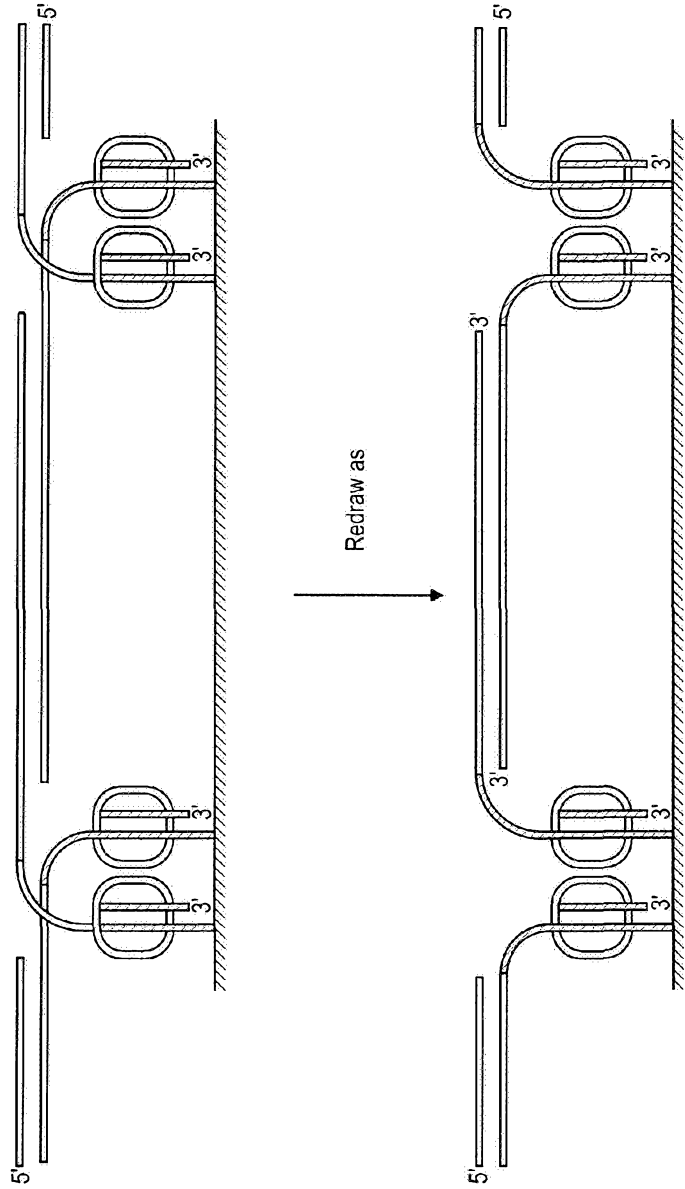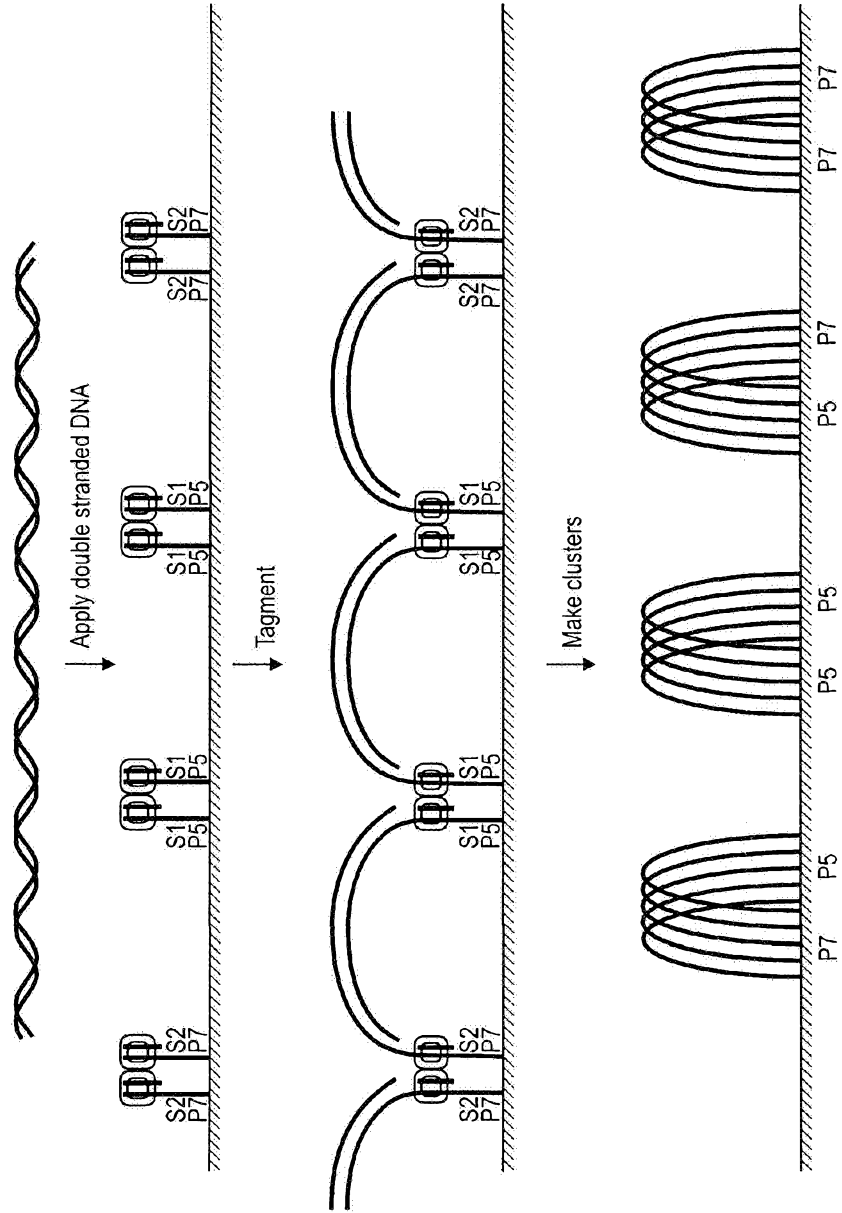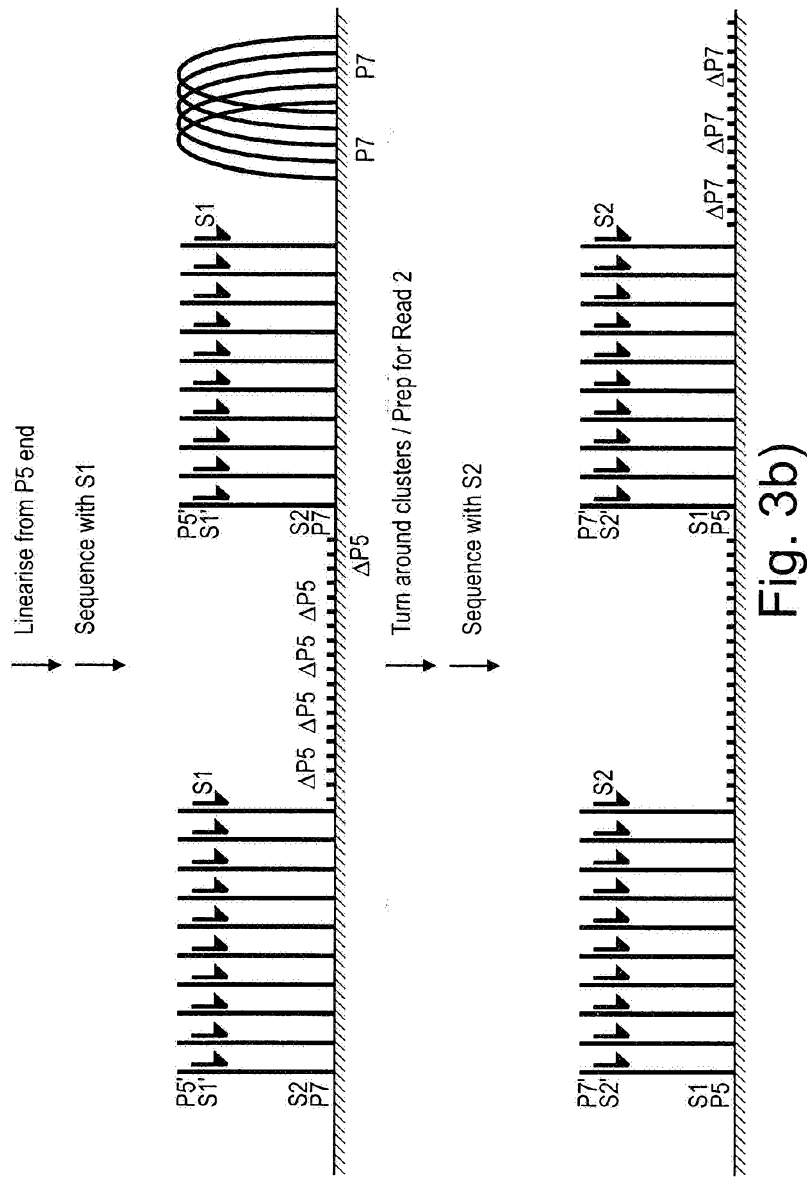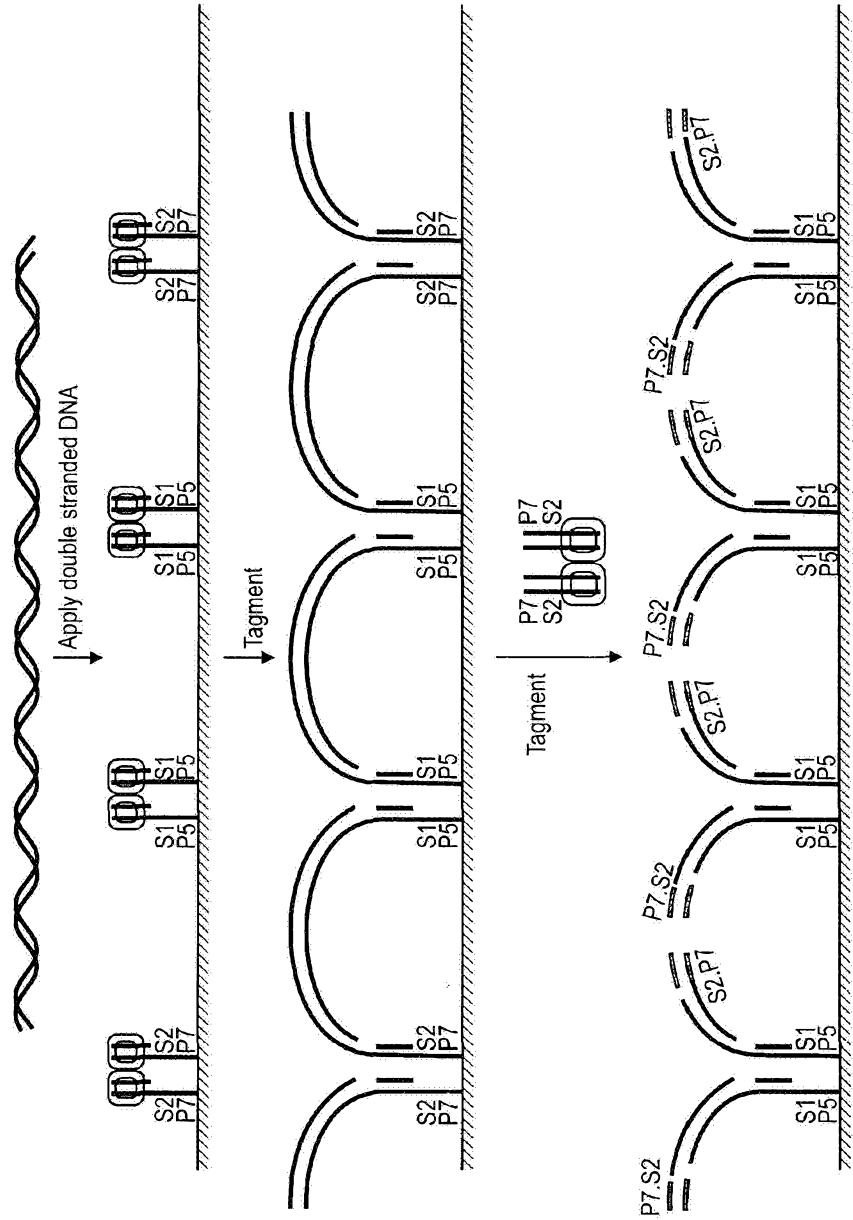
Fig. 4a)
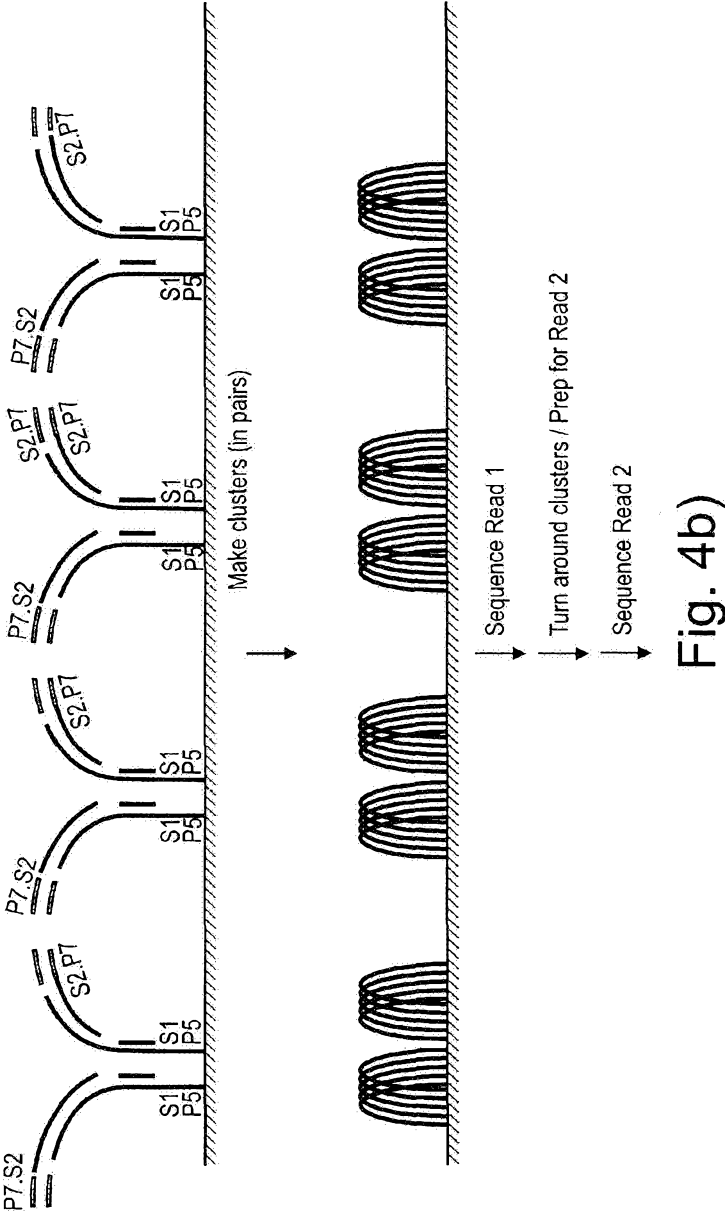
Fig. 4b)
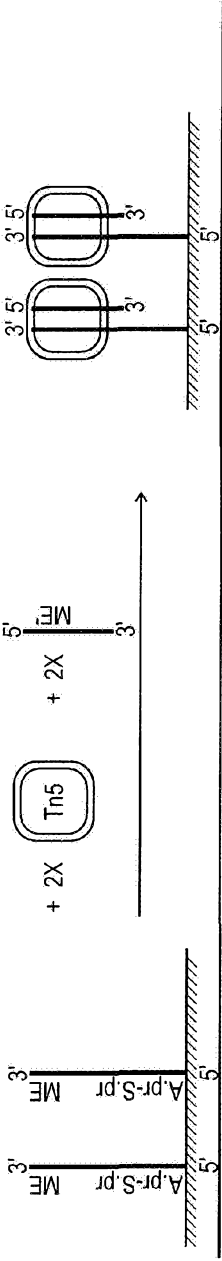
Fig. 5a)

Fig. 5b)

Fig. 5c)
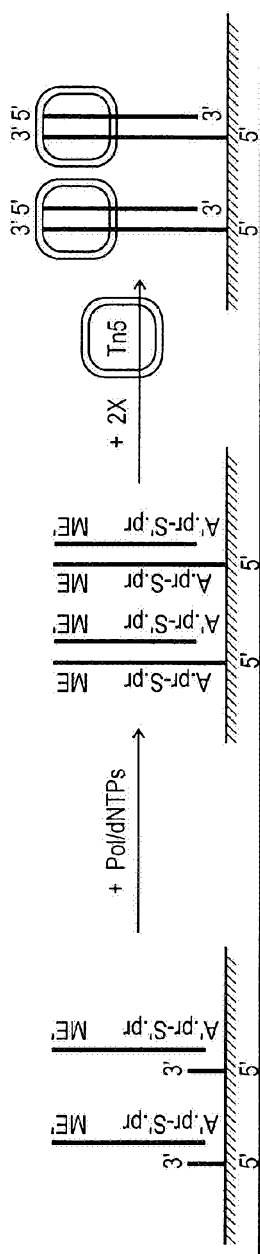
Fig. 5d)
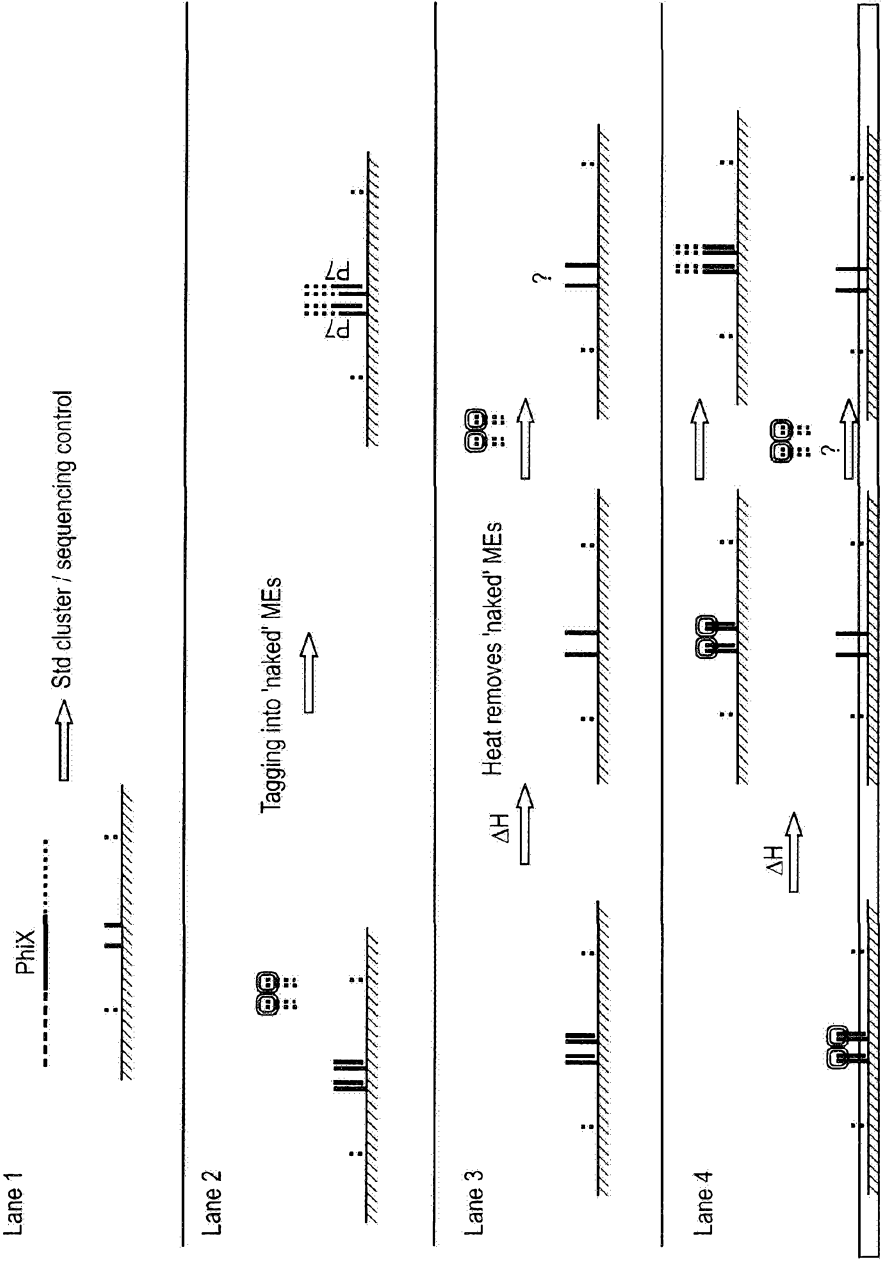
Fig. 6a
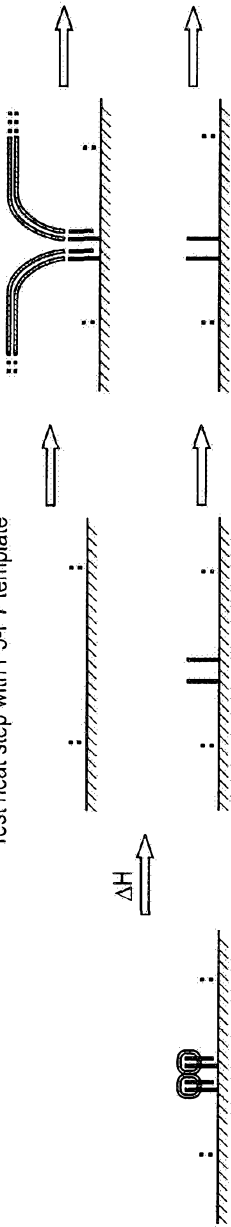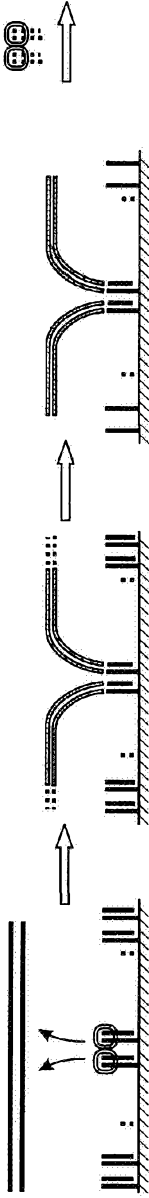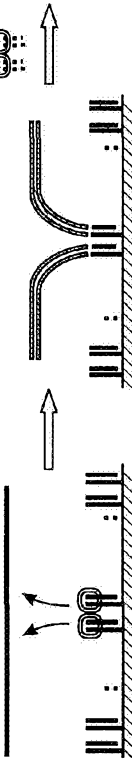
Fig. 6b)

(Lane 6 condition)
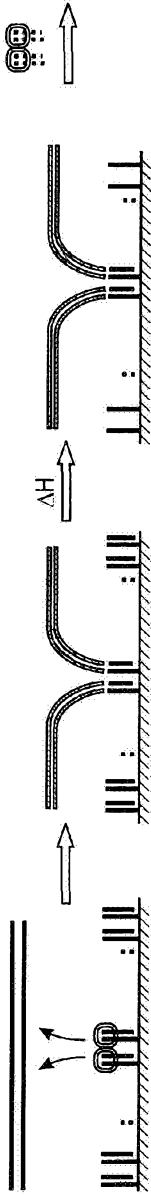
ΔH

Read 1
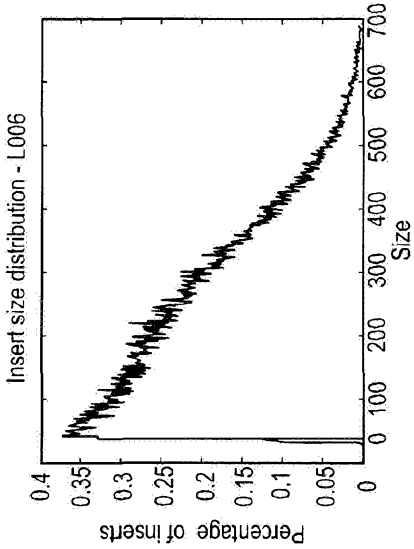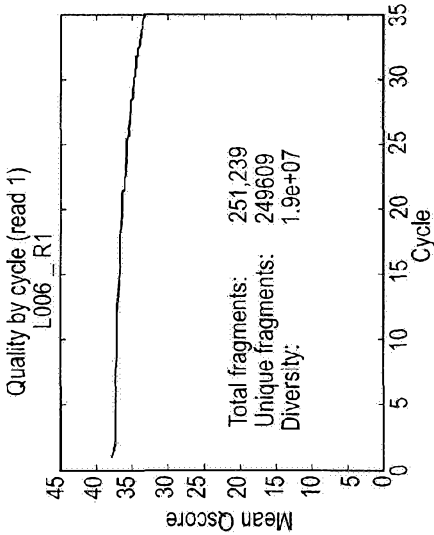
	Sample Yield (Mbases)	Clusters (raw)	Clusters (PF)	1st Cycle Int (PF)	% intensity after 20 cycles (PF)	% PF Clusters	% Align (PF)	% Mismatch Rate (PF)	% >=Q30 bases (PF)	Mean Quality SCore (PF)
lane1	87	2,640,647	2,494,533	335	80.44	94.47	98.29	0.11	99.02	39.38
lane2	142	4,669,765	4,049,565	199	78.47	86.72	0		95.18	37.56
lane3	25	1,012,657	728,288	236	76.56	71.92	0		87.69	35.49
lane4	13	1,660,050	368,331	38	88.69	22.19	0		66.92	28.75
lane5	0	1,651,305	2,790	104	89.81	0.17	4.23	0.97	33.3	18
lane6	51	1,992,375	1,457,197	134	83.8	73.14	74.69	0.35	90.27	36.1
lane7	89	3,508,504	2,545,003	128	80.46	72.54	27.62	0.6	90.51	36.09
lane8	0	1,191,904	6,629	107	73.09	0.56	0		11.23	7.43

Read 2

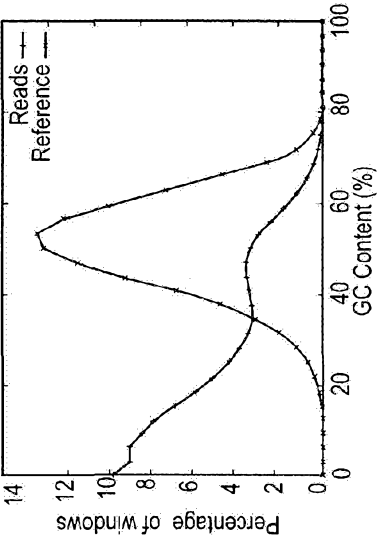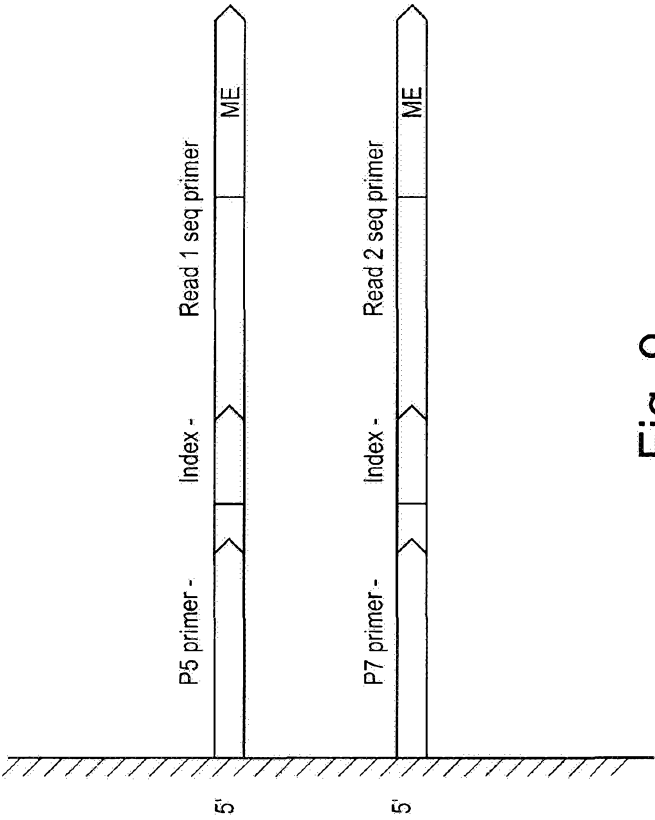|	Sample Yield (Mbases)	Clusters (raw)	Clusters (PF)	1st Cycle Int (PF)	% intensity after 20 cycles (PF)	% PF Clusters	% Align (PF)	% Mismatch Rate (PF)	% >=Q30 bases (PF)	Mean Quality SCore (PF)
lane1	87	2,640,647	2,494,533	96	80.99	94.47	96.86	0.19	95.71	37.93
lane2	142	4,669,765	4,049,565	15	105.29	86.72	0		1.97	2.96
lane3	25	1,012,657	728,288	18	105.22	71.92	0		2.38	3.16
lane4	13	1,660,050	368,331	13	92.27	22.19	0		2.18	3.18
lane5	0	1,651,305	2,790	37	18.39	0.17	0.43	1.43	5.51	5.01
lane6	51	1,992,375	1,457,197	32	119.21	73.14	17.45	2.26	8.55	7.34
lane7	89	3,508,504	2,545,003	26	115.87	72.54	4.96	3.31	6.08	5.77
lane8	0	1,191,904	6,629	81	72.94	0.56	0		5.83	5.40

Fig. 7
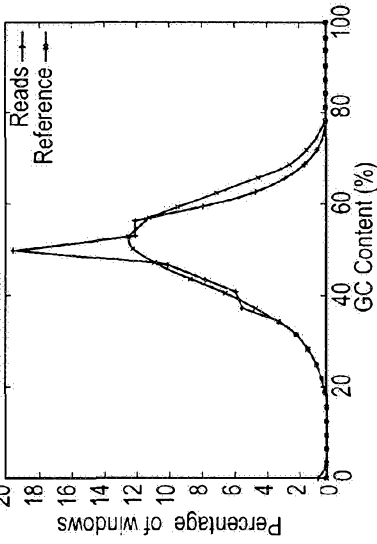
Fig. 8

Fig. 9
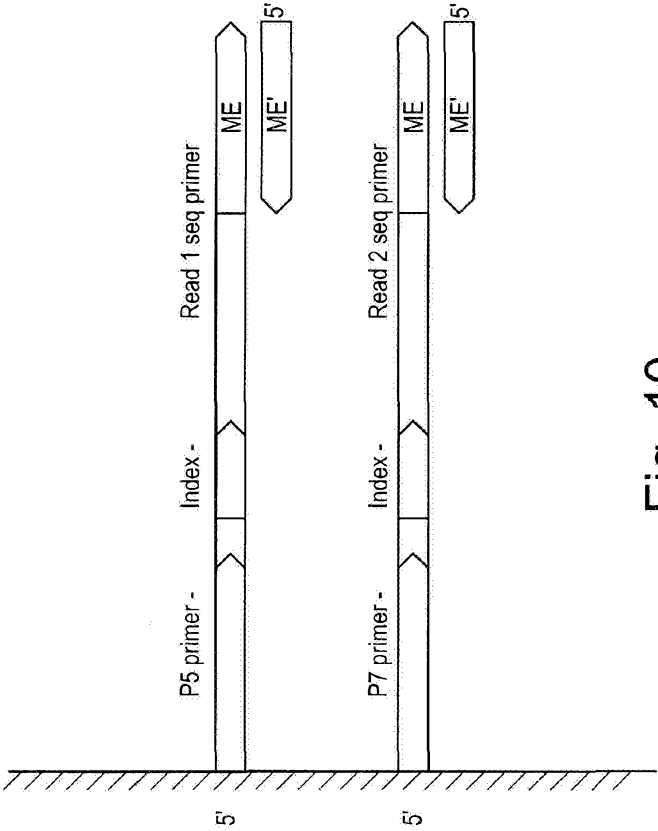
Fig. 10

Fig. 11

Fig. 12

Transposase dimer

MENTS
5' 5'
3' MENTS
Barcode
Primer

MENTS
5' 5'
3' MENTS
Barcode
Primer

5'   5'

*Examples of Primer sequence composition:

Launch pad seq for step-out PCR

Optional cleavage point

P5 - Read 1 sequencing primer

5' X

5'

Fig. 13

Fig. 14

Transfer barcoded nextera sub-fragments off beads
(eg by cleavage or PCR)

Pool and sequence
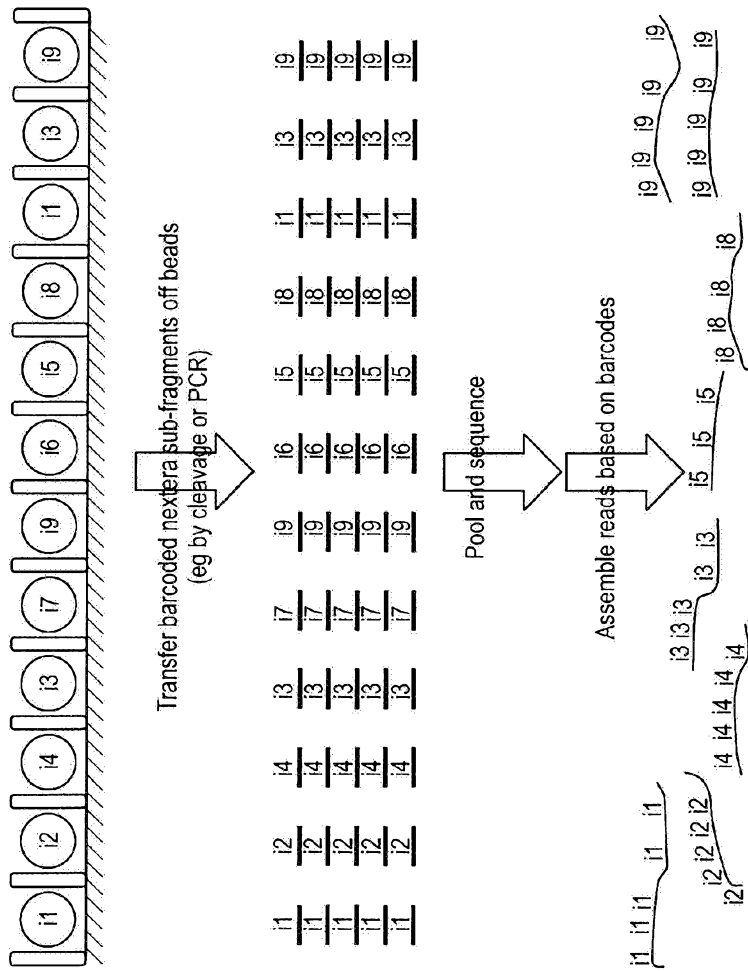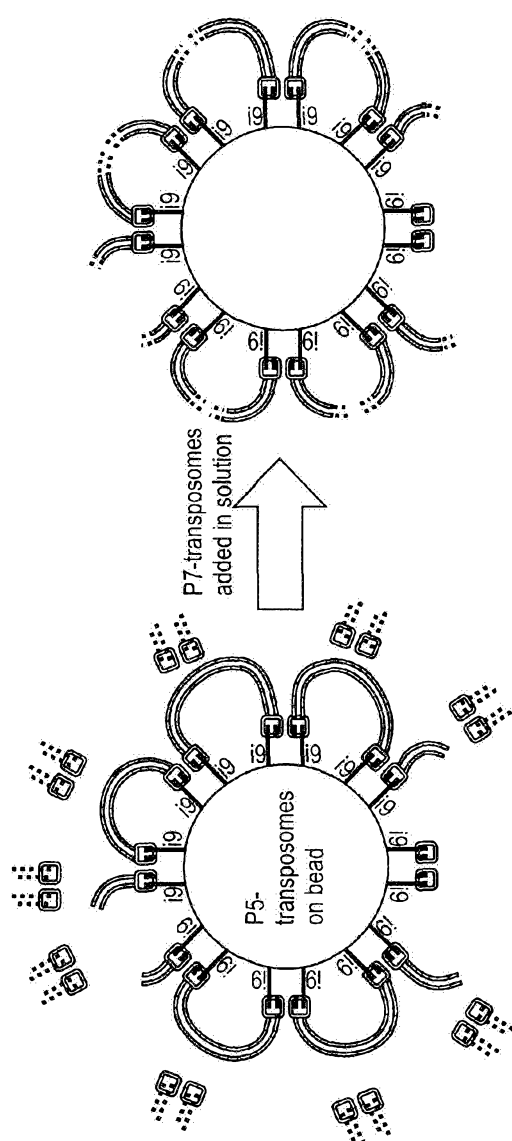
Assemble reads based on barcodes
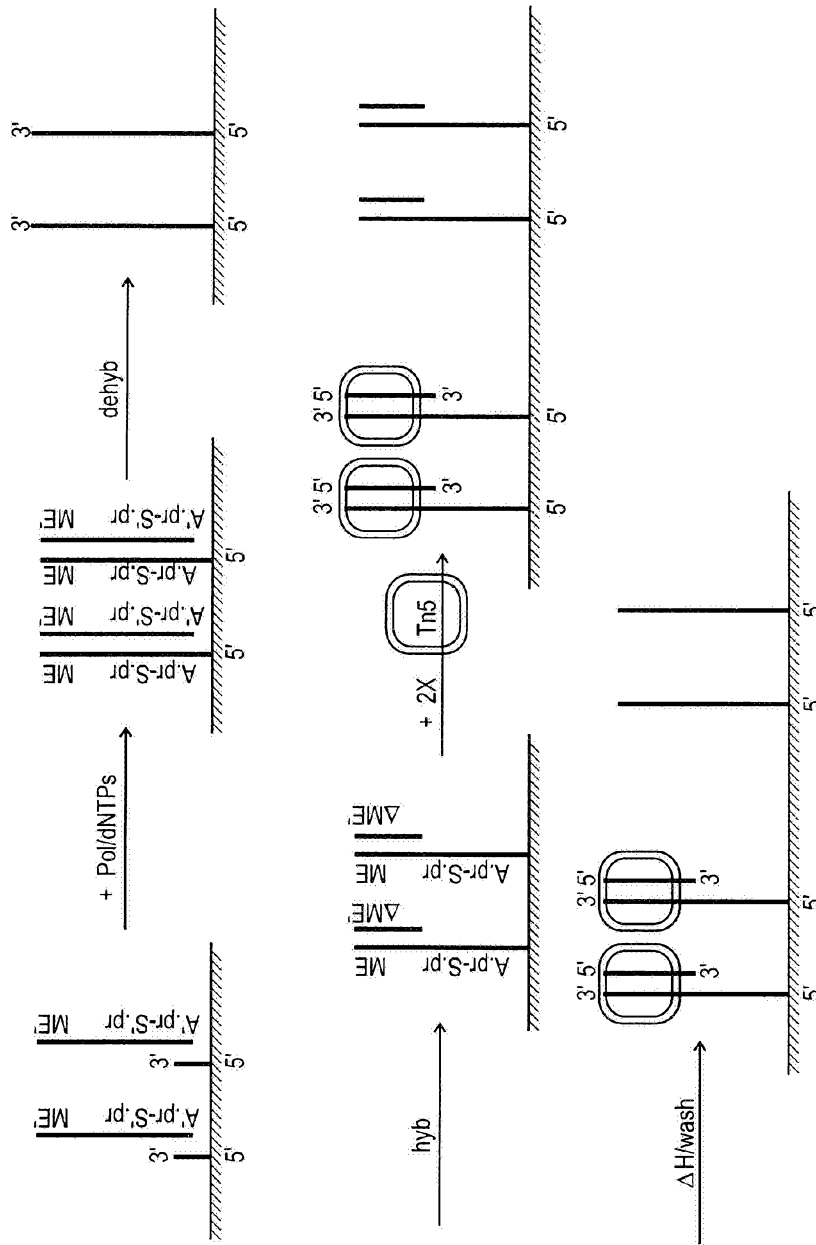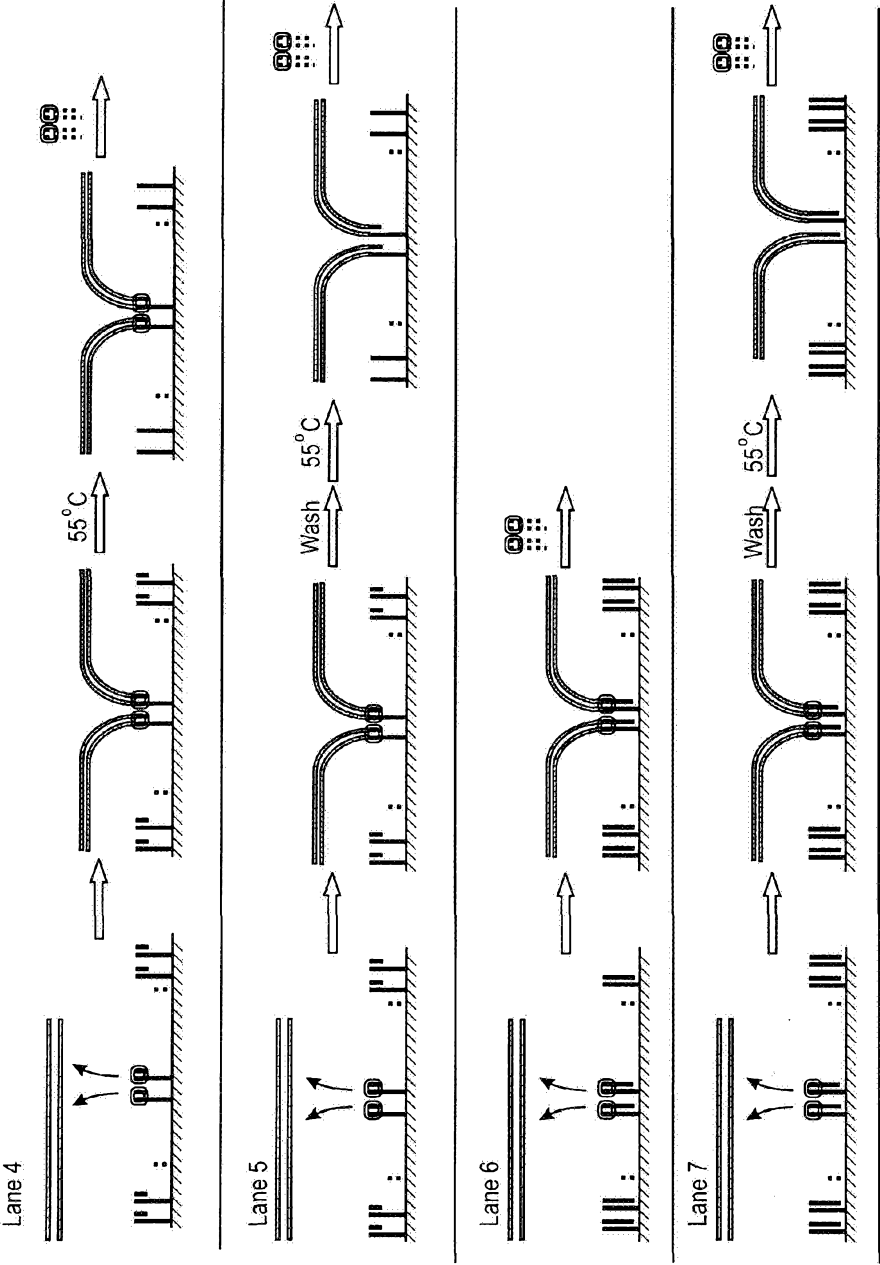
Fig. 15
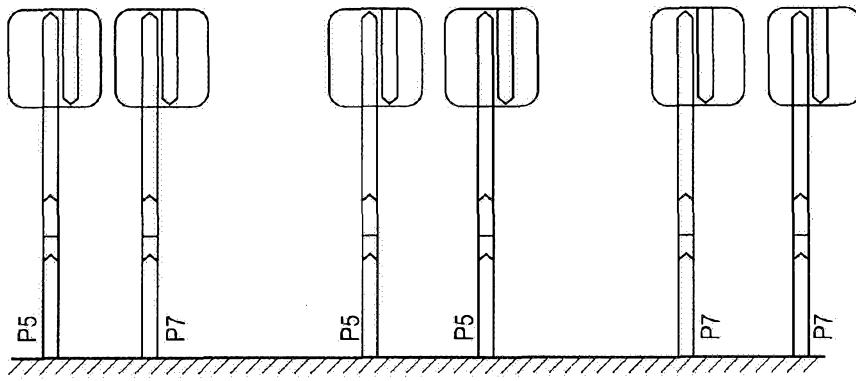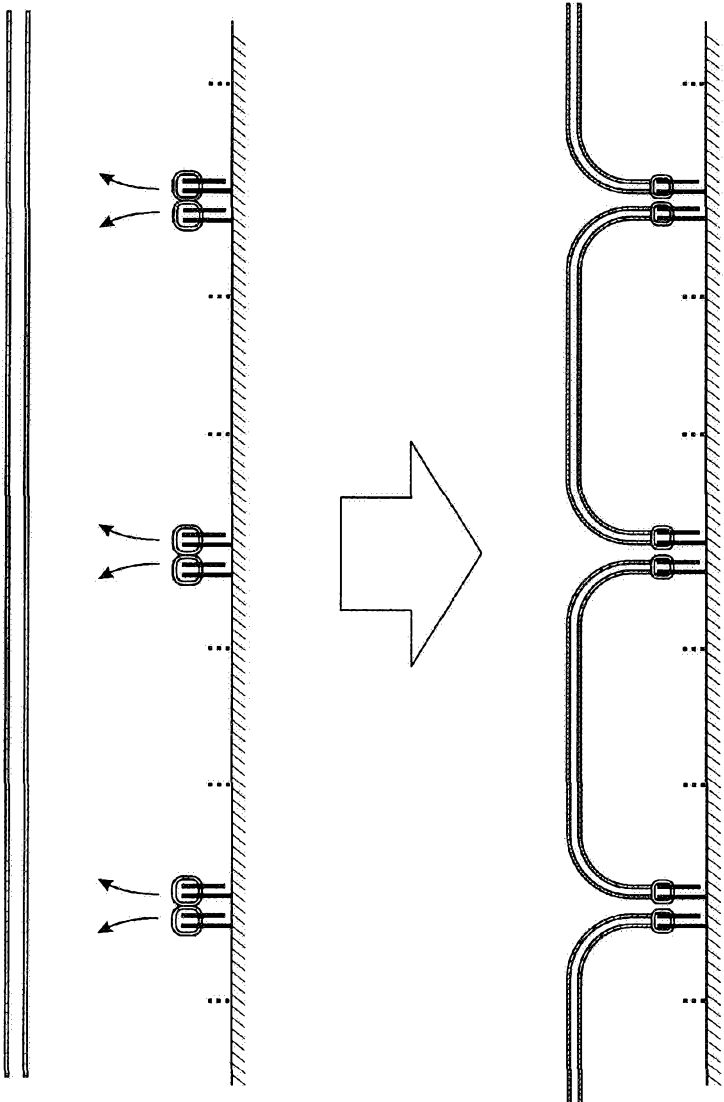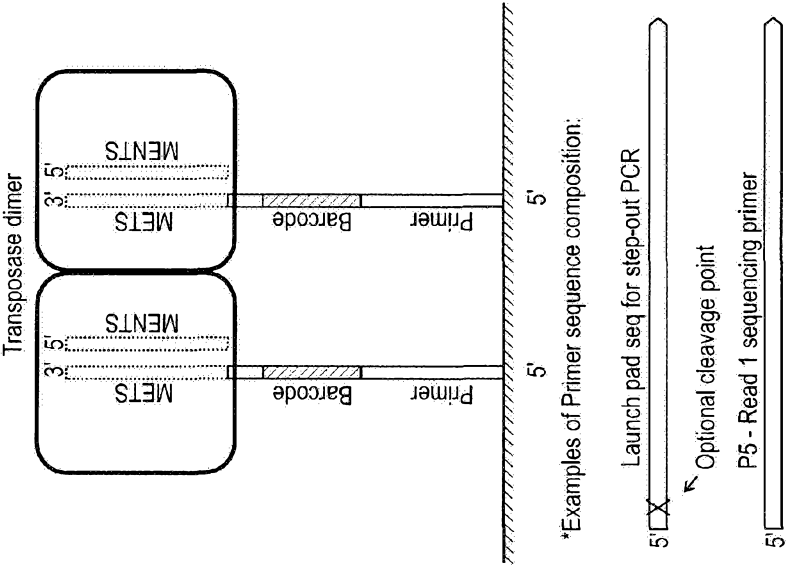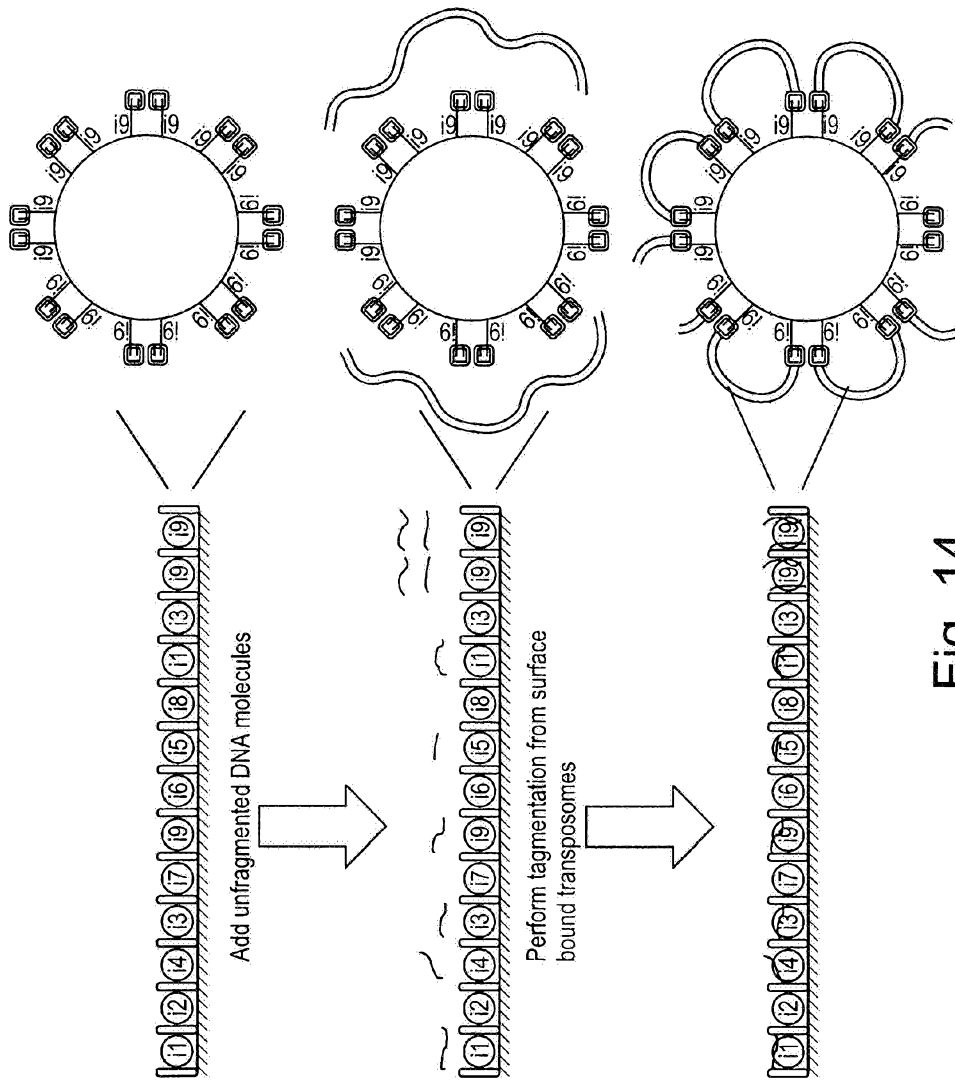
Fig. 16

Fig. 17

Fig. 18a

	Yield (Mb)	Clusters (raw)	Clusters (PF)	1st Cycle Int (PF)	% intensity after 20 cycles (PF)	% PF Clusters	% Align (PF)	% Mismatch Rate (PF)	% >=Q30 bases (PF)	Mean Quality Score (PF)
lane1	204	6,172,822	5,821,910	226	84.42	94.32	98.24	0.04	98.75	39.16
lane2	178	5,676,854	5,096,648	170	84.02	89.78	93.81	0.08	97.78	38.74
lane3	261	8,915,997	7,450,306	168	82.73	83.56	89.9	0.13	95.42	37.7
lane4	89	2,963,760	2,542,838	120	85.56	85.8	96.47	0.11	95.99	38.14
lane5	96	3,223,900	2,728,667	103	83.66	84.64	96.41	0.12	95.22	37.82
lane6	109	3,801,294	3,103,977	105	85.96	81.66	22.59	0.11	95.21	37.69
lane7	123	4,239,062	3,513,343	112	85.31	82.88	15.52	0.17	95.64	37.82
lane8	0	2,230,575	0	0	0	0				

Read 1

Fig. 18b